Accelerating Dissociative Events in Molecular Dynamics Simulations by Selective Potential Scaling

Indrajit Deb[1] **and Aaron T. Frank**[1,2*]

[1]Department of Biophysics, University of Michigan, Ann Arbor, Michigan 48109, USA
[2]Department of Biophysics and Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA
[*]afrankz@umich.edu

ABSTRACT

Molecular dynamics (or MD) simulations can be a powerful tool for modeling complex dissociative processes such as ligand unbinding. However, many biologically relevant dissociative processes occur on timescales that far exceed the timescales of typical MD simulations. Here, we implement and apply an enhanced sampling method in which specific energy terms in the potential energy function are selectively "scaled" to accelerate dissociative events during simulations. Using ligand unbinding as an example of a complex dissociative process, we selectively scaled-up ligand-water interactions in an attempt to increase the rate of ligand unbinding. By applying our selectively scaled MD (or ssMD) approach to three cyclin-dependent kinase 2 (CDK2)-inhibitor complexes, we were able to significantly accelerate ligand unbinding thereby allowing, in some cases, unbinding events to occur within as little as 2 ns. Moreover, we found that we could make realistic estimates of the unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) as well as the binding free energies (ΔG^{sim}) of the three inhibitors from our ssMD simulation data. To accomplish this, we employed a previously described Kramers'-based rate extrapolation (KRE) method and a newly described free energy extrapolation (FEE) method. Because our ssMD approach is general, it should find utility as an easy-to-deploy, enhanced sampling method for modeling other dissociative processes.

Introduction

Dissociation events, such as ligand unbinding, play a critical role in many cellular processes (1, 2). Understanding the microscopic details that govern these dissociative processes require biophysical tools that can monitor the structure of biological systems across a wide range of spatial and temporal scales. In principle, classical molecular dynamics (MD) simulations are well suited to study dissociative processes (3). MD simulations can be used to generate trajectories that describe the spatiotemporal properties of all atoms in a given system, leading to insights into the sequence of structural changes that precede dissociation. Trajectories generated using MD simulation can also be used to estimate the value of key kinetic and thermodynamic parameters associated with dissociative processes (4, 5).

The timescales over which biologically relevant dissociative processes occur, however, are significantly longer than the timescales that are accessible via molecular dynamics (MD) simulations. Currently, simulations on the microsecond (μs) timescale are routinely carried out, but many of the interesting dissociative processes that occur in biology occur on the millisecond (ms) timescale and beyond. The use of specialized computer hardware coupled with optimized software have significantly extend the timescales that are accessible during MD simulations (6–11), but access to such computers is still limited. Fortunately, many enhanced sampling algorithms have been developed that can be used to accelerate dissociative processes in multi-component biomolecular systems during simulations (12–18).

Of these, potential scaled MD simulation (scaled-MD) (19–21) is a simple and reaction coordinate-free enhanced sampling technique that increases the frequency of rare events during simulations by lowering potential energy barriers between metastable states. Typically, during scaled-MD simulations, the potential energy surface is uniformly scaled by a scaling-factor, α, where $0 < \alpha < 1$. Recently, scaled-MD simulations have been used to accelerate ligand unbinding during simulations and to estimate the relative ligand residence times or unbinding times (22–25). However, because the entire potential energy surface is scaled by α during scaled-MD simulations, conformational restraints are needed to maintain the structure of the protein components in the complex; without these restraints, the individual protein components would unfold (22).

Rather than uniformly scaling the entire potential energy surface, we hypothesized that dissociative processes could be accelerated by selectively scaling specific potential energy terms, in a manner similar to the selective accelerated molecular dynamics (selective aMD) approach (26), in which, a boosting potential is selectively applied to specific energy terms to enhance the sampling of important degrees of freedom. For instance, to accelerate ligand unbinding, we reasoned that the interactions between a ligand and water molecules could be scaled-up to strengthen ligand-water interactions. Scaling-up ligand-water interactions should accelerate unbinding by essentially flooding the binding pocket and washing the ligand out.

