Reduced Efficacy of a Src Kinase Inhibitor in Crowded Protein Solution

Kento Kasahara
RIKEN Center for Biosystems Dynamics Research
Suyong Re
RIKEN Center for Biosystems Dynamics Research
Grzegorz Nawrocki
Michigan State University
Hiraku Oshima
Chiemi Mishima-Tsumagari
RIKEN Centre for Life Science Technologies
Mutsuko Kukimoto-Niino
RIKEN Systems and Structural Biology Centre
Isseki Yu
Maebashi Institute of Technology
Mikako Shirouzu
RIKEN Center for Biosystems Dynamics Research https://orcid.org/0000-0002-7997-2149
Michael Feig
Michigan State University
Yuji Sugita (sugita@riken.jp)
RIKEN https://orcid.org/0000-0001-9738-9216

Article

Keywords: Inhibitor Binding, Atomistic Molecular Dynamics, Crowded Cellular Environment

Posted Date: September 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-72438/v1

Version of Record: A version of this preprint was published at Nature Communications on July 2nd, 2021.
See the published version at https://doi.org/10.1038/s41467-021-24349-5.

Reduced Efficacy of a Src Kinase Inhibitor in Crowded Protein Solution

Author Information

Kento Kasahara[a], Suyong Re[a], Grzegorz Nawrocki[b], Hiraku Oshima[a], Chiemi Mishima-Tsumagari[c], Mutsuko Kukimoto-Niino[c], Isseki Yu[d], Mikako Shirouzu[c], Michael Feig[b], and Yuji Sugita[a,e,f]

[a]Laboratory for Biomolecular Function Simulation, RIKEN Center for Biosystems Dynamics Research, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan.

[b]Department of Biochemistry and Molecular Biology, Michigan State University, 630 Wilson Road, 218 Biochemistry Building, East Lansing, Michigan 48824-1319, United States.

[c]Laboratory for Protein Functional and Structural Biology, RIKEN Center for Biosystems Dynamics Research, 1-7-22 Suehiro-cho, Tsurumi, Yokohama, Kanagawa 230-0045, Japan.

[d]Department of Life Science and Informatics, Maebashi Institute of Technology, Kamisadori-machi, Maebashi, Gunma 371-0816, Japan.

[e]Computational Biophysics Research Team, RIKEN Center for Computational Science, 6-7-1 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan.

[f]Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Hirosawa 2-1, Wako, Saitama 351-0198, Japan.

Contributions

K.K., G.N., and H.O. performed simulations. K.K. analyzed data. K.K., S.R., and I.Y. modelled simulation systems. C.M.T., M.K.N., and M.S. carried out experiments. K.K., S.R., M.F., and Y.S. wrote manuscript. M.F. and Y.S. designed and initiated the research.

Corresponding author

Correspondence to: Michael Feig (mfeiglab@gmail.com), Yuji Sugita (sugita@riken.jp)

ORCID

Kento Kasahara (https://orcid.org/0000-0001-7138-250X)

Suyong Re (https://orcid.org/0000-0002-3752-6554)

Hiraku Oshima (https://orcid.org/0000-0001-5626-1291)

Mutsuko Kukimoto-Niino (https://orcid.org/0000-0002-4178-1611)

Michael Feig (https://orcid.org/0000-0001-9380-6422)

Yuji Sugita (https://orcid.org/0000-0001-9738-9216)

Abstract

The inside of a cell is highly crowded with proteins and other biomolecules. How proteins express their specific functions together with many off-target proteins in crowded cellular environments is largely unknown. Here, we investigate an inhibitor binding with c-Src kinase using atomistic molecular dynamics (MD) simulations in dilute as well as crowded protein solution. The populations of the inhibitor, PP1, in bulk solution and on the surface of c-Src kinase are reduced as the concentration of crowder bovine serum albumins (BSAs) increases. This observation is consistent with the reduced PP1 inhibitor efficacy in experimental c-Src kinase assays in addition with BSAs. The crowded environment changes the major binding pathway of PP1 toward c-Src kinase compared to that in dilute solution. This change is explained based on the population shift mechanism of local conformations near the inhibitor binding site in c-Src kinase. Protein functions in a living cell could be examined using atomistic MD simulations with realistic cellular environments.

Introduction

Understanding of protein functions in a living cell is one of the essential issues in molecular biology and biochemistry. *In vitro* experiments on protein or enzyme functions often reveal different results from *in vivo* experiments. In a typical cytoplasm, 25-45% of the total volume is occupied by macromolecules including many proteins, nucleic acids, metabolites, osmolytes, and ions[1,2]. The crowded environment affects protein structure, stability, dynamics, and biological functions slightly or significantly[3,4]. Recent experimental[5–9] and theoretical studies[10–16] have shown that weak and non-specific molecular interactions are equally important with the excluded volume effect that has been emphasized in the traditional theory on macromolecular crowding[2,4]. For instance, in-cell NMR spectroscopy has shown reduced conformational stability of ubiquitin in Hela cells compared to those in dilute solution. The traditional theory, however, suggested that compact native structures of proteins are preferred in crowded cellular environments[7].

Substrate (or ligand) binding to a target protein is an essential process that controls a variety of biological functions including enzymatic reactions, substrate transports, signal transductions, and so on. The well-established concepts like "lock-and-key", "induced-fit", or "conformational selection" have been used to interpret the binding processes[17] as well as to design new drug candidates[18,19]. Protein conformational flexibility has been regarded as more important in substrate bindings so that atomistic molecular dynamics (MD) simulations are often conducted in *in-silico* drug discovery[20]. In spite of these efforts, many designed drug candidates lack *in vivo* efficacy and cannot be applied to the patients who are suffering from diseases. One of the missing factors in current drug discovery protocols might be the effect of cellular environments on protein-ligand bindings and functions. There is increasing experimental evidence that supports the importance of cellular environments for the thermodynamics and kinetics of protein-ligand binding[3,21]. For instance, the efficacy of inhibitors of threonine tyrosine kinases in cells is reduced significantly in many cases compared to that in dilute solution[22].

Recently, the accessible time scale of MD simulation has been greatly extended due to the developments of MD-special purpose supercomputers, Anton/Anton2[23,24] and enhanced conformational sampling algorithms including replica exchange methods[25,26], metadynamics[27,28], Gaussian accelerated MD

(GaMD)[29], and scaled MD[30,31]. Parallel MD simulations based on weighted-ensemble method[32–34], parallel cascade selection MD (PaCS-MD)[35], Markov state model (MSM)[36] are also useful to predict thermodynamic and kinetic properties of biomolecules. Using these advanced techniques, absolute binding free energies and/or kinetic parameters, namely, k_{on} or k_{off}, in dilute solution have been successfully predicted. Atomic MD simulations of multiple proteins in crowded environments revealed different aspects of proteins, such as translational or rotational diffusion[13,37], solvation structure and dynamics[12,14], and the conformational stability of proteins[10,12]. The importance of metabolites including adenosine triphosphate (ATP) on protein structures and interactions was suggested in large-scale MD simulations of the cytoplasm in *Mycoplasma genitalium*[15]. Independent experiments also suggested that ATP works as a hydrotrope in cells[38].

