



Practical Protocols for Efficient Sampling of Kinase-Inhibitor Binding Pathways Using Two-Dimensional Replica-Exchange Molecular Dynamics

Ai Shinobu¹, Suyong Re^{1,2} and Yuji Sugita^{1,3,4*}

¹RIKEN Center for Biosystems Dynamics Research, Kobe, Japan, ²Artificial Intelligence Center for Health and Biomedical Research, National Institutes of Biomedical Innovation, Health, and Nutrition, Ibaraki, Japan, ³Theoretical Molecular Science Laboratory, RIKEN Cluster for Pioneering Research, Saitama, Japan, ⁴RIKEN Center for Computational Science, Kobe, Japan

OPEN ACCESS

Edited by:

Yinglong Miao,
University of Kansas, United States

Reviewed by:

Alex Dickson,
Michigan State University,
United States
Chung Wong,
University of Missouri–St. Louis,
United States

*Correspondence:

Yuji Sugita
sugita@riken.jp

Specialty section:

This article was submitted to
Molecular Recognition,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 February 2022

Accepted: 31 March 2022

Published: 29 April 2022

Citation:

Shinobu A, Re S and Sugita Y (2022)
Practical Protocols for Efficient
Sampling of Kinase-Inhibitor Binding
Pathways Using Two-Dimensional
Replica-Exchange
Molecular Dynamics.
Front. Mol. Biosci. 9:878830.
doi: 10.3389/fmolb.2022.878830

Molecular dynamics (MD) simulations are increasingly used to study various biological processes such as protein folding, conformational changes, and ligand binding. These processes generally involve slow dynamics that occur on the millisecond or longer timescale, which are difficult to simulate by conventional atomistic MD. Recently, we applied a two-dimensional (2D) replica-exchange MD (REMD) method, which combines the generalized replica exchange with solute tempering (gREST) with the replica-exchange umbrella sampling (REUS) in kinase-inhibitor binding simulations, and successfully observed multiple ligand binding/unbinding events. To efficiently apply the gREST/REUS method to other kinase-inhibitor systems, we establish modified, practical protocols with non-trivial simulation parameter tuning. The current gREST/REUS simulation protocols are tested for three kinase-inhibitor systems: c-Src kinase with PP1, c-Src kinase with Dasatinib, and c-Abl kinase with Imatinib. We optimized the definition of kinase-ligand distance as a collective variable (CV), the solute temperatures in gREST, and replica distributions and umbrella forces in the REUS simulations. Also, the initial structures of each replica in the 2D replica space were prepared carefully by pulling each ligand from and toward the protein binding sites for keeping stable kinase conformations. These optimizations were carried out individually in multiple short MD simulations. The current gREST/REUS simulation protocol ensures good random walks in 2D replica spaces, which are required for enhanced sampling of inhibitor dynamics around a target kinase.

Keywords: molecular dynamics simulations, multi-dimensional replica-exchange simulations, generalized replica exchange with solute tempering, replica-exchange umbrella sampling, kinase-inhibitor binding

1 INTRODUCTION

Ligand binding to a target protein or enzyme plays important roles in many biological processes which regulate protein functional activity (Du et al., 2016). Understanding of the binding processes directly contributes to the design of effective drugs which specifically bind to target proteins. Recently, the drug residence time on a protein has been attracting attention in the development of

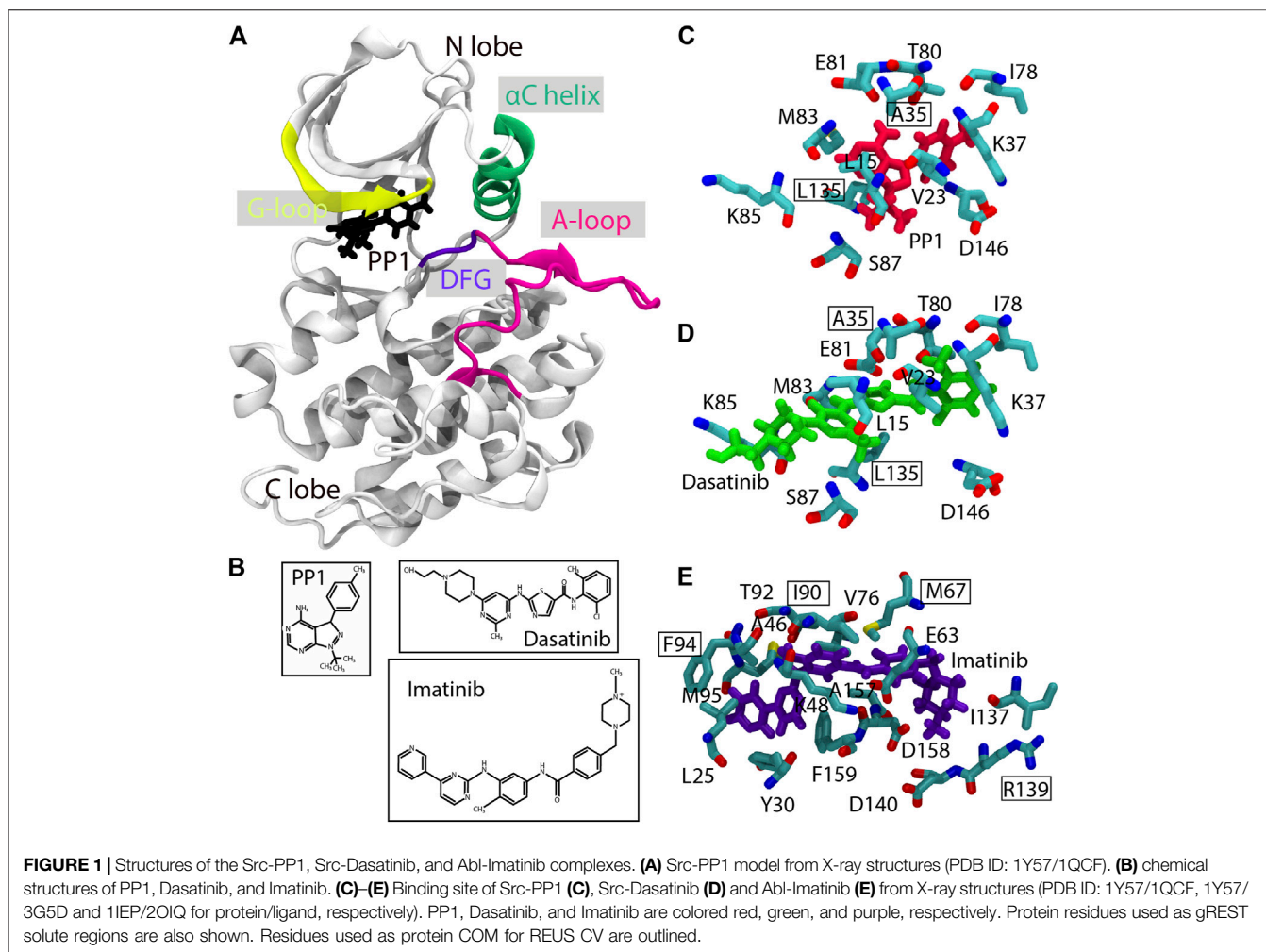
effective drugs (Bernetti et al., 2017; Schuetz et al., 2017). For this purpose, understanding the molecular mechanisms underlying protein-ligand binding processes, namely, binding pathways, transition states, encounter complexes, and binding kinetics, are essential, as well as sampling stable ligand-bound structures. Unlike most stable bound poses, transient and dynamic information is hardly accessible by experiments.

Molecular dynamics (MD) simulations are widely used to investigate conformational dynamics of biomolecules at the atomic level and are applied to many biological processes including protein-ligand binding/unbinding (De Vivo et al., 2016; Dickson et al., 2017; Bruce et al., 2018; Zhang et al., 2022). All-atom MD simulations can easily simulate protein dynamics on the 1–10 ms timescales, while high-performance MD-specialized computers are necessary to explore 1 ms or slower dynamics (Dror et al., 2011; Shan et al., 2011). Thus, conventional MD simulations of a protein-ligand complex are not sufficient for observing multiple binding/unbinding events, which are necessary for obtaining converged thermodynamics or free-energy landscapes. To go beyond, parallel trajectory MD methods (Silva et al., 2011; Plattner and Noé, 2015; Dickson, 2018; Tran et al., 2020) perform multiple short simulations and provide us with large amount of structural data for predicting long timescale dynamics. Another approach is the use of enhanced sampling MD methods such as replica-exchange MD (Sugita and Okamoto, 1999), metadynamics (Valsson et al., 2016), and others (Meng et al., 2015; Miao and McCammon, 2017; Spitaleri et al., 2018; Gobbo et al., 2019; Hénin et al., 2022) to explore a wider conformational space of systems with rugged energy landscapes by overcoming high energy barriers between multiple minimum states. Replica-exchange MD (REMD) (Sugita and Okamoto, 1999; Sugita et al., 2000) effectively overcome energy barriers through the exchange of system parameters between independently running replicas. In temperature REMD, high temperature replicas sample various conformations including unfolded, extended, or other flexible ones, while low temperature replicas explore stable structures existing at different energy minima through the replica exchange. In replica exchange with solute tempering (REST/REST2) (Liu et al., 2005; Terakawa et al., 2011; Wang et al., 2011), a specific region of interest is selected as “solute,” and “solute temperature” exchanges are attempted with a reduced number of replicas. Hence, REST/REST2 is applicable to larger biological systems than temperature-REMD owing to the reduced computational cost. Recently, we generalized the definition of “solute” in REST2 by selecting a part of the molecule of interest and/or a part of the potential energy function terms as “solute”. This method, which we refer to as the generalized REST (gREST) (Kamiya and Sugita, 2018), can reduce the number of replicas even further while observing efficient conformational dynamics of proteins or protein-ligand complexes. For instance, in gREST simulations of protein-ligand binding, the solute is defined as a target ligand as well as amino-acid sidechains near the target protein binding site, which accelerates ligand dynamics more than in REST2 simulations, where only the ligand molecule is selected as “solute”. The gREST method was applied for the prediction of the correct binding pose (Niitsu et al., 2019) and

affinities, when combined with absolute binding free energy calculations (Oshima et al., 2020). The replica-exchange umbrella sampling (REUS) method (Sugita et al., 2000; Fukunishi et al., 2002) exchanges geometrical parameters along a predefined collective variable (CV). This method is also applicable to large biological systems, if a good CV is used for describing the target conformational motion.