Here, we implemented and tested a selectively scaled MD (ssMD) simulation approach and used it to accelerate ligand unbinding during MD simulations. As a proof-of-concept, we used our ssMD approach to study the ligand unbinding of three cyclin-dependent kinase 2 (CDK2) inhibitors. Additionally, we used a previously described Kramers'-based rate extrapolation (KRE) scheme (27–29) and a newly described free energy extrapolation (FEE) scheme to estimate unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) and binding free energies (ΔG^{sim}) of the three CDK2 inhibitors on the unscaled potential from ssMD trajectories. We found that our ssMD approach could speed up unbinding by several orders of magnitude and the resulting simulation data could be used to general *unbiased* estimates of the $\tau^{\text{sim}}_{\text{unbind}}$ and ΔG^{sim} on the unscaled potential.

Methods

Scaled MD simulations have previously been used to accelerate ligand unbinding (22). However, low values of α cause unfolding, which necessitates the use of conformational restraints to maintain the internal structure of the biomolecule(s) in the system. What follows is a description of a simple selective scaling approach that avoids the unfolding problem thus allowing reaction coordinate-free and restraint-free simulation of dissociative processes like ligand unbinding.

Selectively Scaling the Potential Energy Surface

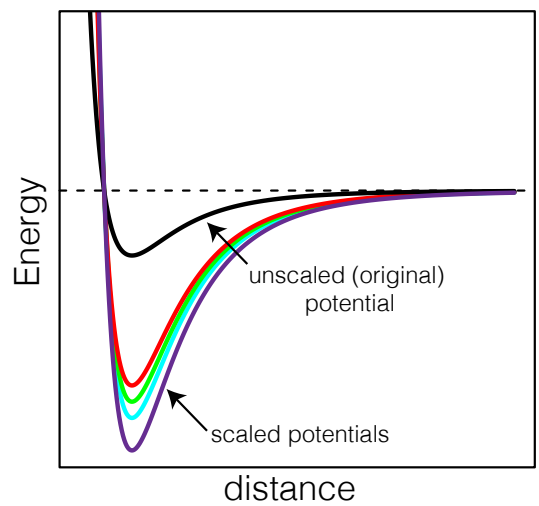

Figure 1. To accelerate ligand unbinding during simulations, ligand-water interactions are strengthened by scaling the well-depth (ε) of the Lennard-Jones potential between ligand and water atoms by a scaling-factor (α'). Shown is a series of scaled potentials (*colored*) and the original, unscaled potential (*black*).

Rather than uniformly scaling the potential energy surface, our selective scaling approach involves leaving a subset of the energy terms in the potential, V_0, unscaled, while scaling the remaining energy terms, V_1, by some factor, α'. To accelerate ligand unbinding, for example, a selective scaling approach could be implemented in which V_1 corresponded to the Lennard-Jones potential between ligand and water atoms. We reasoned that by strengthening or scaling up ligand-water interactions, that is, setting $\alpha' > 1$ (Figure 1), the frequency at which ligand-protein contacts are exchanged for ligand-water contacts would be increased and in so doing ligand unbinding would be accelerated. Formally, this selective scaling could be accomplished by using a modified-Lennard-Jones potential in which the Lennard-Jones potential, V_{ij} (Figure 1), between atoms i and j is given by

$$V_{ij} = \alpha'_{ij}\left(\varepsilon_{ij}\left[\left(\frac{r_{ij,\text{min}}}{r_{ij}}\right)^{12} - \left(\frac{r_{ij,\text{min}}}{r_{ij}}\right)^{6}\right]\right) \tag{1}$$

where ε_{ij} is the well-depth, α'_{ij} is the scaling factor, $r_{ij,\text{min}}$ is the distance at which the potential reaches its minimum, and r_{ij} is the distance between atom i and atom j. To selectively scale the interactions between ligand and water atoms using Eq. 1, ε_{ij} can be scaled by $\alpha' \neq 1$ for all pairwise interactions between atoms in the ligand and atoms in surrounding water molecules while ε_{ij} remains unscaled for all other pairwise interactions (i.e., $\alpha'_{ij} = 1$).