Here, we study protein functions in crowded protein solutions by performing multiple μsec MD simulations using Anton2 and other supercomputers (~100 μsec, in total). In the study, we focus on a small inhibitor (PP1) binding to c-Src kinase in dilute solution and crowded environments with bovine serum albumins (BSAs) as protein crowders. BSA has been shown to not change the conformational stability of the other protein in crowded solution[6]. c-Src kinase (Supplementary Fig. 1) regulates various signal transduction pathways by catalyzing phosphate transfer from ATP to a target substrate[39]. Dysregulation of this kinase function is associated with many diseases like cancer, making it an important therapeutic target. To design drug candidates, both the active and inactive conformations of c-Src kinase[40–45] as well as the ligand binding process[43,46–48] have been extensively investigated in dilute solution. PP1 was designed to inhibit the kinase function by binding to the ATP binding site (the canonical binding site) of c-Src kinase (Supplementary Fig. 2)[49]. A recent computational study using the two-dimensional replica-exchange method characterizes the multiple binding poses and binding pathways of PP1 to c-Src kinase from the encounter state[48,50]. We performed atomistic MD simulations of four c-Src kinase/PP1/BSA systems, each of which contains 0 (dilute solution), 2, 4, and 8 BSAs in a simulation box with the same size (referred to as Src2BSA, Src4BSA, and Src8BSA, respectively). The observations in the simulations provide a microscopic picture of protein-

inhibitor binding in crowded solution, deepening our understanding of protein function in a living cell and proposing a new strategy for *in-silico* drug discovery.

Results

Structural stability and flexibility of c-Src kinase

We first examine the effect of crowded environments on the structural stability and flexibility of c-Src kinase. The Cα root mean square deviations (RMSDs) from the X-ray structure are observed around 2-3 Å in all the simulations, indicating that the kinase remains in the active conformation (Supplementary Fig. 3). The time averages of the RMSD values in the presence of BSA (~2.4 Å) are slightly smaller than that in the absence of BSA (~2.6 Å) (Supplementary Fig. 4(a)). e in the inhibitor-unbound state are also not affected by the presence of the BSAs (Supplementary Fig. 4(b)), where the unbound state is defined when the protein-PP1 distance, ξ, is longer than 15 Å (Supplementary Fig. 5).

PP1 distributions in water and on the protein surfaces

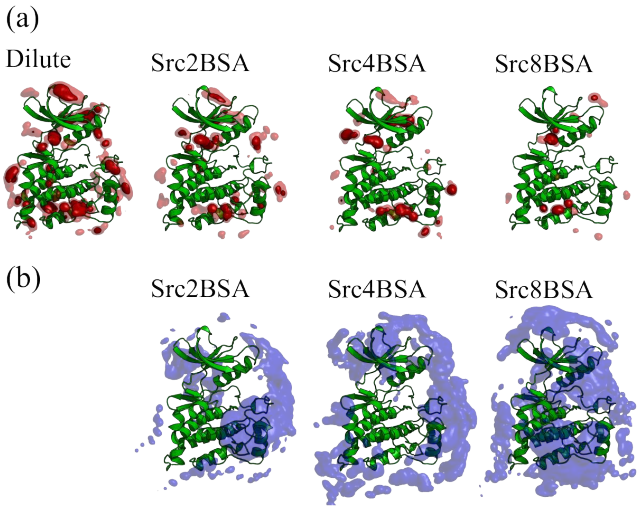


Fig. 1: Spatial distribution functions (SDF) of (a) PP1 and (b) BSA around c-Src kinase. In the case of BSA-SDF, Cα atoms are used for analysis. The PP1-SDF is shown as isosurface at 0.5% (transparent) and 1.5% occupancies (solid). For BSA-SDF, the isosurface at 1.0% occupancy is shown.

The interaction between c-Src kinase and PP1 was examined based on the spatial distribution function (SDF) of PP1 around the kinase. The SDF in the absence of BSA illustrates the existence of many non-canonical binding sites beside the canonical ATP-binding one, which is consistent with previous findings in long MD simulations using Anton[51] (Fig. 1(a) and Supplementary Fig. 6). The PP1 non-canonical binding sites eventually disappear as the BSA concentrations increase. The surface region of c-Src kinase in the crowded solutions is largely occupied by BSAs (Fig. 1(b) and Supplementary Fig. 7). Since the canonical binding site exists at the cleft formed between the N- and C-lobes, BSA cannot cover it. A lot of PP1 stays on the surface of BSA in the crowded solutions, due to weak and non-specific interactions between PP1 and BSAs (Supplementary Fig. 8). The reduced PP1 distribution around c-Src kinase in the crowded solutions is, thus, explained via both the steric hinderance of BSAs (Supplementary Fig. 7) and the weak and non-specific interactions between PP1 and BSAs (Supplementary Fig. 8).

We also investigated the population of PP1 in the bulk water region, which is at least 5 Å away from any heavy atom of c-Src kinase or BSAs (Supplementary Fig. 9). The space within 5 Å of any heavy atom of c-Src kinase (or BSA) is defined as the surface region of c-Src kinase (or BSA). In dilute solution, 79.9% of PP1 exist in the water phase, while the probability dropped to 1.8% in Src8BSA (Fig. 2 (a)). The PP1 population on the surface of c-Src kinase also decreased from 20.1% (dilute) to 8.3% (Src8BSA) as the protein volume fraction increased. Contrarily, the population on the BSA surfaces increased from 55.3% (Src2BSA) to 89.9% (Src8BSA). If we divide the probability by the number of BSAs in a simulation box, the values for Src2BSA, Src4BSA, and Src8BSA become 27.7, 20.0, and 11.2%, respectively, which are

still larger than those on the c-Src kinase surface at the same protein concentration (Supplementary Tables 4 and 5, Supplementary Fig. 10).

In vitro inhibitory assays in the presence or absence of BSA

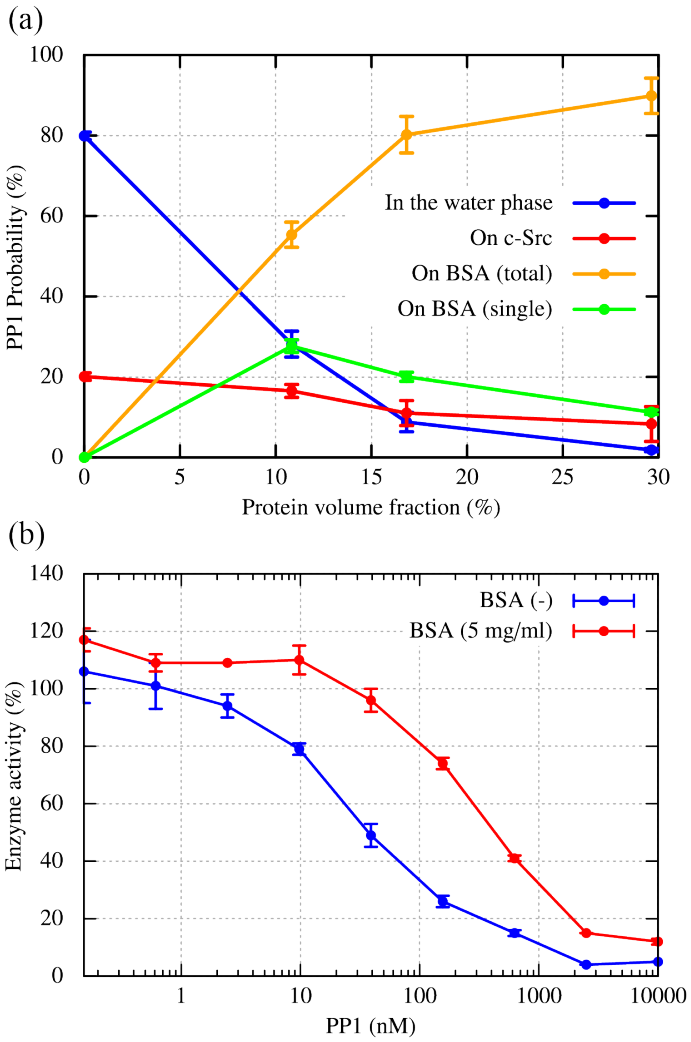


Fig. 2: Reduced efficacy of PP1 inhibitor in the presence of the crowder (BSA) proteins. (a) The probability of finding PP1 in the bulk water region, on c-Src kinase, and on the BSAs as functions of the protein volume fraction. (b) Enzymatic activity of c-Src kinase in the dilute solution ($[\text{c-Src}] = 65.7$ nM) and in the