It is noteworthy that different parameters can be exchanged in a multidimensional fashion to further enhance conformational sampling of various biological systems (Sugita et al., 2000). Multidimensional REMD was first applied in protein-ligand binding simulations by Kokubo et al. (2013) where they combined REST2 with REUS (the REST2/REUS method). In their study, a target ligand was selected as solute in REST2 and the protein-ligand distance was used as a CV in REUS. After the success of this approach, we replaced REST2 with gREST and applied the gREST/REUS method to inhibitor binding/unbinding in the c-Src kinase/PP1 complex (gREST/REUS) (Re et al., 2019). We briefly describe the gREST/REUS method in the **Supplementary Text** and **Figure S1**. The simulations could enhance inhibitor dynamics around c-Src kinase and we observed a total of about 100 binding/unbinding events for all replicas. Using the well-converged free-energy landscapes of protein-ligand binding processes, multiple binding pathways, transition states, encounter complex structures, and other atomistic insights were obtained for the c-Src kinase-PP1 complex in solution.

The gREST/REUS method is theoretically applicable to any biological system for studying molecular mechanisms of protein-ligand binding/unbinding processes. However, the size and flexibility of the ligand increase the computational difficulty. Here, we re-examine the practical protocols of the two-dimensional (2D) gREST/REUS protein-ligand binding simulations and apply them for three kinase-inhibitor systems: c-Src kinase with PP1 (Src-PP1), c-Src kinase with Dasatinib (Src-Dasatinib), and c-Abl kinase with Imatinib (Abl-Imatinib) (**Figure 1**). As the size and flexibility of the ligand increases in the aforementioned order, binding simulations are expected to be more challenging. Kinase-inhibitor binding processes have been subjected to both long-time conventional MD (Shan et al., 2011; Morando et al., 2016; Paul et al., 2020; Sohraby et al., 2020) and enhanced sampling MD simulations (Yang et al., 2009; Lin et al., 2013; Tiwary et al., 2017; Gobbo et al., 2019; Koneru et al., 2019; Narayan et al., 2020; Spitaleri et al., 2020; Narayan et al., 2021; Shekhar et al., 2021). However, to gain more atomistic insight on protein-ligand binding processes, better computational algorithms and practical protocols are necessary. In this paper, we describe how to optimize parameters and procedures for the setup of gREST/REUS simulations and target biomolecular systems. The role of flexible inhibitor binding in c-Src/c-Abl kinases will be discussed in a separate paper, thus here we focus on the practical issues and the protocols, which are non-trivial when performing gREST/REUS simulations with more than a hundred replicas. The protocols presented here can be useful for carrying out ligand binding/unbinding simulations of various biomolecular systems with the gREST/REUS method on massively parallel supercomputers or GPU clusters.



2 METHODS

2.1 The gREST/REUS Simulation Protocols

The 2D-REMD methods such as gREST/REUS typically require a large number of replicas (i.e., more than 100 replicas), while they can realize better random walks in replica space including the bound, intermediate, and unbound states of the protein/ligand complexes. The preparation of replicas and the choice of solute temperatures in gREST and/or collective variables in REUS directly affects the conformational sampling efficiency. For instance, if there exist large distribution gaps between replicas, we cannot observe good random walks in replica space. This situation is equivalent to performing multiple independent REMD simulations with smaller number of replicas, which might lead to missing important intermediate structures and slow convergences of thermodynamic data. Initial setups of the gREST/REUS simulations are thus, essential for successful gREST/REUS calculations and for obtaining reliable simulation results.

In gREST/REUS, replica random walks are necessary in both the gREST and REUS dimensions. The former is realized only

when the solute region and replica temperatures are defined appropriately, and we can observe sufficient overlaps in potential energies between replicas at neighboring solute temperatures. In REUS simulations, the choice of CVs, replica distributions along the CV, and proper force constants in US potentials are all important. There are many parameters and choices of procedures in gREST/REUS simulations with more than 100 replicas. For simplifying the parameter optimization, we tuned the parameters in each dimension separately using multiple short MD or gREST/REUS simulations, as described below.

2.1.1 Definition of the Protein-Ligand Distance as a CV for REUS

The protein-ligand distance is commonly used in binding MD simulation studies. The distance is usually measured as that between the centers of mass (COMs) of the backbone atoms of the selected binding site residues (protein anchor sites) and ligand heavy atoms (ligand COM). For Src-PP1 and Src-Dasatinib, the backbone atoms of Ala35 and Leu135 in c-Src kinase are used as the protein anchor site. All the heavy atoms in

PP1 and Dasatinib were used for obtaining the ligand COM, since they are small compounds with less conformational flexibilities than Imatinib, which is composed of five rings. There are multiple choices for Abl-Imatinib for the protein anchor sites and the ligand COM. As the former, we tested four choices: “2 sites” (Ile90 and Arg139), “3 sites” (Ile90, Arg139, and Phe94), “4 sites” (Ile90, Arg139, Phe94, and Met67), and “5 sites” (Ile90, Arg139, Phe94, Met67, and Phe159), respectively. For the ligand COM, four definitions including a single ring (“Ring3”), three rings (“Ring135” and “Ring 234”), and all rings (“Ring all”) were examined. We expect that Imatinib flexibility is important not only near the binding site but also in the intermediate or unbound structures. A good combination of the protein anchor sites and the ligand COM may reduce the number of possible protein-ligand complex structures near the binding sites. In our protocols, the ligand COM definition was first examined for Imatinib and then, multiple choices of the protein anchor sites were tested for Abl-Imatinib simulations.

2.1.2 Preparation of Initial Structures in REUS

Thirty replicas were used for covering the protein-ligand distance in the range of 3–18 Å for Src-PP1, and 3–23 Å for Src-Dasatinib and Abl-Imatinib. We obtained the initial structure of each replica using two US simulations: In the “forward pull” simulation, the ligand was gradually pulled away from the protein binding site, while it was subsequently pulled back to the bound pose in the “reverse pull” simulation. Each replica was simulated for 300 ps with a force constant of 4 kcal/mol/Å². Positional restraints on the protein C α atoms with a force constant of 1 kcal/mol/Å² were necessary during the pulling simulations to prevent artificial deformations of the protein. In this stage, the 30 initial structures were set in equidistance in the REUS dimension.

2.1.3 Tuning of Solute Temperatures in gREST

The solute region in gREST was defined as the dihedral angle and the nonbonded energy terms of the ligands and binding-site residues of the proteins (ca. 10 residues defined as SITE residues in the X-ray structure as shown in **Figures 1C–E**, and listed in **Supplementary Table S1**). We determined the solute temperatures using the automatic parameter tuning tool in the GENESIS MD program (Kobayashi et al., 2017). Given initial temperatures and desired acceptance ratio as inputs, the tool finds a set of solute temperatures which satisfies the desired acceptance ratio. The initial temperatures and the target acceptance ratio were set in the range of 310–663 K and 0.2, respectively. We performed five rounds of the tuning simulations (1.1 ns for each replica), by gradually increasing the frequency of exchange attempts (from every 0.21 ps in the first round to every 2.1 ps for final round), until temperature values were converged. The tuning was performed in 1D-gREST simulations at the bound (protein-ligand distance of 3.0 Å), intermediate (10.3 Å for Src-PP1, and 15.0 Å for Src-Dasatinib and Abl-Imatinib), and the unbound states (18.1 Å for Src-PP1, and 23 Å for Src-Dasatinib and Abl-Imatinib). The final temperature values were taken as the average of those obtained at the above three states.

2.1.4 Determination of REUS Parameters

To ensure sufficient potential energy overlaps between adjacent replicas, which is a pre-requisite for good REUS performance, we conducted several short trial simulations, while manually tweaking the location and force constants. At each round, we assessed the distribution overlaps between replicas and the acceptance ratios, and accordingly modified the REUS parameters, namely, the center position and the force constant of each harmonic umbrella potential. The tuning procedures were repeated in 1D-REUS simulations at three solute temperatures: 310 K (at lowest) for all three systems, 478 K, 471 K, and 440 K (at middle), and 692 K, 663 K, 590 K (at highest) for Src-PP1, Src-Dasatinib, and Abl-Imatinib, respectively.

2.2 System Preparation

The initial structure of Src-PP1 was taken from our previous work (Re et al., 2019). In brief, we extracted the kinase domain (residues 260–533, renumbered 2–275 in this work) from the X-ray crystal structure of the active-like c-Src kinase (PDB ID: 1Y57) (Cowan-Jacob et al., 2005) and replaced the co-crystallized ligand with PP1 bound to c-Src (PDB ID: 1QCF) (Schindler et al., 1999). The initial structures of Src-Dasatinib and Abl-Imatinib were constructed with the same modeling protocol used for Src-PP1. For Src-Dasatinib, we used the kinase domain of c-Src kinase (PDB ID: 1Y57) (Cowan-Jacob et al., 2005) and replaced the co-crystallized ligand with Dasatinib from an X-ray structure (PDB ID: 3G5D) (Getlik et al., 2009). Similarly, for Abl-Imatinib, we used the kinase domain (residues 225–498, renumbered 2–275 in this work) of c-Abl kinase (PDB ID: 1IEP) (Nagar et al., 2002) and the ligand structure from an X-ray structure (PDB ID: 2OIQ) (Seeliger et al., 2007). Each kinase-inhibitor complex was solvated with water molecules, where the number of water molecules was 7,698, 13,992, and 17,485 for Src-PP1, Src-Dasatinib, and Abl-Imatinib, respectively. The size of the simulation boxes for Src-Dasatinib and Abl-Imatinib was larger than for Src-PP1 because the farthest REUS replica (created by the US pulling simulations) was placed farther from the binding site (23 Å vs 18 Å). The systems were neutralized by adding sodium counter ions (six for Src-PP1 and Src-Dasatinib and eight for Abl-Imatinib). Each system was minimized for 1,000 steps while applying a positional restraint of 10.0 kcal/mol/Å² on protein backbone atoms. Then it was gradually heated to 310 K in the NVT ensemble for 105 ps, followed by equilibration in the NPT ensemble for 105 ps. Finally, the system was equilibrated for 1.05 ns in the NPT ensemble without restraining the protein atoms. Modeling was performed using AmberTools16 (Case et al., 2021).