Selectively Scaled Molecular Dynamics (ssMD) Simulations of Cyclin-Dependent Kinase 2 (CDK2)-Inhibitor Complexes

Selectively scaled MD (ssMD) simulations, which employed the modified (or scaled) Lennard-Jones potential (Eq. 1), were used to simulate three CDK2-inhibitor complexes using the CHARMM simulation package (30). The three inhibitors are shown in Figure 2 and are referred to as inhibitor **1**, **2**, and **3**, respectively. The initial structures for the simulations of CDK2 in complex with inhibitors **1**, **2**, and **3** were obtained from PDB via the accession codes 5IEY, 5IEX, and 5IEV, respectively (31). All the missing loop residues in CDK2 of the initial structures were added using MODELLER 9.18 (32) within UCSF Chimera 1.11.2 (33) interface to built the starting structures. The simulation systems and the initial CHARMM simulation scripts were generated using the starting structures in CHARMM-GUI webserver (34, 35). For CDK2, we used the CHARMM36 protein force field (36) and for the CDK2 inhibitors, we used the CHARMM general force field (CGenFF) (37). The CDK2-inhibitor complexes were solvated in a cubic TIP3P (38) water box maintaining a minimum distance of 10 Å between any solute atom and the edge of the solvent box. Neutralizing ions were added to obtain zero net-charge of the solvated systems. Each system was then energy minimized by 200 steps of steepest descent (SD) followed by 500 steps of adopted basis Newton-Raphson (ABNR) algorithms, and equilibrated in the NVT ensemble with Langevin integrator for 2 ns with a timestep of 1 fs. During the energy minimization and equilibration steps, the sidechain heavy atoms of the protein were harmonically restrained with a force constant of 0.1 kcal/mol/Å^2 and the backbone heavy atoms of the protein and the heavy atoms of the ligand were harmonically restrained with a force constant of 1.0 kcal/mol/Å^2. ssMD simulations of the three CDK2-inhibitor complexes were then initiated from the equilibrated structures. In these ssMD simulations, the Lennard-Jones potential associated with all inhibitor-water pairs were scaled with the $\alpha'_{i,j}$ set to 3.00, 3.25, 3.50, and 4.00, respectively (Eq. 1). To selectively scale the Lennard-Jones potential associated with all inhibitor-water pairs, the well-depth (ε_{ij}) of all unique ligand-water atom type pairs were scaled using the NBFIX facility in CHARMM. For $\alpha' = 3.00$ and 3.25, 50 independent 50 ns ssMD trajectories were generated and for $\alpha' = 3.50$ and 4.00, 50 independent 20 ns ssMD trajectories were generated. All energy minimization, equilibration, and production runs were carried out using graphics processing unit (GPU) through CHARMM/OpenMM interface. Periodic boundary conditions (PBC) were applied during all the calculations. Non-bonded interactions were calculated using 16.0 Å cutoff and long-range electrostatics were treated using particle mesh Ewald (PME) method (39). All bonds involving hydrogens were constrained using SHAKE algorithm (40) during both the equilibration and production run simulations. During the production run simulations, constant pressure/constant temperature (CPT) Langevin dynamics were propagated at 303.15 K with a timestep of 2 fs with Monte Carlo (MC) barostat turned on at a reference pressure of 1 atm and sampled at a frequency of 15. The Langevin friction coefficient was set to 10 ps^{-1} for both the equilibration and production run simulations.

Estimating Unbinding Times From Selectively Scaled Molecular Dynamics (ssMD) Simulations Using Kramers'-based Rate Extrapolation (KRE)

To estimate the unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) on the unscaled ($\alpha' = 1$) potential from ssMD simulations, we used a previously described Kramers'-based rate extrapolation (KRE) method (41). Extrapolation approaches have been used previously to estimate of kinetic parameters from simulation data generated on modified potentials (27, 29, 42). In the KRE method, the Kramers' rate equation is applied to a potential surface scaled by a factor α to estimate the mean first passage time, $\tau(\alpha)$, for a transition from a reactant region ($\mathscr{F}$) to a product region ($\mathscr{U}$) using the expression:

$$\tau(\alpha) \cong \frac{\tau(1)}{\alpha^2}\left(\exp\left(\frac{\Delta H^{\dagger}(\alpha-1)}{k_B T}\right)\right) \tag{2}$$

where $\tau(1)$ is the mean first passage for the transition from $\mathscr{F}$ to $\mathscr{U}$ on the unscaled potential (i.e., when α = 1), k_B is the Boltzmann constant, T is temperature, and $\Delta H^{\dagger}$ is the activation energy (or barrier height) for the transition. Eq 2 is applicable in the high barrier, high friction limit ($\gg k_B T$) (28).