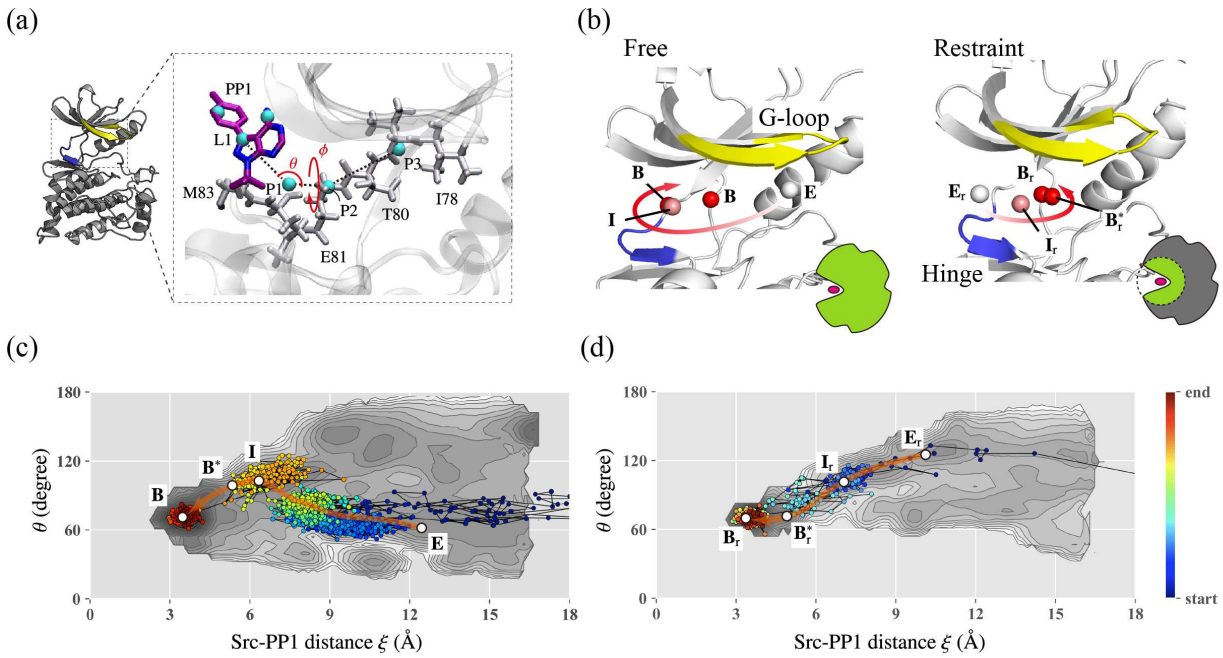
crowded solution ([c-Src] = 65.7 nM, [BSA] = 12.7 μM (5 mg/ml)). Error bars of (a) and (b) represent the standard error of the mean from seven trajectories and three indepdent assays, respectively.

Based on the PP1 populations in the bulk water region and on the protein surfaces, we predict that the inhibitor efficacy is reduced in crowded environments compared to that in dilute solution. *In vitro* PP1 inhibition assays with c-Src kinase, Src substrate, and ATP, were performed in the presence or absence of BSA to test this hypothesis. To avoid aggregation in the protein solution, the protein concentration in the experiments was much lower than in the simulations. The c-Src kinase concentration was set to 65.7 nM in the absence of BSA, while the BSA concentration was much larger (12.7 μM = 5 mg/ml) than kinase (65.7 nM) to realize a crowded protein solution. The 50% inhibition concentration (IC_{50}) of the PP1 considerably increased from 33.80 nM to 290.4 nM in the presence of BSA (Fig. 2(b)). This validates our hypothesis of a reduced inhibitor efficacy in crowded protein solutions.

Slow-down of PP1 diffusion in crowded solution

The mean square displacements (MSDs) of PP1 were analyzed in the bulk water region, on the surfaces of c-Src kinase or BSA. PP1 diffusion in the bulk water region slowed down as the protein volume fraction increased (Supplementary Fig. 11(a)). PP1 diffusion on the protein surfaces is nearly identical to diffusion of c-Src kinase or BSA in Src4BSA and Src8BSA (Supplementary Figs. 11(b) and (c)). In the crowded solutions, the probability of finding PP1 on the protein surfaces is much higher than that in the bulk water region (Fig. 2(a)). PP1 diffusion is thus significantly affected by proteins in the crowded solutions as PP1 mostly diffuses along with the proteins while bound to it. This is further manifested by long residence times on the 1-100 ns time scale for PP1 association with the c-Src kinase and BSA as shown in Supplementary Table 6.

Ligand-binding pathways in crowded environments



164

165 **Fig. 3:** Binding trajectories projected onto the two-dimensional free energy landscapes (FELs) obtained
166 from the gREST/REUS simulations in the dilute solution. (a) Definition of azimuth angles (θ, ϕ) for
167 describing the relative position (See Supporting information for the detail). (b) The major binding
168 pathway in the dilute solution obtained from the gREST/REUS simulation in different conditions. Left
169 panel: simulation without any restraints (free), Right panel: simulation in which the residues outside the
170 binding pocket is kept close to the crystal structure. **E**, **I**, **B***, and **B** respectively indicate encounter
171 complex, intermediate, pre-bound, and bound states. The subscript r denotes the presence of the restraints.
172 (c, d) The binding trajectories in the dilute solution and in the crowded solution (Src8BSA) projected onto
173 the 2D-FEL from 'free' simulation and that from 'restraint' simulation, respectively.

174

175 In the present simulations, we observed four and two PP1 binding events to the canonical binding site in
176 the dilute solution and Src8BSA, respectively (Supplementary Figs. 12 and 13). In dilute solution, PP1
177 approaches to the canonical binding site from the glycine rich loop (G-loop) in c-Src kinase. This binding
178 pathway is different from that observed in Src8BSA: PP1 first reaches the hinge region of the kinase and

then intrudes into the canonical binding site. These binding trajectories are projected onto the two free energy landscapes (FELs) that were obtained in our previous study on the same system in dilute solution using the gREST/REUS simulations[48] (Fig. 3 and Supplementary Figs. 14-16). The two-dimensional FELs illustrate the changes in free energy along the c-Src kinase-PP1 distance (ξ) and one of the coordinates (an azimuth angle θ) representing the PP1 position with respect to c-Src kinase (Fig. 3(a)). Two FELs were computed in two different conditions: in one simulation, c-Src kinase fluctuated freely (referred to as 'free') (Fig. 3(b) and Supplementary Fig. 14(a)), while strong positional restraints to the kinase prohibited the kinase flexibilities in the other simulation (referred to as 'restraint') (Fig. 3(b) and Supplementary Fig. 14(b)). On the major binding pathway in the 'free' FEL, PP1 approaches c-Src kinase from the small $\theta \sim 60^\circ$ (Supplementary Fig. 14(a)), while the binding pathway of the 'restraint' FEL starts from the large $\theta \sim 120^\circ$ (Supplementary Fig. 14(b)).