2.3 MD Simulation

Simulations were performed using the GENESIS MD program (Jung et al., 2015; Kobayashi et al., 2017) version 2.0 beta (Jung et al., 2021). The AMBER ff99SB-ILDN (Hornak et al., 2006; Lindorff-Larsen et al., 2010) force field was used for the proteins, GAFF (Wang et al., 2004) (with AM1-BCC) for the ligands, and the TIP3P (Jorgensen et al., 1983) was used for water molecules.

TABLE 1 | System models and simulation details.

System	Src-PP1 (Forward/Reverse ^a)	Src-dasatinib	Abl-imatinib
Protein structure	1Y57 Cowan-Jacob et al. (2005)	1Y57 Cowan-Jacob et al. (2005)	1IEP Nagar et al. (2002)
Ligand structure	1QCF Schindler et al. (1999)	3G5D Getlik et al. (2009)	2OIQ Seeliger et al. (2007)
Number of atoms	27,549 (7,698 waters)	46,240 (13,992)	56,952 (17,485)
gREST solute temperature range, K	310–692	310–663	310–590
REUS distance range, Å	3.0–17.9/3.0–18.05 ^b	3.0–23.1	2.7–23.0
Simulation time per replica, ns	500	750	1,000

^aFor simulations that were initiated from REUS, replicas obtained from the forward and reverse simulations.

^bRange of values for forward simulations/range of values for reverse simulations.

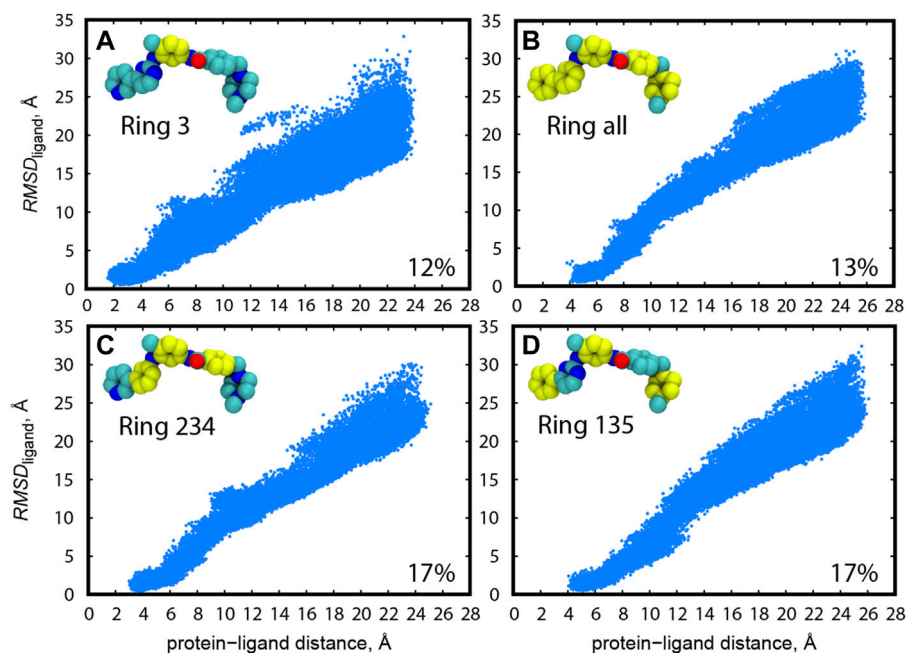


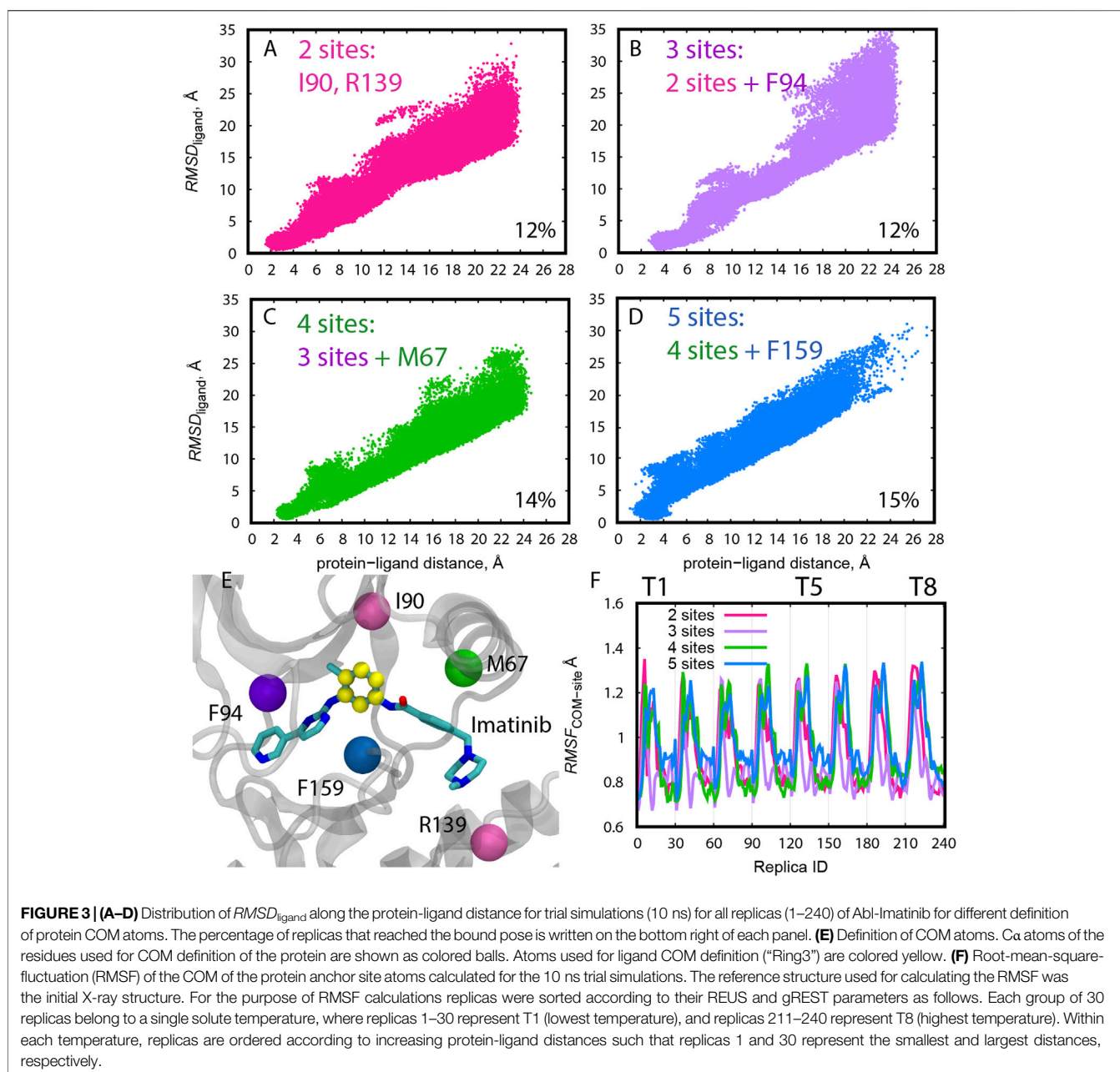
FIGURE 2 | Distribution of $RMSD_{ligand}$ along the protein-ligand distance for trial simulations (10 ns) for all replicas (1–240) of Abl-Imatinib for different definition of ligand COM atoms: (A) “Ring 3”, (B) “Ring all”, (C) “Ring 234”, and (D) “Ring 135”. Atoms used for ligand COM definitions are colored yellow. Ligand rings are numbered from 1 to 5, starting from the left. Protein atoms used for COM are backbone atoms of I90 and R139. The percentage of replicas that reached the bound pose is written on the bottom right of each panel.

Bonds involving hydrogen atoms were constrained using the SHAKE algorithm (Ryckaert et al., 1977). Water molecules were kept rigid using the SETTLE algorithm (Miyamoto and Kollman, 1992). Long-range electrostatic interactions were evaluated using the Particle-mesh Ewald summation (Darden et al., 1993; Essmann et al., 1995). The cutoff distance for non-bonded interaction was 8 Å. The NVT ensemble was used with the Bussi thermostat (Bussi et al., 2007) for keeping the temperature at 310 K, with a temperature coupling time of 5 ps. A timestep of 3.5 fs was used with the RESPA integrator (Tuckerman et al., 1992) and hydrogen mass repartitioning (HMR) (Feenstra et al., 1999) was applied on solute atoms with an HMR ratio of 3.0 (Jung et al., 2021).

Eight gREST replicas and 30 REUS replicas were used in the 2D-gREST/REUS simulation. In total, the number of replicas in each run was 240. All replicas were equilibrated for 1.05 ns without exchange attempts, followed by production runs.

Exchanges were attempted every 2.1 ps alternatively in the gREST and the REUS dimensions. The gREST/REUS simulations were performed for 500 ns per replica for Src-PP1, 750 ns per replica for Src-Dasatinib, and 1,000 ns per replica for Abl-imatinib. For Src-PP1, two simulations using initial replicas from either the “forward pull” or “reverse pull” US simulations were performed (referred to as Src-PP1 and Src-PP1-Rev, respectively). All simulation steps (except the US pulling simulations for creating the initial REUS replicas) were performed without any restraints on protein atoms. The total simulation time in the current work was 660 μ s. Frames for analysis were written every 10.5 ps. Simulations were performed on the supercomputer Fugaku¹ using 480 nodes.

¹<https://www.r-ccs.riken.jp/en/fugaku/>.



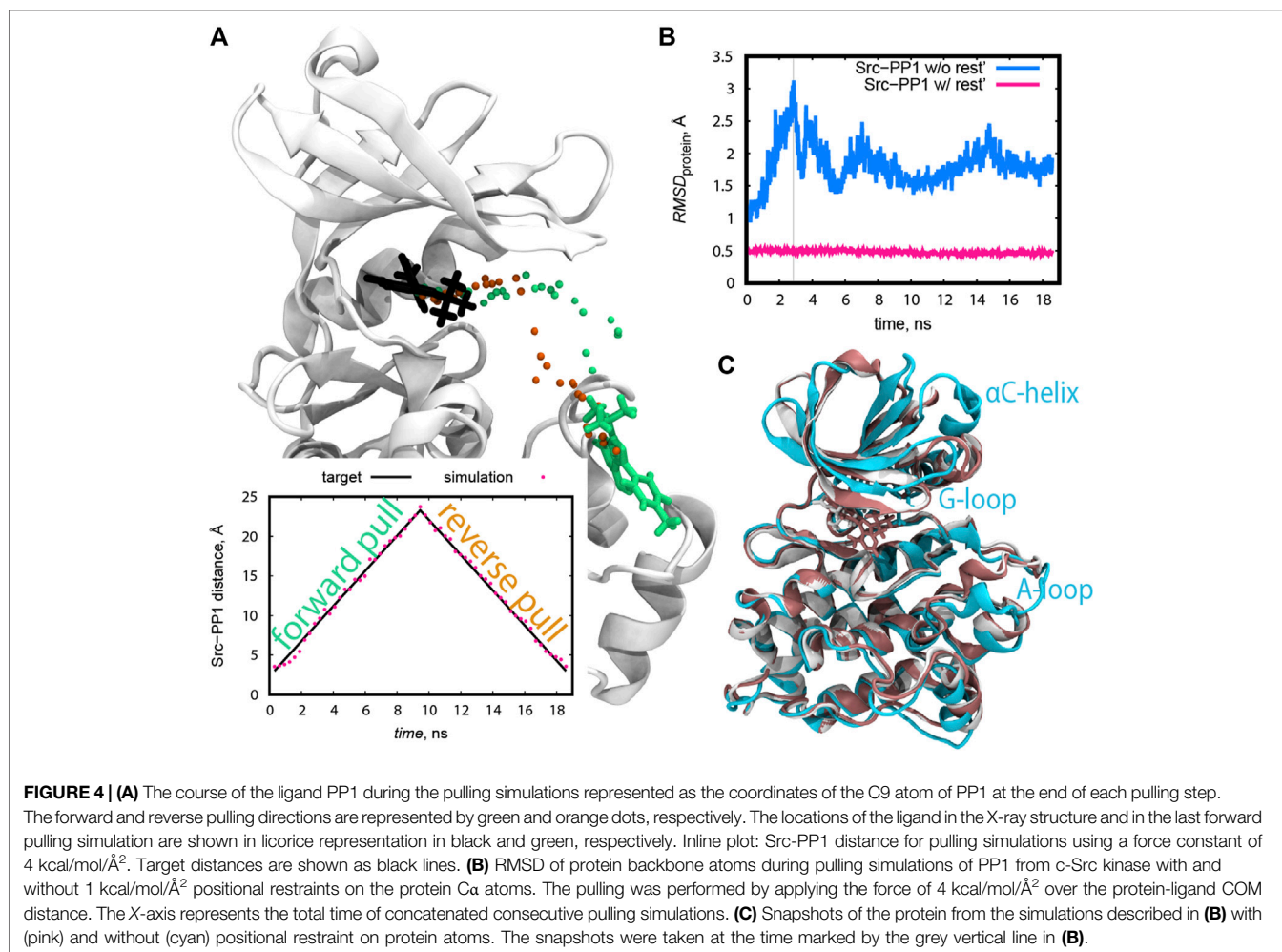
The GENESIS 2.0 beta version was optimized to run on Fugaku and obtained a speed of >50 ns/day. The details of the systems and the simulation conditions are summarized in **Table 1**.

3 RESULTS

3.1 Tuning the Definition of Protein-Ligand Distance in the REUS Dimension

As for Src-PP1 and Src-Dasatinib, we defined the protein-ligand distance using the “2 site” model (Ala 35 and Leu135) in c-Src kinase for the protein anchor sites and all the heavy atoms for calculating the ligand COM. Due to the inhibitor size and

flexibility, we tested multiple choices of the protein anchor sites and the ligand COMs for Abl-Imatinib by short (10 ns) gREST/REUS trial simulations. **Figure 2** shows the distribution of the ligand RMSD ($RMSD_{ligand}$) with respect to the bound pose of the X-ray crystal structure (2OIQ) (Seeliger et al., 2007) along the protein-ligand distance. As for the ligand COM, three rings (“Ring 135” and “Ring 234”), a single ring (“Ring 3”) and all rings (“Ring all”) were tested when we used “2 site” (Ile90 and Arg139) as the protein anchor site in c-Abl kinase. The probability of finding the bound pose, which we defined as the percentage of replicas that reached $RMSD_{ligand} < 1 \text{ \AA}$ at least once during the simulation, is also shown. An efficient pose sampling should give a linear correlation with narrow distribution. CVs



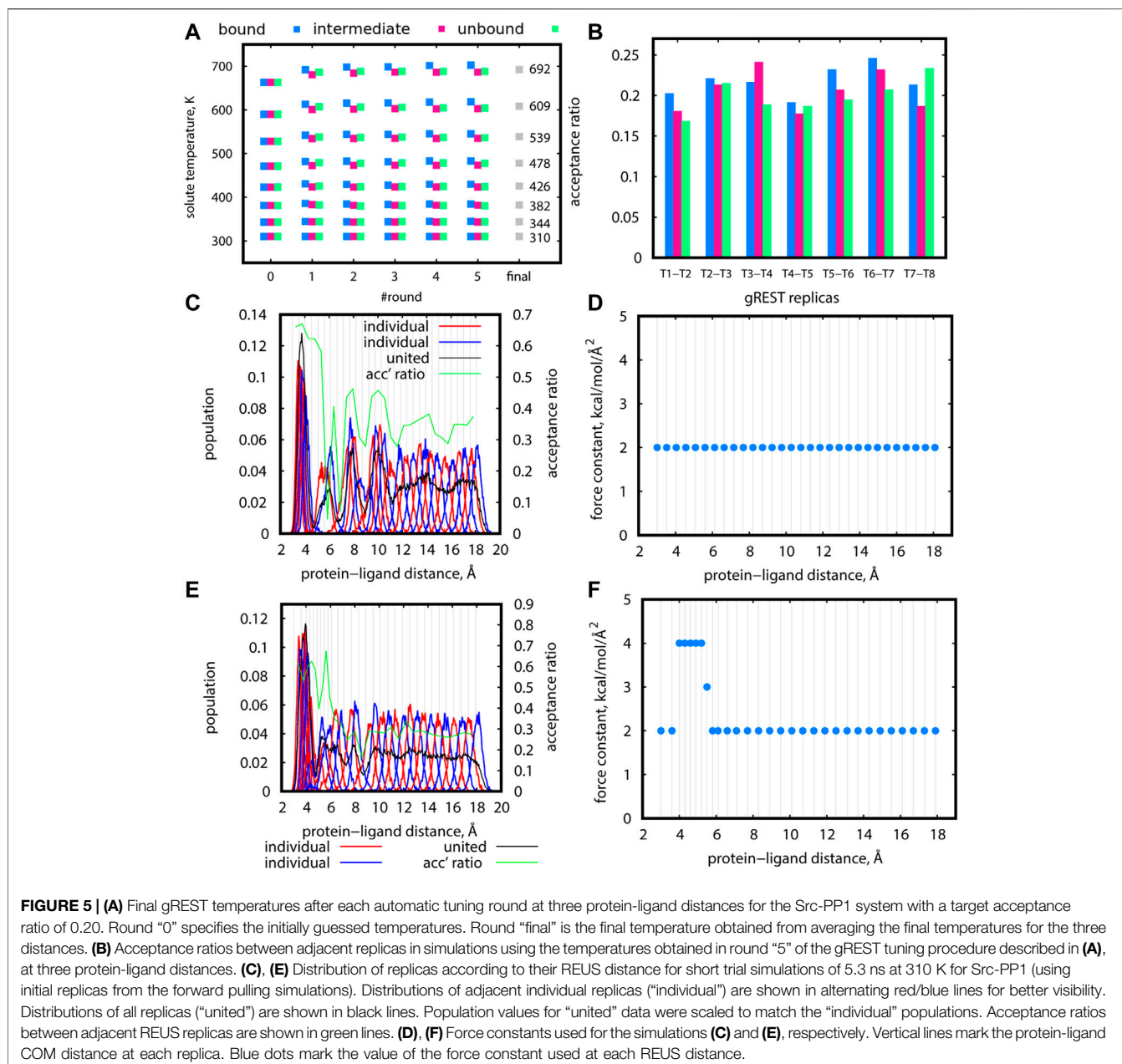
with multiple rings in Imatinib (“Ring all”, “Ring 234”, and “Ring 135”, **Figures 2B–D**) display linear and narrow distributions overall, compared to the single ring (“Ring 3”, **Figure 2A**). The latter could possess various conformations at the same distance, likely worsening the efficiency. “Ring 234” (**Figure 2C**) and “Ring 135” (**Figure 2D**) both have higher probabilities of finding the bound pose, while the latter shows slightly narrower distribution in the range of short protein-ligand distances. These results suggest that three anchor sites (a molecular center and both edges, “Ring 135”) is the practical choice for Abl-Imatinib.

At the same time, we tested four choices for the protein anchor sites using “Ring 3” as the ligand COM in Abl-Imatinib simulations (**Figure 3**). The overall distribution of $RMSD_{ligand}$ becomes narrow with increasing number of protein anchor sites. The probability of finding the bound pose is higher for “4 sites” and “5 sites” (14–15%, **Figures 3C,D**) than “2 sites” and “3 sites” (12%, **Figures 3A,B**), suggesting that two or three anchor sites are not sufficient to resolve the bound conformations of a ligand as large as Imatinib. “5 sites” produces a relatively wide distribution compared to “4 sites” at the bound region (~ 4 Å, **Figure 3D**). Increasing the number of residues in the protein anchor sites (**Figure 3E**) is effective for resolving the ligand position and

orientation but bears the risk of making protein anchor sites unstable. The COMs with “5 sites” indeed fluctuated more than the others through the replicas (**Figure 3F**). Consequently, the “4 sites–Ring135” pair was chosen as the best combination for Abl-Imatinib.