In the present simulations, we observed four and two PP1 binding events to the canonical binding site in the dilute solution and Src8BSA, respectively (Supplementary Figs. 12 and 13). In dilute solution, PP1 approaches to the canonical binding site from the glycine rich loop (G-loop) in c-Src kinase. This binding pathway is different from that observed in Src8BSA: PP1 first reaches the hinge region of the kinase and then intrudes into the canonical binding site. These binding trajectories are projected onto the two free energy landscapes (FELs) that were obtained in our previous study on the same system in dilute solution using the gREST/REUS simulations[48] (Fig. 3 and Supplementary Figs. 14-16). The two-dimensional FELs illustrate the changes in free energy along the c-Src kinase-PP1 distance (ξ) and one of the coordinates (an azimuth angle θ) representing the PP1 position with respect to c-Src kinase (Fig. 3(a)). Two FELs were computed in two different conditions: in one simulation, c-Src kinase fluctuated freely (referred to as 'free') (Fig. 3(b) and Supplementary Fig. 14(a)), while strong positional restraints to the kinase prohibited the kinase flexibilities in the other simulation (referred to as 'restraint') (Fig. 3(b) and Supplementary Fig. 14(b)). On the major binding pathway in the 'free' FEL, PP1 approaches c-Src kinase from the small $\theta \sim$

60° (Supplementary Fig. 14(a)), while the binding pathway of the 'restraint' FEL starts from the large $\theta \sim$ 120° (Supplementary Fig. 14(b)).

Conformatioal shifts of a Tyr sidechain upon crowding

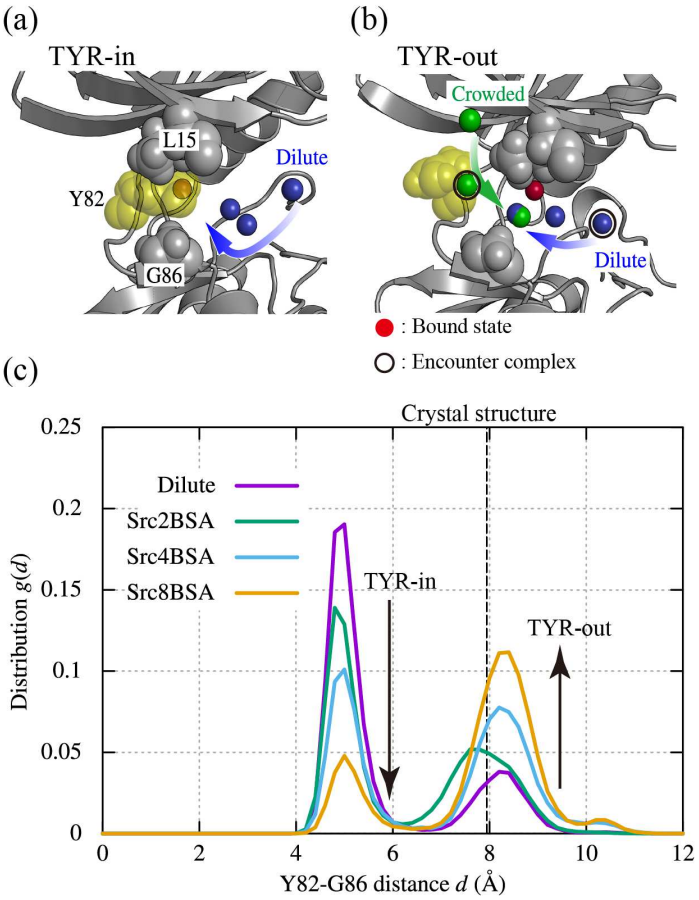


Fig. 4: Distribution of the TYR-in and TYR-out states of c-Src kinase. (a) Tyr-in state structure together with a binding trajectory in the dilute solution. (b) Tyr-out state structure with binding trajectory in the dilute and crowded solution (Src8BSA). Red, blue, and green spheres represent the PP1 bound position in the crystal structure, representative binding pathways in dilute and crowded solutions, respectively. Encounter states in the pathways are circled. (c) Distribution for Tyr82-Gly86 distance, $g(d)$. The distance between Tyr82 and Gly86 in a crystal structure of c-Src kinase in complex with an inhibitor (ID: 1Y57) is shown with dashed line for comparison[52].

To examine the molecular mechanisms for why the binding pathways differ between dilute and crowded solutions, we investigated local conformational changes in c-Src kinase as protein volume fraction increased. We found two characteristic sidechain conformations of Tyr82 at the hinge region, which we refer to as TYR-in and TYR-out. In TYR-in, two hydrophobic residues Leu15 and Gly86 sandwich Tyr82, thereby blocking the pathway from the hinge region (Fig. 4(a)). PP1 thus interacts with G-loop to form the encounter complex **E**. In TYR-out, on the other hand, G-loop shifts downward (~1 Å, Supplementary Fig. 17). Consequently, Tyr82 is positioned away from Leu15 and Gly86 (Fig. 4(b)), allowing PP1 access to the hinge region at the encounter complex $\mathbf{E_r}$. The distribution of Tyr82-Gly86 distance, $g(d)$, is centered around d = 5 Å in TYR-in and d = 8.5 Å in TYR-out (Fig. 4(c)). TYR-in is dominant in the dilute condition, while we observe TYR-out as a minor population. As the protein concentration increases (in the order of Src2BSA, Src4BSA, and Src8BSA), the ratio between TRY-in and TYR-out is gradually changed. Finally, TYR-out becomes the major population in Src8BSA.

In a crystal structure of c-Src kinase in complex with an inhibitor (PDB ID: 1Y57), the Tyr82-Gly86 distance is close to, but slightly different from that in TYR-out (Fig. 4(c)). Interestingly, a water molecule exists between Tyr82 and Gly86 in the crystal structure of c-Src kinase in the apo form (PDB ID: 1YOJ) (Supplementary Fig. 18). In the same position, we observe water densities in the MD snapshots where the Tyr82 sidechain assumes the TYR-out structure (Supplementary Fig. 18).

Discussion

The crowding effects on c-Src kinase/PP1 binding can be categorized into two different types: 1) The slowdown of PP1 diffusion, the decrease of non-canonical PP1 binding sites in the kinase, and the reduced efficacy of the inhibitor are due to generic crowding effects. We expect similar effects with other proteins under crowded conditions. 2) The different PP1 binding pathways in crowded vs. solution are specific to c-Src kinase. Whether such an effect is present for other ligands and other enzymes is highly dependent on

238 the molecular details of a given system. Hereafter, we reconsider these two crowding effects based on the
239 traditional excluded volume effects as well as weak and non-specific molecular interactions in crowded and
240 cellular environments.