3.2 Preparation of Initial Structures in REUS From the Pulling Simulations

The initial structures along the protein-ligand distance CV were prepared from the following pulling simulations. Both forward (pulling away from the bound pose) and reverse directions (pulling back to the bound pose) were examined in the case of Src-PP1. The resulting initial pathways differed from each other (**Figure 4A**), suggesting that “dual direction pulling” could reduce the initial structure dependence and improve the convergence of the simulation results. To prepare the initial coordinates at each of the desired protein-ligand distances along the path in short simulations (9 ns per each), a rather strong force constant (4 kcal/mol/Å²) of the umbrella potential was required (**Figure 4A**). Note that the pulling simulations can introduce an artificial structure change in the protein. In the case of Src-PP1, the structures

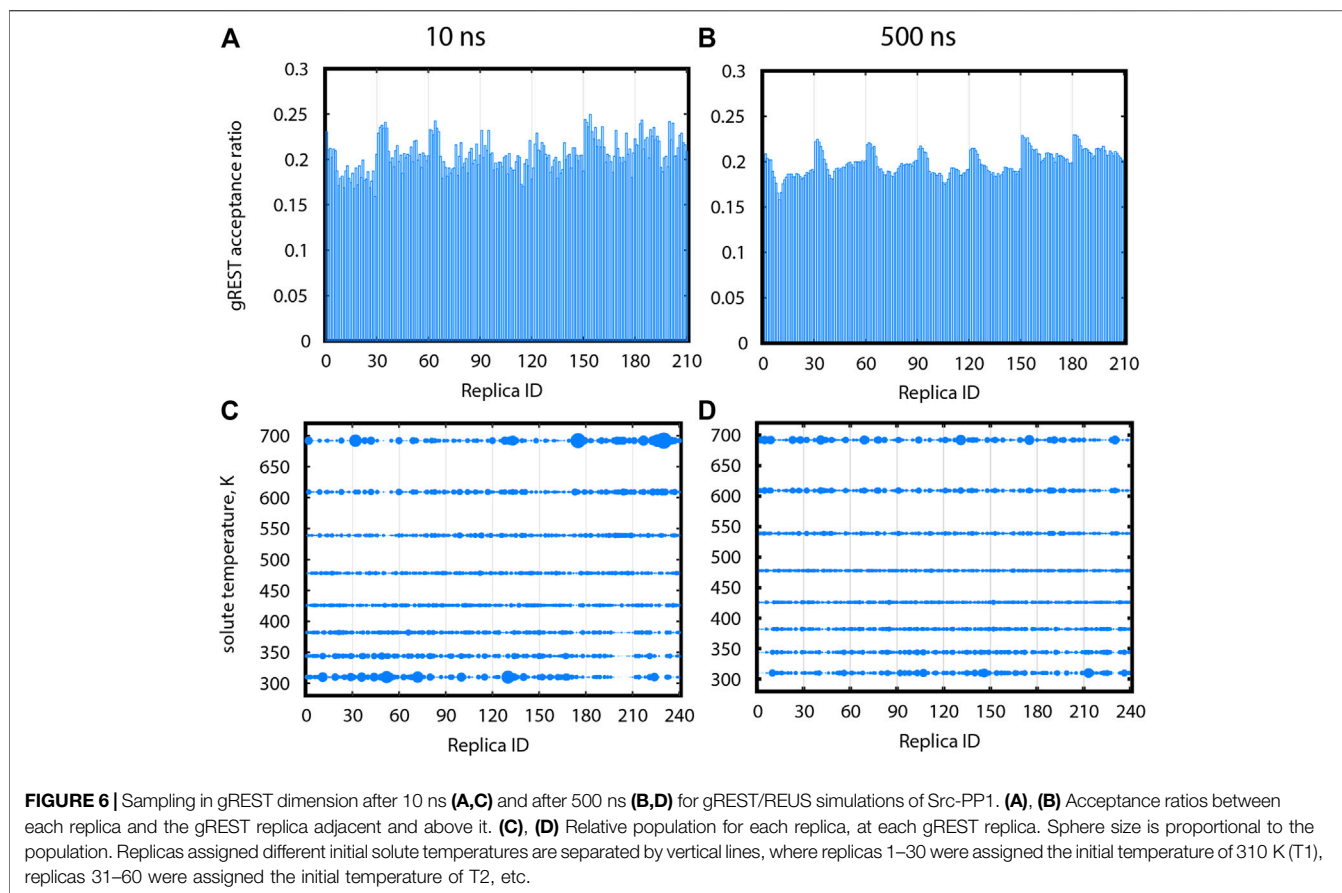


around the α C-helix, the G-loop, and the A-loop region significantly deviated from the X-ray crystal structure (Figures 4B,C). Since these regions directly relate to the binding mechanism, we also applied 1 kcal/mol/Å² restraints on the protein Ca atoms to avoid the artificial structure changes.

3.3 Tuning of Solute Temperatures in gREST

Solute temperatures in gREST could be determined rather effortlessly using the automatic tuning tool in GENESIS (Kobayashi et al., 2017). For Src-PP1, we set the initial temperatures in the range of 310–663 K, which is much narrower than our previous work (Re et al., 2019). This change markedly improved the sampling along the solute

temperature space. In addition, there are two key points in determining the temperatures. First, multiple rounds of tuning are desired. Figure 5A shows the solute temperatures determined at each of the five tuning rounds, where we set the final temperature at each round as the initial temperature for the subsequent round. The temperature values changed for the first few rounds and converged. Second, tuning at different protein-ligand distances are desired. For Src-PP1, we performed the tunings at protein-ligand distances of 3.0 Å (“bound”), 10.3 Å (“intermediate”), and 18.1 Å (“unbound”) distances. The resulting temperatures slightly differ in the three states (Figure 5A), and we therefore took their average at the final round to obtain the final set of solute temperatures. The resulting



solute temperatures provided uniform acceptance ratios along the gREST replicas (**Figure 5B**). Temperature tuning for Src-Dasatinib and Abl-Imatinib was performed using the same scheme. In practice, we manually changed the value of target acceptance ratios at each round for obtaining the final acceptance ratio of 0.2. The final sets of solute temperatures are listed in **Supplementary Table S1**.

3.4 Tuning of the REUS Parameters

The tuning of distance replicas and force constants in REUS simulations was more challenging, and several trial rounds were required for obtaining proper values. For Src-PP1, we started with even-spaced distance replicas and a uniform force constant of 2 kcal/mol/Å² at 310 K (**Figures 5C,D**). The sampled distance distribution was uneven. For example, the regions around 4 Å, 6 Å, and 9 Å are poorly covered, while there is an overlap in the region under 4 Å giving a large population in that region. The acceptance ratios around 6 Å drops to nearly zero, indicating almost no exchanges between replicas in that region. Accordingly, we put more replicas in poorly covered regions and set the force constants in those replicas to larger values (3 and 4 kcal/mol/Å²) (**Figures 5E,F**). This modification resulted in better coverage of the REUS space and acceptance ratios of above 0.2 for most replicas, ensuring the occurrence of replica exchanges throughout the REUS dimension. Nevertheless, we still observed an overly large population of replicas in the bound region of under 4 Å

alongside regions with poor coverage. The gap cannot be eliminated altogether since poorly covered regions represent high energy regions on the free energy landscape. This demonstrates the necessity of performing replica exchanges in two dimensions where increasing the temperature will facilitate the crossing of high energy barriers. The final REUS replica placements and force constant values are given in **Supplementary Table S1**.

3.5 Sampling Efficiency of the gREST/REUS Simulations After Parameters Tuning

As production runs, gREST/REUS simulations with 240 replicas were executed on the three systems, Src-PP1 (500 ns), Src-Dasatinib (750 ns), and Abl-Imatinib (1,000 ns), using the optimal parameters determined as described in previous sections. In the following sections, we quantify their sampling efficiencies in replica space and in the conformational space of the kinase-inhibitor complexes.

3.5.1 Random Walks in the gREST Dimension

Proper exchanges in the gREST dimension will allow replicas to go back and forth between low and high solute temperatures to sample high energy conformations. **Figures 6A,B** show acceptance ratios between adjacent gREST replicas for Src-PP1 after 10 and 500 ns, respectively. Acceptance ratios average around 0.2 as early as 10 ns.

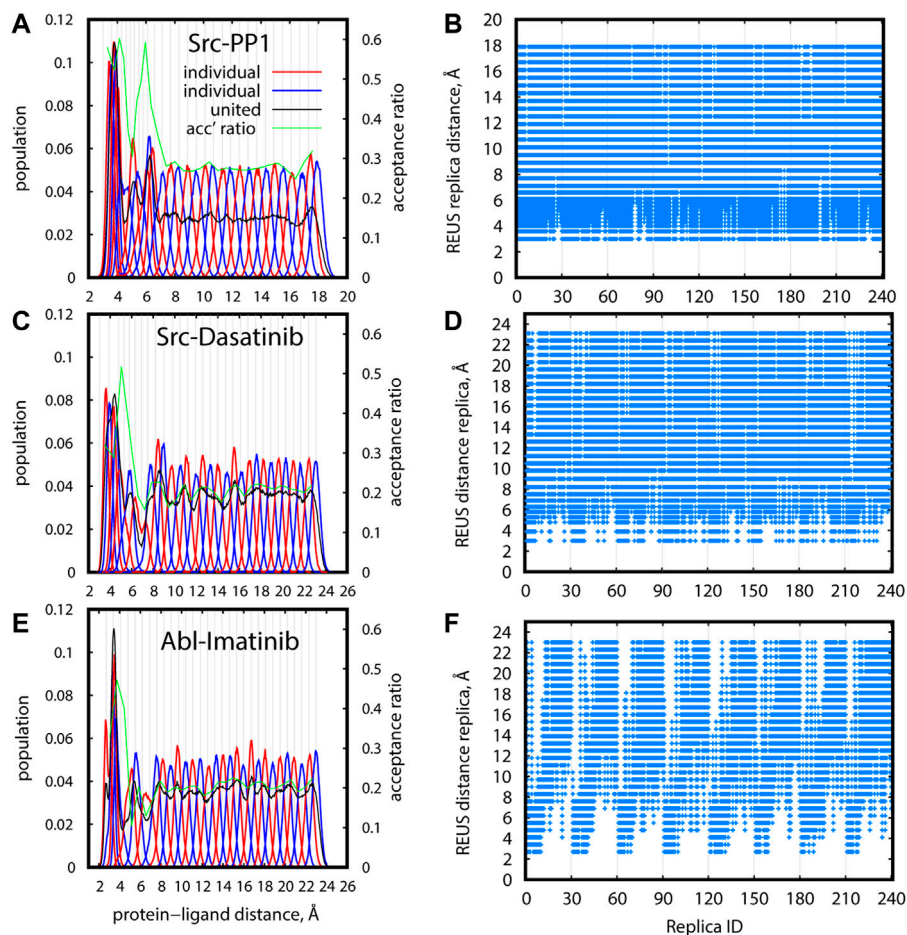


FIGURE 7 | Efficiency of sampling in REUS space for gREST/REUS simulations at 310 K for 500-ns Src-PP1 (**A,B**) 750-ns Src-Dasatinib (**C,D**), and 1,000-ns Abl-Imatinib (**E,F**). (**A**), (**C**), (**E**) Distribution of replicas according to their REUS distance. Distributions of adjacent individual replicas (“individual”) are shown in alternating red/blue lines for better visibility. Distributions of all replicas (“united”) are shown in black lines. Population values for “united” data were scaled to match the “individual” populations. Acceptance ratios between adjacent REUS replicas are shown in green lines. (**B**), (**D**), (**F**) REUS replicas visited at least once by individual replicas.

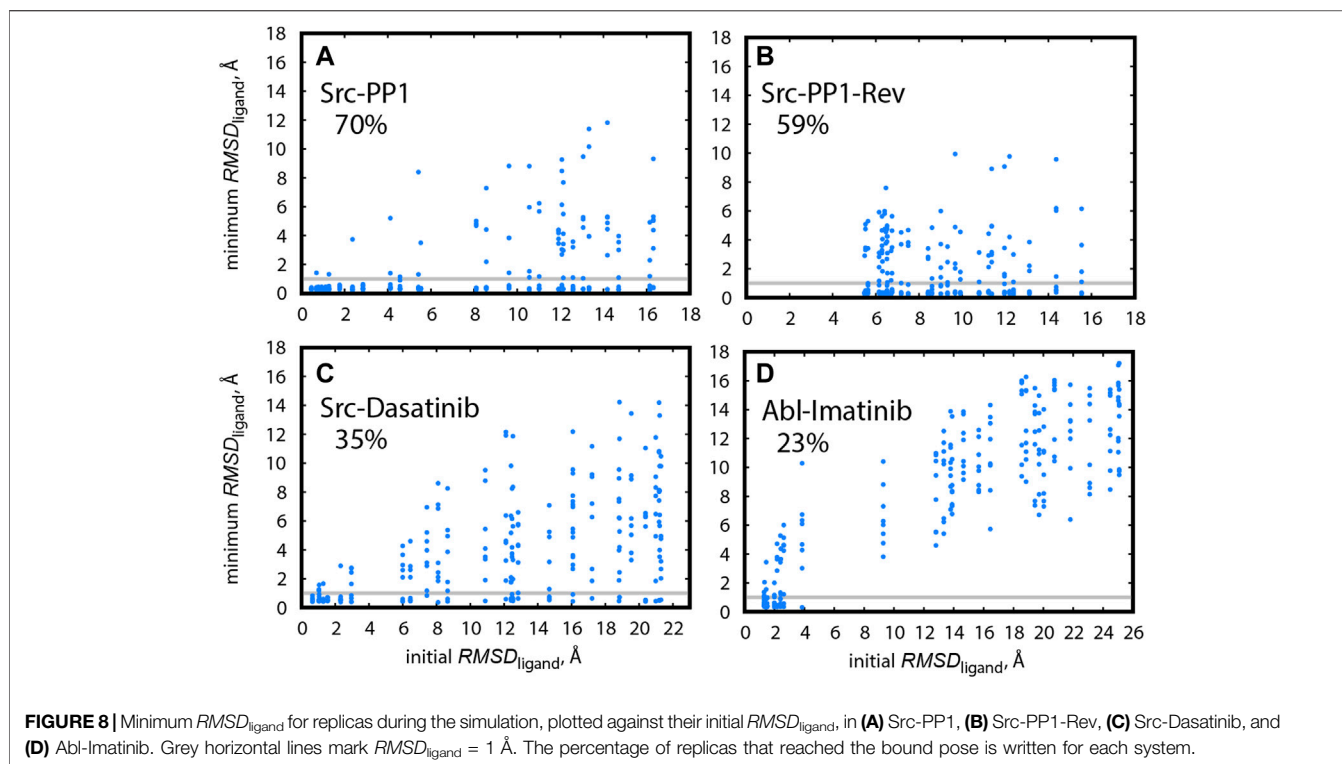
Figures 6C,D show the relative population at each solute temperature visited by individual replicas. Ideally, a uniform sampling is desired, in which each replica visits each temperature evenly. In a short simulation time (10 ns), all solute temperatures were already visited, although the replicas preferred the lowest and highest temperatures. After 500 ns, the sampling becomes more uniform, where the excessive population at the lowest and highest temperatures for some of the replicas flattens out. We found similar trends in the other two systems (Supplementary Figure S2), while the convergence becomes slower with increasing ligand size from PP1 to Dasatinib to Imatinib.

3.5.2 Random Walks in the REUS Dimension

The distribution of distance replicas along the REUS dimension at 310 K is shown in Figures 7A,C,E for Src-PP1, Src-Dasatinib, and Abl-Imatinib, respectively (Supplementary Figure S3A for Src-PP1-rev). All systems show a similar trend, where the population is large at short distances, drops once as the distance increases, and then converges to a constant value. Despite the drop in population, owing to the intensive tuning of the replica

parameters, all regions were sampled to an acceptable extent, maintaining a constant overlap and good acceptance ratios between adjacent replicas. For example, in the case of Src-Dasatinib, the initial lack of population in the region of protein-ligand distances 4–8 Å was gradually filled with increasing sampling time, and converged at 250 ns (Supplementary Figures S4A,C,E,G). The lack of population is much less significant at higher temperature replicas (Supplementary Figure S5), indicating that two-dimensional replica exchanges improve the sampling at 310 K.

Figures 7B,D,F (Supplementary Figure S3B) demonstrate the random walks along the REUS dimension. Each of the 240 replicas visited all REUS distances almost perfectly for Src-PP1 and moderately for Src-Dasatinib. In contrast, for Abl-Imatinib, random walks in the vicinity of each region are rather good but the overall random walks are not as efficient, namely, replicas which started at small distances could not reach far distances and vice versa (Figure 7F). This suggests that a large and flexible ligand can be trapped in the vicinity of its starting configuration due to either specific or non-specific interactions with the protein.



3.5.3 Finding the X-Ray Bound Pose in gREST/REUS Simulations

Finally, we compared the efficiencies of finding the X-ray bound pose for Src-PP1, Src-Dasatinib, and Abl-Imatinib. **Figure 8** shows the minimum $RMSD_{ligand}$ for individual replicas as a function of the initial $RMSD_{ligand}$ for all simulated systems. We define that a replica reached the bound pose if it had a $RMSD_{ligand} < 1 \text{ \AA}$ at least once during the simulation. The hit ratios along the sampling time are also summarized in **Supplementary Table S2**. For Src-PP1, 70% of the replicas, including those starting from far distances (large initial $RMSD_{ligand}$), found the X-ray bound pose (**Figure 8A**). Notably, the hit ratio was slightly low (59%) in the reverse pulling simulation (Src-PP1-Rev). We find that the initial $RMSD_{ligand}$ values are larger than 5 \AA , indicating that the Src-PP1-Rev simulation did not include the bound pose, which is nearly identical to the X-ray crystal structure, in its initial structures. Nevertheless, many replicas starting from large $RMSD_{ligand}$ values found the bound pose within 500 ns simulations, demonstrating that the gREST/REUS method can efficiently find an unknown bound pose.

The hit ratio drops to 35% for Src-Dasatinib (**Figure 8C**). However, a fraction of replicas with an initial $RMSD_{ligand}$ of $\sim 6 \text{ \AA}$ and above still finds the bound pose. For Abl-Imatinib (**Figure 8D**), which is the most challenging case, the hit ratio was only 23% even though its simulation time (1,000 ns) was the longest among the three systems. There is a gap in $RMSD_{ligand}$ values between $\sim 4 \text{ \AA}$ and $\sim 9 \text{ \AA}$. Unlike the case of Src-Dasatinib, the replicas above $\sim 9 \text{ \AA}$ cannot even reach the vicinity of the

binding site (**Figure 8D**). Therefore, the hit ratio stays around 20% after 250 ns and until 1,000 ns (**Supplementary Table S2**). These results suggest that Imatinib binding is a considerably rare event and that Imatinib can be trapped at various locations in the vicinity of the binding region before fully entering deep inside the binding pocket. **Supplementary Movies S1–S3** show binding events for a single replica for Src-PP1, Src-Dasatinib, and Abl-Imatinib, respectively, and demonstrate the difference in the efficiency of finding the bound pose. Whereas for Src-PP1 the ligand binds and unbinds several times during 500 ns, for Src-Dasatinib, a single binding event of a replica that started from a far distance is observed after $\sim 550 \text{ ns}$, and for Abl-Imatinib, a replica that started from an intermediate distance binds after $\sim 250 \text{ ns}$ and does not leave the binding site during the rest of the simulation time.

Here we followed the definition of Re et al. (2019) for hitting the bound pose, who deliberately set a strict cutoff of $RMSD_{ligand} < 1 \text{ \AA}$. We could set the cutoff slightly larger (for example 1.5 \AA) to consider ligand fluctuations around the bound pose. In this case, we obtain hit ratios of 75, 65, 38, and 27% for Src-PP1, Src-PP1-rev, Src-Dasatinib, and Abl-Imatinib, respectively.

4 DISCUSSION AND CONCLUSION

In this work, we described a step-by-step procedure for obtaining the optimal parameter settings for efficient gREST/REUS simulations of protein-ligand binding. The protocol, which was demonstrated here for three kinase-inhibitor systems, was

validated through an extensive analysis of sampling efficiency based on a total of 660 μ s of simulation time and can be applied to protein-ligand systems in general. We demonstrated that while the determination of gREST parameters is rather straightforward and nearly automatic, a particular care is needed in the determination of REUS parameters. First, a proper definition of the protein-ligand distance as the REUS CV, and second, careful tuning of replica space and force constants. Both of these practices can enhance the sampling efficiency. Taking care of these points, gREST/REUS simulations can sample binding events with high statistical accuracy and the obtained trajectories can be used to characterize binding poses and pathways on the free-energy landscape.