241 The slowdown of PP1 diffusion in crowded solutions is consistent with the traditional understanding of
242 crowding, which predicts that the kinetics of protein-ligand binding becomes slower under crowded
243 conditions due to increased solution viscosity[3]. Indeed, we found that the translational diffusion of PP1 and
244 water even in the bulk water region of the crowded solutions become slower (Supplementary Fig. 19). On
245 the protein surfaces, ligand diffusion was much more retarded due to non-specific interactions with proteins.
246 In our previous MD simulations of the *MG* cytoplasm, charged and/or large metabolites were also found to
247 stay on the protein surfaces for a long time, while hydrophobic metabolites like Val were found to remain
248 solvated in larger fractions[15]. The importance of non-specific interactions between macromolecular
249 crowders and substrates were also seen in a recent enzyme assay[53]. Therefore, it is becoming evident that
250 the heterogeneous environments for proteins and ligands in crowded cellular conditions must be taken into
251 account to correctly estimate protein-ligand binding kinetics in a living cell. On the computational side, this
252 insight could be incorporated into Brownian dynamics (BD) based approaches[54–57], which are commonly
253 used to predict protein-ligand binding kinetics for dilute conditions. The traditional theory of
254 macromolecular crowding predicts that protein-ligand and protein-protein interactions in crowded
255 environments would be stronger than those in dilute solution. The volume exclusion effect has been
256 considered as a main driving force[21,58]. This effect suggests an increase of the effective concentration of the
257 substrates around the target macromolecule. For example, the activity of a DNA ligase correlates with an
258 increase in the effective concentration of the substrate due to crowding[59]. Contrary, the present study shows
259 that the effective ligand concentration is reduced as a function of protein volume fraction, because target
260 protein-crowder interactions sterically block non-canonical binding sites. Ligands might be trapped on the
261 surfaces of crowder proteins. This is again consistent with the observations in our previous simulations of
262 the *MG* cytoplasm.[15] Based on these common observations, we hypothesize that an effectively reduced

ligand concentration in soluble regions is common in crowded solution as well as cellular environments like the cytoplasm. A decrease in the ligand concentration is expected to reduce the inhibitor efficacy in these environments, as our kinase assay showed. In future work, binding free energy calculations under crowded conditions could quantitatively address the relation between the effective inhibitor concentration and its efficacy.

In Src8BSA simulations, the PP1 binding pathway is entirely different from that in dilute simulations. This suggests that crowding effects could change protein-ligand binding mechanisms. The crystal structure of c-Src kinase has a local structure similar to what is dominant in Src8BSA. However, the crystal environments and crowded solutions are different conditions in terms of their flexibilities and molecular interactions with nearby molecules. We observed an increasing shift in population of the local conformation of c-Src kinase as crowding was increased. As shown in Supplementary Fig. 20, the transitions between Tyr-in and Tyr-out forms occur dynamically. Even in the dilute solution, Tyr-out exists as a minor conformation. Due to non-specific molecular interactions between crowder BSAs and the G-loop of c-Src kinase, Tyr-out conformations are preferred in the crowded environment (see Supplementary Fig. 17). Therefore, in addition to the macromolecular crowding effects on open-to-close conformations or folding/unfolding equilibrium[60,61], small and local conformational shifts near the active site of proteins due to crowding may change molecular mechanisms underlying their biological functions as we describe in this study.

We showed that the sidechain position of Tyr82 driven by the downward shift of the G-loop is the key determinant for changing the PP1 binding pathway. It has been suggested previously that the downward shift of the G-loop could impact drug efficacy[22]. Also, the sequence alignment data set of 490 kinases[62] (Supplementary Fig. 21) shows that about 44% of kinases share the tyrosine near the hinge region. In addition, about 16% of the kinases have phenylalanine instead of tyrosine at the same position. Therefore, the suggested role of Tyr82 as a key amino acid in the binding of ligands and inhibitors - and the possibility to be sensitive to crowding effects suggested here - may be shared by a broad range of kinases.

Methods

MD simulations

The initial structure of c-Src kinase was prepared from an X-ray structure of unphosphorylated active-like c-Src kinase (PDB: 1Y57)[52]. Similar to the previous conventional MD[51] and the gREST/REUS simulations[48], only the kinase domain (residues 82-258) was used in this study. A c-Src kinase and 24 mM PP1 inhibitors were placed and then solvated by 150 mM NaCl aqueous solution (dilute). To represent crowded protein solutions, bovine serum albumin (BSA) proteins (PDB: 4F5S)[63] are also included as crowders. In Src2BSA, Src4BSA, and Src8BSA, two, four, and eight BSAs were simulated with c-Src kinase and PP1. We used the AMBER ff99SB-ILDN[64,65] and TIP3P[66] parameters for the proteins and water molecules, respectively. The PP1 parameters were prepared with GAFF with AM1-BCC[67]. For each system, a set of different initial configurations was prepared. The simulations were performed with the GENESIS software[68,69] and with Anton2[24] in Pittsburg supercomputing center (PSC). Further detail is available as a part of Supplementary Information.

Protein expressions and purification

The gene encoding human Src kinase domain (residues 260-533) was cloned into the pCR2.1 vector (Invitrogen) and expressed as a fusion with N-terminal histidine and GST tags using *Escherichia coli* cell-free reaction supplemented with chaperones and YopH[70,71]. The protein was purified by affinity chromatography using a HisTrap column (GE Healthcare) and further by ion exchange on a HiTrap Q column (GE Healthcare) and used for *in vitro* kinase assay.

In vitro inhibition assay

In vitro inhibition assay was performed using the ADP-Glo Kinase Assay (Promega) under the conditions based on the manufacturer's protocol. Briefly, purified Src protein (20 ng) was incubated with the indicated concentrations of PP1 (Cayman Chemical), 100 μg/ml Src substrate (SignalChem Pharmaceuticals), and 10 μM ATP for 60 min at room temperature in a reaction solution containing 0.2% DMSO, 40 mM Tris-HCl

(pH7.5), 20 mM $MgCl_2$, 2 mM $MnCl_2$, 50 μM DTT with and without 5.0 mg/ml BSA. The reaction was terminated by adding the ADP-Glo Reagent, and then ADP generation was detected using the luciferase reaction by monitoring the luminescence on an ARVO X3 microscale luminometer (PerkinElmer). IC_{50} was determined by curve fitting the data using the GraphPad Prism8 program (GraphPad software).

Data Availability

The trajectories are available from the corresponding authors upon reasonable request.

Code Availability

The source code of GENESIS (https://www.r-ccs.riken.jp/labs/cbrt/) is distributed under the GNU Lesser General Public License version 3.

References

1. Minton, A. P. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* **276**, 10577–10580 (2001).

2. Rivas, G. & Minton, A. P. Macromolecular crowding in vitro, in vivo, and in between. *Trends Biochem. Sci.* **41**, 970–981 (2016).

3. Kuznetsova, I., Zaslavsky, B., Breydo, L., Turoverov, K. & Uversky, V. Beyond the excluded volume effects: mechanistic complexity of the crowded milieu. *Molecules* **20**, 1377–1409 (2015).

4. Zimmerman, S. B. & Minton, A. P. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 27–65 (1993).

5. Wang, Y., Li, C. & Pielak, G. J. Effects of proteins on protein diffusion. *J. Am. Chem. Soc.* **132**, 9392–9397 (2010).

6. Miklos, A. C., Sarkar, M., Wang, Y. & Pielak, G. J. Protein crowding tunes protein stability. *J. Am. Chem. Soc.* **133**, 7116–7120 (2011).

7. Inomata, K. *et al.* High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **458**, 106–109 (2009).

8. Smith, A. E., Zhou, L. Z. & Pielak, G. J. Hydrogen exchange of disordered proteins inEscherichia coli. *Protein Sci.* **24**, 706–713 (2015).

9. Ikeya, T., Güntert, P. & Ito, Y. Protein Structure Determination in Living Cells. *Int. J. Mol. Sci.* **20**, 2442 (2019).

10. Feig, M. & Sugita, Y. Variable interactions between protein crowders and biomolecular solutes are important in understanding cellular crowding. *J. Phys. Chem. B* **116**, 599–605 (2011).

11. Harada, R., Sugita, Y. & Feig, M. Protein crowding affects hydration structure and dynamics. *J. Am. Chem. Soc.* **134**, 4842–4849 (2012).

12. Harada, R., Tochio, N., Kigawa, T., Sugita, Y. & Feig, M. Reduced native state stability in crowded cellular environment due to protein--protein interactions. *J. Am. Chem. Soc.* **135**, 3696–3701 (2013).

13. Nawrocki, G., Wang, P., Yu, I., Sugita, Y. & Feig, M. Slow-Down in Diffusion in Crowded Protein Solutions Correlates with Transient Cluster Formation. *J. Phys. Chem. B* (2017).

14. Wang, P., Yu, I., Feig, M. & Sugita, Y. Influence of protein crowder size on hydration structure and dynamics in macromolecular crowding. *Chem. Phys. Lett.* **671**, 63–70 (2017).

15. Yu, I. *et al.* Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *Elife* **5**, (2016).

16. von Bülow, S., Siggel, M., Linke, M. & Hummer, G. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. *Proc. Natl. Acad. Sci.* **116**, 9843–9852 (2019).

17. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789 (2009).

18. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

19. Forli, S. *et al.* Computational protein--ligand docking and virtual drug screening with the AutoDock suite. *Nat. Protoc.* **11**, 905–919 (2016).

20. De Vivo, M., Masetti, M., Bottegoni, G. & Cavalli, A. Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* **59**, 4035–4061 (2016).

21. Kuznetsova, I., Turoverov, K. & Uversky, V. What macromolecular crowding can do to a protein. *Int. J. Mol. Sci.* **15**, 23090–23140 (2014).

22. Uitdehaag, J. C. M. *et al.* Target residence time-guided optimization on TTK kinase results in inhibitors with potent anti-proliferative activity. *J. Mol. Biol.* **429**, 2211–2230 (2017).

23. Shaw, D. E. *et al.* Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **51**, 91–97 (2008).

24. Shaw, D. E. *et al.* Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. in *Proceedings of the international conference for high performance computing, networking, storage and analysis* 41–53 (2014).

25. Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141–151 (1999).

26. Kokubo, H., Tanaka, T. & Okamoto, Y. Two-dimensional replica-exchange method for predicting protein--ligand binding structures. *J. Comput. Chem.* **34**, 2601–2614 (2013).

27. Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **99**, 12562–12566 (2002).

28. Tiwary, P. & Parrinello, M. From Metadynamics to Dynamics. *Phys. Rev. Lett.* **111**, (2013).

29. Miao, Y., Feher, V. A. & McCammon, J. A. Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *J. Chem. Theory Comput.* **11**, 3584–3595 (2015).

30. Sinko, W., Miao, Y., de Oliveira, C. A. F. & McCammon, J. A. Population based reweighting of scaled molecular dynamics. *J. Phys. Chem. B* **117**, 12759–12768 (2013).

31. Schuetz, D. A. *et al.* Predicting Residence Time and Drug Unbinding Pathway through Scaled Molecular Dynamics. *J. Chem. Inf. Model.* **59**, 535–549 (2018).

32. Huber, G. A. & Kim, S. Weighted-ensemble Brownian dynamics simulations for protein association reactions. *Biophys. J.* **70**, 97–110 (1996).

33. Zuckerman, D. M. & Chong, L. T. Weighted ensemble simulation: review of methodology, applications, and software. *Annu. Rev. Biophys.* **46**, 43–57 (2017).

34. Dickson, A. & Brooks III, C. L. WExplore: hierarchical exploration of high-dimensional spaces using the weighted ensemble algorithm. *J. Phys. Chem. B* **118**, 3532–3542 (2014).

35. Harada, R. & Kitao, A. Parallel cascade selection molecular dynamics (PaCS-MD) to generate conformational transition pathway. *J. Chem. Phys.* **139**, 07B611_1 (2013).

36. Bowman, G. R., Pande, V. S. & Noé, F. *An introduction to Markov state models and their application to long timescale molecular simulation*. vol. 797 (Springer Science & Business Media, 2013).

37. Nawrocki, G., Karaboga, A., Sugita, Y. & Feig, M. Effect of protein--protein interactions and solvent viscosity on the rotational diffusion of proteins in crowded environments. *Phys. Chem. Chem. Phys.* **21**, 876–883 (2019).

38. Patel, A. *et al.* ATP as a biological hydrotrope. *Science (80-.).* **356**, 753–756 (2017).

39. Thomas, S. M. & Brugge, J. S. Cellular functions regulated by Src family kinases. *Annu. Rev. Cell Dev. Biol.* **13**, 513–609 (1997).

40. Yang, S., Banavali, N. K. & Roux, B. Mapping the conformational transition in Src activation by cumulating the information from multiple molecular dynamics trajectories. *Proc. Natl. Acad. Sci.* **106**, 3776–3781 (2009).

41. Meng, Y., Lin, Y. & Roux, B. Computational Study of the ``DFG-Flip'' Conformational Transition in c-Abl and c-Src Tyrosine Kinases. *J. Phys. Chem. B* **119**, 1443–1456 (2015).

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

43. Foda, Z. H., Shan, Y., Kim, E. T., Shaw, D. E. & Seeliger, M. A. A dynamically coupled allosteric network underlies binding cooperativity in Src kinase. *Nat. Commun.* **6**, 5939 (2015).

44. Sultan, M. M., Kiss, G. & Pande, V. S. Towards simple kinetic models of functional dynamics for a kinase subfamily. *Nat. Chem.* **10**, 903–909 (2018).

45. Tsai, C.-C., Yue, Z. & Shen, J. How Electrostatic Coupling Enables Conformational Plasticity in a Tyrosine Kinase. *J. Am. Chem. Soc.* **141**, 15092–15101 (2019).

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

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

48. Re, S., Oshima, H., Kasahara, K., Kamiya, M. & Sugita, Y. Encounter complexes and hidden poses of kinase-inhibitor binding on the free-energy landscape. *Proc. Natl. Acad. Sci.* **116**, 18404–18409 (2019).

49. Hanke, J. H. *et al.* Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor Study of Lck-and FynT-dependent T cell activation. *J. Biol. Chem.* **271**, 695–701 (1996).

50. Sugita, Y., Kitao, A. & Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **113**, 6042–6051 (2000).

51. Shan, Y. *et al.* How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* **133**, 9181–9183 (2011).

52. Cowan-Jacob, S. W. *et al.* The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861–871 (2005).

53. Aumiller Jr, W. M., Davis, B. W., Hatzakis, E. & Keating, C. D. Interactions of macromolecular crowding agents and cosolutes with small-molecule substrates: effect on horseradish peroxidase activity with two different substrates. *J. Phys. Chem. B* **118**, 10624–10632 (2014).

54. Northrup, S. H., Allison, S. A. & McCammon, J. A. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys.* **80**, 1517–1524 (1984).