The use of protein-ligand distance as CV is a common practice for simulating binding events. Typically, the distance is determined using the COMs of the binding site and the ligand. For flexible ligands with molecular weight of few hundreds, as in the case of Imatinib, the determination of the CV significantly affects replica exchanges in REUS dimension. A lesson from this work is that each COM of the binding site and the ligand should be determined using multiple anchor sites for taking the flexibilities and orientation into account. This is because the flexible ligand can interact with the protein in different conformations and at different parts of the molecule. Even with a proper definition of the protein-ligand distance and well-tuned REUS parameters (replica spacing and force constants), the realization of constant acceptance ratios throughout the REUS dimension is quite difficult as shown for Imatinib. Here, we must add that applying too stiff umbrella potentials during the pulling simulations for obtaining the initial REUS replicas or during the REUS simulation may affect the obtained binding pathways. Thus, we must find the right balance of parameters that will not excessively bias the simulation but will still result in efficient sampling.

We showed that gREST/REUS can fill this gap with the aid of solute temperature exchange. Our results justify performing exchanges in two dimensions while using non-negligible computational resources. Using this protocol, protein-ligand binding simulations, in particular ligands or inhibitors of small or medium sizes, would be successfully performed on massively parallel supercomputers or GPU clusters.

Although good random walks in the replica space were observed in all three cases, simulation results of Abl-Imatinib suggest that efficient conformational sampling of Imatinib around the binding site of c-Abl kinase is still challenging. Unlike for Src-PP1 and Src-Dasatinib simulations, we could not observe many binding/unbinding events for Imatinib, especially of replicas initiated from far distances. Observing efficient random walks along the whole REUS range is important for visualizing the binding pathway. We learned that the problem is harder as the ligand size increased from PP1 (easy) to Dasatinib (moderate) to Imatinib (difficult), especially for obtaining the whole binding pathway. To further enhance the sampling for flexible ligands, consideration of a CV other than protein-ligand distance or an extension of the current scheme would be necessary. Considering the very slow unbinding rate of Imatinib, more drastic acceleration, such as simulations at higher solvent temperatures or enhancement of the c-Abl kinase

domain motions might be introduced in the gREST/REUS simulations.

Another practical drawback of the gREST/REUS ligand-binding simulations is that huge computational resources are required for them. In this study, we used 240 replicas in the 2D-REMD for each of the three cases. Without the use of Fugaku or other massively parallel supercomputers, it is not easy to access such huge resources. One way to overcome the problem is to replace gREST or REUS with other enhanced sampling methods with less computational costs. We previously developed GaREUS (Gaussian accelerated replica-exchange umbrella sampling) (Oshima et al., 2019) by replacing gREST in gREST/REUS into GaMD (Miao and McCammon, 2017). We were able to significantly reduce the number of replicas using GaREUS while keeping the sampling strategy and efficiency, because GaREUS requires the same number of replicas as 1D-REUS. The use of such low-cost enhanced sampling methods is necessary for investigating molecular mechanisms for many other kinase-inhibitor binding processes.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

Conceptualization, YS; Methodology, AS, SR and YS; Investigation, AS; Resources, YS; Writing–Original Draft, AS; Writing–Review and Editing, AS, SR and YS; Funding Acquisition, YS; Supervision, YS.

FUNDING

We used the computational resources provided by the HPCI System Research Project (Project ID: hp200129, hp200135, hp210172, and hp210177) and those in RIKEN Advanced Center for Computing and Communication (HOKUSAI BigWaterfall). This work was supported by MEXT/JSPS KAKENHI Grant Number 19H05645 (to YS), 21H05249 (to YS), 19K12229 (to SR), RIKEN pioneering projects in “Biology of Intracellular Environment,” “Dynamic Structural Biology,” and “Glycolipidlogue” (to YS), MEXT “Program for Promoting Research on the Supercomputer Fugaku (Biomolecular dynamics in a living cell (JPMXP1020200101)/MD-driven Precision Medicine (JPMXP1020200201)).”

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.878830/full#supplementary-material>

REFERENCES

- Bernetti, M., Cavalli, A., and Mollica, L. (2017). Protein-Ligand (Un)binding Kinetics as a New Paradigm for Drug Discovery at the Crossroad between Experiments and Modelling. *Med. Chem. Commun.* 8 (3), 534–550. doi:10.1039/c6md00581k
- Bruce, N. J., Ganotra, G. K., Kokh, D. B., Sadiq, S. K., and Wade, R. C. (2018). New Approaches for Computing Ligand-Receptor Binding Kinetics. *Curr. Opin. Struct. Biol.* 49, 1–10. doi:10.1016/j.sbi.2017.10.001
- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* 126 (1), 014101. doi:10.1063/1.2408420
- Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I., Brozell, S. R., Cerutti, D. S., et al. (2021). *Amber 2021*. San Francisco: University of California.
- Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J., et al. (2005). The crystal Structure of a C-Src Complex in an Active Conformation Suggests Possible Steps in C-Src Activation. *Structure*. 13 (6), 861–871. doi:10.1016/j.str.2005.03.012
- Darden, T., York, D., and Pedersen, L. (1993). Particle Mesh Ewald: AnN-Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 98 (12), 10089–10092. doi:10.1063/1.464397
- De Vivo, M., Masetti, M., Bottegoni, G., and Cavalli, A. (2016). Role of Molecular Dynamics and Related Methods in Drug Discovery. *J. Med. Chem.* 59 (9), 4035–4061. doi:10.1021/acs.jmedchem.5b01684
- Dickson, A., Tiwary, P., and Vashisth, H. (2017). Kinetics of Ligand Binding through Advanced Computational Approaches: a Review. *Curr. Top. Med. Chem.* 17 (23), 2626–2641. doi:10.2174/1568026617666170414142908
- Dickson, A. (2018). Mapping the Ligand Binding Landscape. *Biophysical J.* 115 (9), 1707–1719. doi:10.1016/j.bpj.2018.09.021
- Dror, R. O., Pan, A. C., Arlow, D. H., Borhani, D. W., Maragakis, P., Shan, Y., et al. (2011). Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. U.S.A.* 108 (32), 13118–13123. doi:10.1073/pnas.1104614108
- Du, X., Li, Y., Xia, Y.-L., Ai, S.-M., Liang, J., Sang, P., et al. (2016). Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* 17 (2), 144. doi:10.3390/ijms17020144
- Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995). A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* 103 (19), 8577–8593. doi:10.1063/1.470117
- Feenstra, K. A., Hess, B., and Berendsen, H. J. C. (1999). Improving Efficiency of Large Time-Scale Molecular Dynamics Simulations of Hydrogen-Rich Systems. *J. Comput. Chem.* 20 (8), 786–798. doi:10.1002/(sici)1096-987x(199906)20:8<786::aid-jcc5>3.0.co;2-b
- Fukunishi, H., Watanabe, O., and Takada, S. (2002). On the Hamiltonian Replica Exchange Method for Efficient Sampling of Biomolecular Systems: Application to Protein Structure Prediction. *J. Chem. Phys.* 116 (20), 9058–9067. doi:10.1063/1.1472510
- Getlik, M., Grütter, C., Simard, J. R., Klüter, S., Rabiller, M., Rode, H. B., et al. (2009). Hybrid Compound Design to Overcome the Gatekeeper T338M Mutation in cSrc. *J. Med. Chem.* 52 (13), 3915–3926. doi:10.1021/jm9002928
- Gobbo, D., Piretti, V., Di Martino, R. M. C., Tripathi, S. K., Giabbai, B., Storici, P., et al. (2019). Investigating Drug-Target Residence Time in Kinases through Enhanced Sampling Simulations. *J. Chem. Theor. Comput.* 15 (8), 4646–4659. doi:10.1021/acs.jctc.9b00104
- Hénin, J., Lelièvre, T., Shirts, M. R., Valsson, O., and Delemotte, L. (2022). Enhanced Sampling Methods for Molecular Dynamics Simulations. arXiv preprint arXiv:2202.04164v1. doi:10.48550/arXiv.2202.04164
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of Multiple Amber Force fields and Development of Improved Protein Backbone Parameters. *Proteins*. 65 (3), 712–725. doi:10.1002/prot.21123
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79 (2), 926–935. doi:10.1063/1.445869
- Jung, J., Kasahara, K., Kobayashi, C., Oshima, H., Mori, T., and Sugita, Y. (2021). Optimized Hydrogen Mass Repartitioning Scheme Combined with Accurate Temperature/Pressure Evaluations for Thermodynamic and Kinetic Properties of Biological Systems. *J. Chem. Theor. Comput.* 17 (8), 5312–5321. doi:10.1021/acs.jctc.1c00185
- Jung, J., Mori, T., Kobayashi, C., Matsunaga, Y., Yoda, T., Feig, M., et al. (2015). GENESIS: a Hybrid-Parallel and Multi-Scale Molecular Dynamics Simulator with Enhanced Sampling Algorithms for Biomolecular and Cellular Simulations. *Wires Comput. Mol. Sci.* 5 (4), 310–323. doi:10.1002/wcms.1220
- Kamiya, M., and Sugita, Y. (2018). Flexible Selection of the Solute Region in Replica Exchange with Solute Tempering: Application to Protein-Folding Simulations. *J. Chem. Phys.* 149 (7), 072304. doi:10.1063/1.5016222
- Kobayashi, C., Jung, J., Matsunaga, Y., Mori, T., Ando, T., Tamura, K., et al. (2017). *GENESIS 1.1: A Hybrid-parallel Molecular Dynamics Simulator with Enhanced Sampling Algorithms on Multiple Computational Platforms*. Wiley Online Library.
- Kokubo, H., Tanaka, T., and Okamoto, Y. (2013). Two-dimensional Replica-Exchange Method for Predicting Protein-Ligand Binding Structures. *J. Comput. Chem.* 34 (30), 2601–2614. doi:10.1002/jcc.23427
- Koneru, J. K., Sinha, S., and Mondal, J. (2019). In Silico reoptimization of Binding Affinity and Drug-Resistance Circumvention Ability in Kinase Inhibitors: a Case Study with RL-45 and Src Kinase. *J. Phys. Chem. B*. 123 (31), 6664–6672. doi:10.1021/acs.jpcc.9b02883
- Lin, Y.-L., Meng, Y., Jiang, W., and Roux, B. (2013). Explaining Why Gleevec Is a Specific and Potent Inhibitor of Abl Kinase. *Proc. Natl. Acad. Sci. U.S.A.* 110 (5), 1664–1669. doi:10.1073/pnas.1214330110
- Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., et al. (2010). Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins*. 78 (8), 1950–1958. doi:10.1002/prot.22711
- Liu, P., Kim, B., Friesner, R. A., and Berne, B. J. (2005). Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water. *Proc. Natl. Acad. Sci. U.S.A.* 102 (39), 13749–13754. doi:10.1073/pnas.0506346102
- Meng, Y., Lin, Y.-L., and Roux, B. (2015). Computational Study of the “DFG-Flip” Conformational Transition in C-Abl and C-Src Tyrosine Kinases. *J. Phys. Chem. B*. 119 (4), 1443–1456. doi:10.1021/jp511792a
- Miao, Y., and McCammon, J. A. (2017). “Gaussian Accelerated Molecular Dynamics: Theory, Implementation, and Applications,” in *Annual Reports in Computational Chemistry* (Elsevier), 231–278. doi:10.1016/bs.arcc.2017.06.005
- Miyamoto, S., and Kollman, P. A. (1992). Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* 13 (8), 952–962. doi:10.1002/jcc.540130805
- Morando, M. A., Saladino, G., D’Amelio, N., Pucheta-Martinez, E., Lovera, S., Lelli, M., et al. (2016). Conformational Selection and Induced Fit Mechanisms in the Binding of an Anticancer Drug to the C-Src Kinase. *Sci. Rep.* 6 (1), 24439–9. doi:10.1038/srep24439
- Nagar, B., Bornmann, W. G., Pellicena, P., Schindler, T., Veach, D. R., Miller, W. T., et al. (2002). Crystal Structures of the Kinase Domain of C-Abl in Complex with the Small Molecule Inhibitors PD173955 and Imatinib (STI-571). *Cancer Res.* 62 (15), 4236–4243.
- Narayan, B., Buchete, N.-V., and Elber, R. (2021). Computer Simulations of the Dissociation Mechanism of Gleevec from Abl Kinase with Milestoning. *J. Phys. Chem. B*. 125 (22), 5706–5715. doi:10.1021/acs.jpcc.1c00264
- Narayan, B., Fathizadeh, A., Templeton, C., He, P., Arasteh, S., Elber, R., et al. (2020). The Transition between Active and Inactive Conformations of Abl Kinase Studied by Rock Climbing and Milestoning. *Biochim. Biophys. Acta (Bba) - Gen. Subjects*. 1864 (4), 129508. doi:10.1016/j.bbagen.2019.129508
- Niitsu, A., Re, S., Oshima, H., Kamiya, M., and Sugita, Y. (2019). De Novo prediction of Binders and Nonbinders for T4 Lysozyme by gREST Simulations. *J. Chem. Inf. Model.* 59 (9), 3879–3888. doi:10.1021/acs.jcim.9b00416
- Oshima, H., Re, S., and Sugita, Y. (2020). Prediction of Protein-Ligand Binding Pose and Affinity Using the gREST+FEPP Method. *J. Chem. Inf. Model.* 60 (11), 5382–5394. doi:10.1021/acs.jcim.0c00338
- Oshima, H., Re, S., and Sugita, Y. (2019). Replica-exchange Umbrella Sampling Combined with Gaussian Accelerated Molecular Dynamics for Free-Energy Calculation of Biomolecules. *J. Chem. Theor. Comput.* 15 (10), 5199–5208. doi:10.1021/acs.jctc.9b00761
- Paul, F., Thomas, T., and Roux, B. (2020). Diversity of Long-Lived Intermediates along the Binding Pathway of Imatinib to Abl Kinase Revealed by MD

- Simulations. *J. Chem. Theor. Comput.* 16 (12), 7852–7865. doi:10.1021/acs.jctc.0c00739
- Plattner, N., and Noé, F. (2015). Protein Conformational Plasticity and Complex Ligand-Binding Kinetics Explored by Atomistic Simulations and Markov Models. *Nat. Commun.* 6 (1), 7653–7710. doi:10.1038/ncomms8653
- Re, S., Oshima, H., Kasahara, K., Kamiya, M., and Sugita, Y. (2019). Encounter Complexes and Hidden Poses of Kinase-Inhibitor Binding on the Free-Energy Landscape. *Proc. Natl. Acad. Sci. U.S.A.* 116 (37), 18404–18409. doi:10.1073/pnas.1904707116
- Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* 23 (3), 327–341. doi:10.1016/0021-9991(77)90098-5
- Schindler, T., Sichi, F., Pico, A., Gazit, A., Levitzki, A., and Kuriyan, J. (1999). Crystal Structure of Hck in Complex with a Src Family-Selective Tyrosine Kinase Inhibitor. *Mol. Cell.* 3 (5), 639–648. doi:10.1016/s1097-2765(00)80357-3
- Schuetz, D. A., de Witte, W. E. A., Wong, Y. C., Knasmueller, B., Richter, L., Kokh, D. B., et al. (2017). Kinetics for Drug Discovery: an Industry-Driven Effort to Target Drug Residence Time. *Drug Discov. Today*. 22 (6), 896–911. doi:10.1016/j.drudis.2017.02.002
- Seeliger, M. A., Nagar, B., Frank, F., Cao, X., Henderson, M. N., and Kuriyan, J. (2007). c-Src Binds to the Cancer Drug Imatinib with an Inactive Abl/c-Kit Conformation and a Distributed Thermodynamic Penalty. *Structure*. 15 (3), 299–311. doi:10.1016/j.str.2007.01.015
- Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. (2011). How Does a Drug Molecule Find its Target Binding Site? *J. Am. Chem. Soc.* 133 (24), 9181–9183. doi:10.1021/ja202726y
- Shekhar, M., Smith, Z., Seeliger, M., and Tiwary, P. (2021). Protein Flexibility and Dissociation Pathway Differentiation Can Explain Onset of Resistance Mutations in Kinases. *BioRxiv*. doi:10.1101/2021.07.02.450932
- Silva, D.-A., Bowman, G. R., Sosa-Peinado, A., and Huang, X. (2011). A Role for Both Conformational Selection and Induced Fit in Ligand Binding by the Lao Protein. *Plos Comput. Biol.* 7 (5), e1002054. doi:10.1371/journal.pcbi.1002054
- Sohraby, F., Javaheri Moghadam, M., Aliyar, M., and Aryapour, H. (2020). A Boosted Unbiased Molecular Dynamics Method for Predicting Ligands Binding Mechanisms: Probing the Binding Pathway of Dasatinib to Src-Kinase. *Bioinformatics*. 36 (18), 4714–4720. doi:10.1093/bioinformatics/btaa565
- Spitaleri, A., Decherchi, S., Cavalli, A., and Rocchia, W. (2018). Fast Dynamic Docking Guided by Adaptive Electrostatic Bias: The MD-binding Approach. *J. Chem. Theor. Comput.* 14 (3), 1727–1736. doi:10.1021/acs.jctc.7b01088
- Spitaleri, A., Zia, S. R., Di Micco, P., Al-Lazikani, B., Soler, M. A., and Rocchia, W. (2020). Tuning Local Hydration Enables a Deeper Understanding of Protein-Ligand Binding: The PP1-Src Kinase Case. *J. Phys. Chem. Lett.* 12 (1), 49–58. doi:10.1021/acs.jpcllett.0c03075
- Sugita, Y., Kitao, A., and Okamoto, Y. (2000). Multidimensional Replica-Exchange Method for Free-Energy Calculations. *J. Chem. Phys.* 113 (15), 6042–6051. doi:10.1063/1.1308516
- Sugita, Y., and Okamoto, Y. (1999). Replica-exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* 314 (1-2), 141–151. doi:10.1016/s0009-2614(99)01123-9
- Terakawa, T., Kameda, T., and Takada, S. (2011). On Easy Implementation of a Variant of the Replica Exchange with Solute Tempering in GROMACS. *J. Comput. Chem.* 32 (7), 1228–1234. doi:10.1002/jcc.21703
- Tiwary, P., Mondal, J., and Berne, B. J. (2017). How and when Does an Anticancer Drug Leave its Binding Site? *Sci. Adv.* 3 (5), e1700014. doi:10.1126/sciadv.1700014
- Tran, D. P., Hata, H., Ogawa, T., Taira, Y., and Kitao, A. (2020). PaCS-MD/MSM: Parallel Cascade Selection Molecular Dynamic Simulation in Combination with Markov State Model as an Efficient Non-bias Sampling Method. *Ensemble*. 22 (2), 151–156. doi:10.11436/mssj.22.151
- Tuckerman, M., Berne, B. J., and Martyna, G. J. (1992). Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* 97 (3), 1990–2001. doi:10.1063/1.463137
- Valsson, O., Tiwary, P., and Parrinello, M. (2016). Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annu. Rev. Phys. Chem.* 67, 159–184. doi:10.1146/annurev-physchem-040215-112229
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25 (9), 1157–1174. doi:10.1002/jcc.20035
- Wang, L., Friesner, R. A., and Berne, B. J. (2011). Replica Exchange with Solute Scaling: a More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B*. 115 (30), 9431–9438. doi:10.1021/jp204407d
- Yang, L.-J., Zou, J., Xie, H.-Z., Li, L.-L., Wei, Y.-Q., and Yang, S.-Y. (2009). Steered Molecular Dynamics Simulations Reveal the Likelier Dissociation Pathway of Imatinib from its Targeting Kinases C-Kit and Abl. *PLoS One*. 4 (12), e8470. doi:10.1371/journal.pone.0008470
- Zhang, Q., Zhao, N., Meng, X., Yu, F., Yao, X., and Liu, H. (2022). The Prediction of Protein-Ligand Unbinding for Modern Drug Discovery. *Expert Opin. Drug Discov.* 17 (2), 191–205. doi:10.1080/17460441.2022.2002298

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Shinobu, Re and Sugita. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.