55. Luty, B. A., McCammon, J. A. & Zhou, H.-X. Diffusive reaction rates from Brownian dynamics simulations: Replacing the outer cutoff surface by an analytical treatment. *J. Chem. Phys.* **97**, 5682–5686 (1992).

56. Votapka, L. W. & Amaro, R. E. Multiscale Estimation of Binding Kinetics Using Brownian Dynamics, Molecular Dynamics and Milestoning. *PLOS Comput. Biol.* **11**, e1004381 (2015).

57. Votapka, L. W., Jagger, B. R., Heyneman, A. L. & Amaro, R. E. SEEKR: Simulation Enabled Estimation of Kinetic Rates, A Computational Tool to Estimate Molecular Kinetics and Its Application to Trypsin--Benzamidine Binding. *J. Phys. Chem. B* **121**, 3597–3606 (2017).

58. Gomez, D. & Klumpp, S. Biochemical reactions in crowded environments: revisiting the effects of volume exclusion with simulations. *Front. Phys.* **3**, (2015).

59. Zimmerman, S. B. & Pheiffer, B. H. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or Escherichia coli. *Proc. Natl. Acad. Sci.* **80**, 5852–5856 (1983).

60. Zhou, H.-X. Effect of mixed macromolecular crowding agents on protein folding. *Proteins Struct. Funct. Bioinforma.* **72**, 1109–1113 (2008).

61. Dong, H., Qin, S. & Zhou, H.-X. Effects of Macromolecular Crowding on Protein Conformational Changes. *PLoS Comput. Biol.* **6**, e1000833 (2010).

62. Moebitz, H. The ABC of protein kinase conformations. *Biochim. Biophys. Acta (BBA)-Proteins Proteomics* **1854**, 1555–1566 (2015).

63. Bujacz, A. Structures of bovine, equine and leporine serum albumin. *Acta Crystallogr. Sect. D - Biol. Crystallogr.* **68**, 1278–1289 (2012).

64. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinforma.* **78**, 1950–1958 (2010).

65. Hornak, V. *et al.* Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinforma.* **65**, 712–725 (2006).

66. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

67. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

68. Jung, J. *et al.* GENESIS: a hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **5**, 310–323 (2015).

69. Kobayashi, C. *et al.* GENESIS 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms. *J. Comput. Chem.* **38**, 2193–2206 (2017).

70. Katsura, K. *et al.* A reproducible and scalable procedure for preparing bacterial extracts for cell-free protein synthesis. *J. Biochem.* **162**, 357–369 (2017).

71. Katsura, K. *et al.* Phosphorylated and non-phosphorylated HCK kinase domains produced by cell-free protein expression. *Protein Expr. Purif.* **150**, 92–99 (2018).

Acknowledgements (optional)

Keep acknowledgements brief and do not include thanks to anonymous referees or editors, or effusive comments. Grant or contribution numbers may be acknowledged. We acknowledge the computing time granted by RIKEN Advanced Center for Computing and Communication (HOKUSAI GreatWave and BigWaterfall), RIKEN Center for Computational Science (K computer), and HPCI system (Project ID: hp170254, hp180201, hp180274, hp190097, hp190181, hp200129, hp200135). This research was supported by MEXT as "Priority Issue on Post-K computer" (Building Innovative Drug Discovery Infrastructure Through Functional Control of Biomolecular Systems) (to Y.S.), "Program for Promoting Researches on the Supercomputer Fugaku" (Biomolecular dynamics in a living cell/MD-driven Precision Medicine) (to Y.S.), MEXT Grant-in-Aid for KIBAN (S) (19H05645) (to Y.S.), RIKEN pioneering projects "Dynamic Structural Biology" and "Glycolipidologue Initiative" (to Y.S.). Funding was provided by the

National Science Foundation grant MCB 1817307 (to M.F. and G.N.), by the National Institutes of Health grant R35 GM126948 (to M.F.). Computer time was used on the Anton2 special-purpose supercomputer at the Pittsburgh Supercomputing Center (PSCA18053P). We thank Mr. Kazuharu Hanada and Dr. Kazushige Katsura for their technical assistance on the experiments.

Ethics declarations

Competing interests

The authors declare no competing interests.

Supplementary Information

Supporting information is available for this paper.

Figures

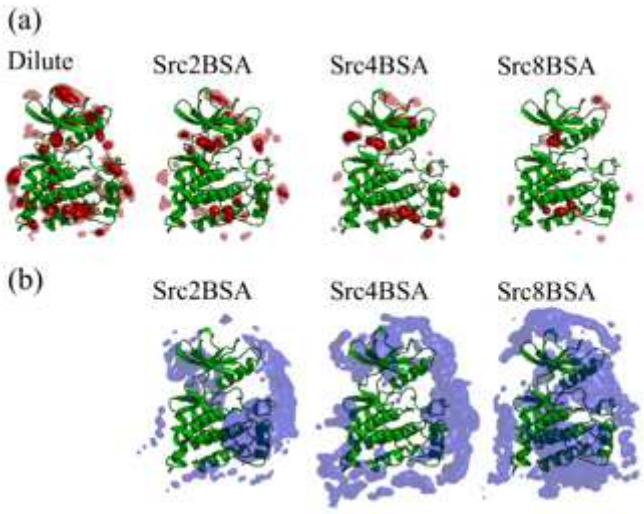


Figure 1

Spatial distribution functions (SDF) of (a) PP1 and (b) BSA around c-Src kinase. In the case of BSA-SDF, C□ atoms are used for analysis. The PP1-SDF is shown as isosurface at 0.5% (transparent) and 1.5% occupancies (solid). For BSA-SDF, the isosurface at 1.0% occupancy is shown.

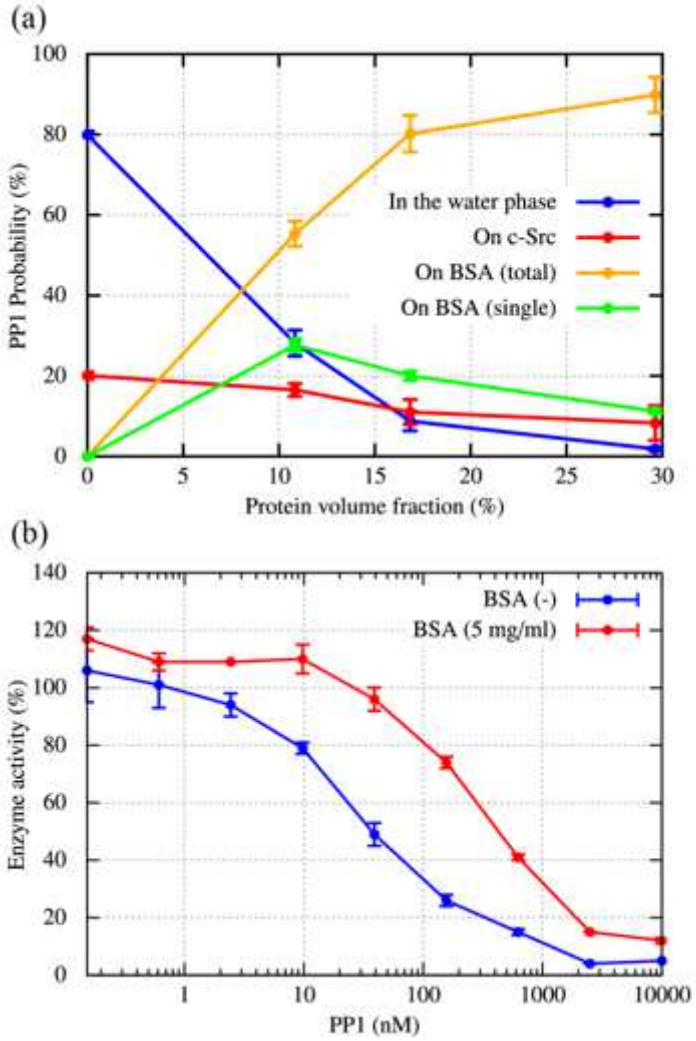


Figure 2

Reduced efficacy of PP1 inhibitor in the presence of the crowder (BSA) proteins. (a) The probability of finding PP1 in the bulk water region, on c-Src kinase, and on the BSAs as functions of the protein volume fraction. (b) Enzymatic activity of c-Src kinase in the dilute solution ([c"-" Src]=65.7 nM) and in the crowded solution ([c"-" Src]=65.7 nM, [BSA]=12.7 μM (5 mg/ml)). Error bars of (a) and (b) represent the standard error of the mean from seven trajectories and three indepdent assays, respectively.

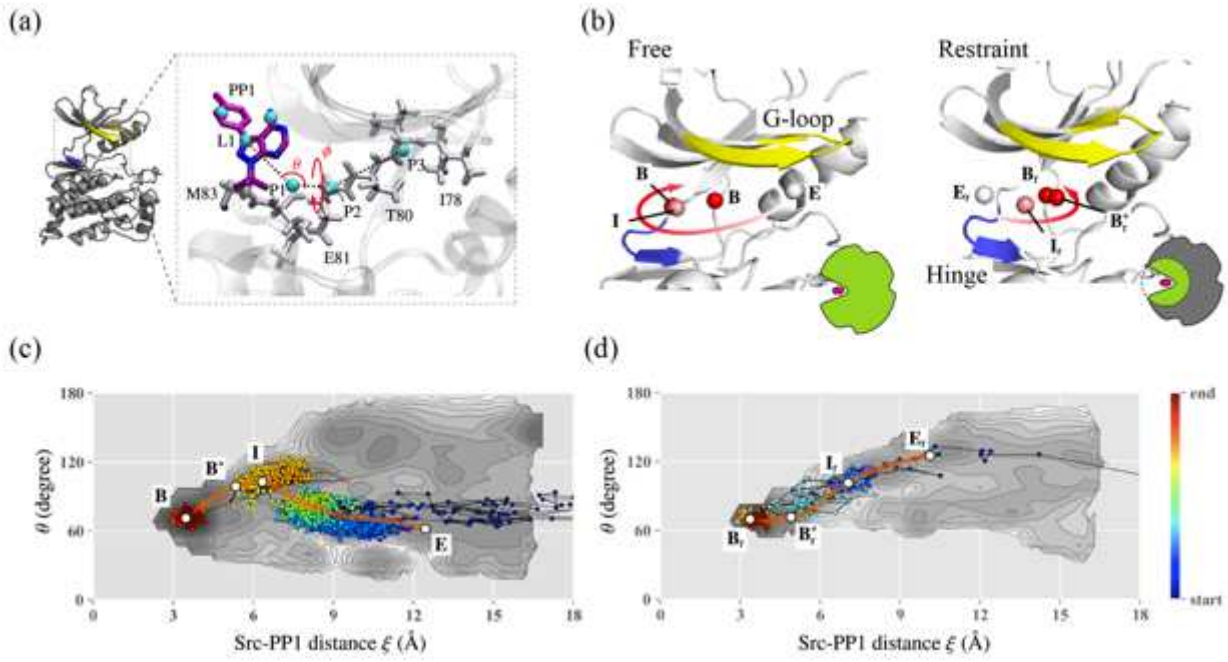


Figure 3

Binding trajectories projected onto the two-dimensional free energy landscapes (FELs) obtained from the gREST/REUS simulations in the dilute solution. (a) Definition of azimuth angles (□, □) for describing the relative position (See Supporting information for the detail). (b) The major binding pathway in the dilute solution obtained from the gREST/REUS simulation in different conditions. Left panel: simulation without any restraints (free), Right panel: simulation in which the residues outside the binding pocket is kept close to the crystal structure. E, I, B*, and B respectively indicate encounter complex, intermediate, pre-bound, and bound states. The subscript r denotes the presence of the restraints. (c, d) The binding trajectories in the dilute solution and in the crowded solution (Src8BSA) projected onto the 2D-FEL from 'free' simulation and that from 'restraint' simulation, respectively.

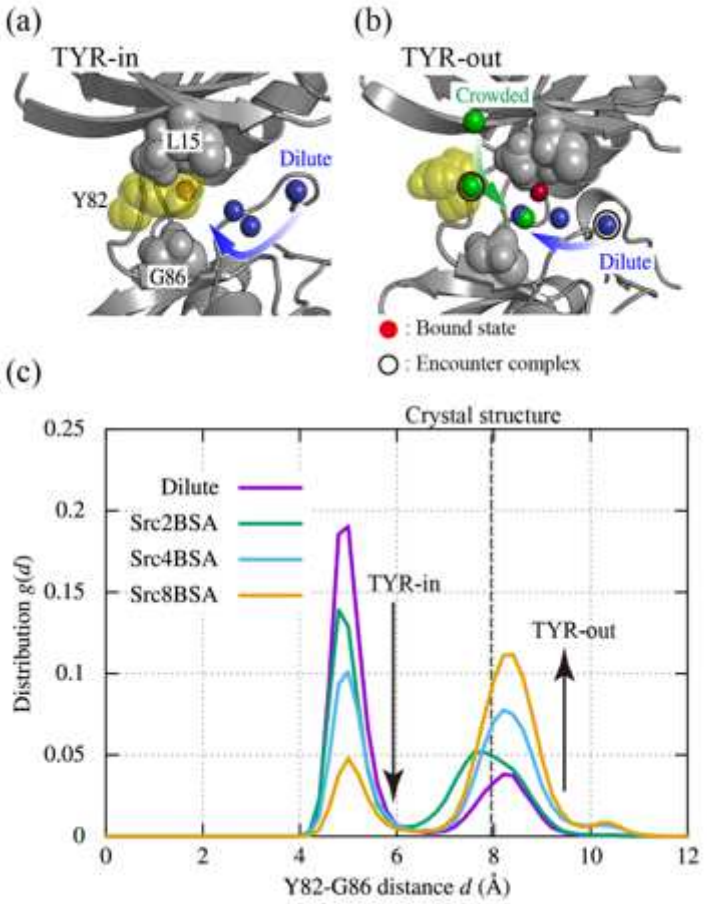


Figure 4

Distribution of the TYR-in and TYR-out states of c-Src kinase. (a) Tyr-in state structure together with a binding trajectory in the dilute solution. (b) Tyr-out state structure with binding trajectory in the dilute and crowded solution (Src8BSA). Red, blue, and green spheres represent the PP1 bound position in the crystal structure, representative binding pathways in dilute and crowded solutions, respectively. Encounter states in the pathways are circled. (c) Distribution for Tyr82-Gly86 distance, g(d). The distance between Tyr82 and Gly86 in a crystal structure of c-Src kinase in complex with an inhibitor (ID: 1Y57) is shown with dashed line for comparison52.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MovieS1.mov
- MovieS2.mov
- SINatComm20200904.docx