For ligand unbinding, we define $\mathscr{F}$ and $\mathscr{U}$ as the ligand-bound and unbound states, respectively, and we assume that ligand unbinding is a first-order transition governed by a single high barrier ($\gg k_B T$). By determining unbinding times, $\tau^{\text{sim}}_{\text{unbind}}$, as function of α (i.e., $\tau(\alpha)^{\text{sim}}_{\text{unbind}}$), and then fitting $\tau(\alpha)^{\text{sim}}_{\text{unbind}}$ to Eq. 2, the $\tau(\alpha = 1)^{\text{sim}}_{\text{unbind}}$ and the $\Delta H^{\dagger}_{\text{unbind}}$ for ligand unbinding on the unscaled ($\alpha = 1$) potential can be determined.

To estimate the $\tau^{\text{sim}}_{\text{unbind}}$ (i.e., the unbinding times on the unscaled potential), we first identified the set of native contacts $\{r_{ij}\}$ between CDK2 and inhibitors **1**, **2**, and **3**, respectively; a contact was defined as native if the distance between atom i in CDK2 and atom j in the inhibitor (r_{ij}) was $\leq$5 Å in the initial (crystal) structure. We then defined $\mathscr{F}$ and $\mathscr{U}$ as the region in which the shortest native contact r_{ij} was <10 Å and $\geq$10 Å, respectively. Next, from the kth simulation in which the well-depth (ε_{ij}) in the Lennard-Jones potential for all unique ligand-water interactions was selectively scaled by $\alpha' = 3.00$, 3.25, 3.50, or 4.00, respectively, $\tau(\alpha')^{\text{sim}}_{k,\text{unbind}}$ was determined as the transition from $\mathscr{F}$ to $\mathscr{U}$. For the mean unbinding times each α' over the set of independent simulations (i.e., $\tau(\alpha')^{\text{sim}}_{\text{unbind}} = \langle\tau(\alpha')^{\text{sim}}_{k,\text{unbind}}\rangle$) were then computed.

To fit Eq. 2, we used a weighted non-linear least-squared fitting technique using the PORT algorithm implemented in the `nls` function in `R` (43). Here the error function used to fit Eq. 2 was:

$$\chi^2 = \sum_{i=1}^{4} w(\alpha_i)^2 \times \left(\langle \tau(\alpha_i)^{\text{sim}}_{\text{unbind}} \rangle - \tau(\alpha_i)^{\text{fit}}_{\text{unbind}}\right)^2 \tag{3}$$

where the index i runs over the series of ssMD simulations; $\alpha_i = 1/\alpha'_i$, where α'_i is equal to 3.00, 3.25, 3.50, and 4.00, respectively; $\langle \tau(\alpha_i)^{\text{sim}}_{\text{unbind}} \rangle$ is the mean unbinding time calculated over the set of independent ssMD simulations associated with the scaling factor α_i; $\tau(\alpha_i)^{\text{fit}}_{\text{unbind}}$ is the corresponding unbinding time calculated using Eq. 2; and $w(\alpha_i)$ is the inverse of the standard deviation in $\tau(\alpha_i)^{\text{sim}}_{\text{unbind}}$ over the set of independent ssMD simulations.

For inhibitors **1**, **2**, and **3**, $\tau(\alpha = 1)^{\text{sim}}_{\text{unbind}}$ was determined as an average over 100 independent fits in which the starting values of the fitting parameters, $\tau(\alpha = 1)^{\text{sim}}_{\text{unbind}}$ and $\Delta H^{\dagger}(\alpha = 1)_{\text{unbind}}$, were randomly chosen. For each fit, the starting value for $\Delta H^{\dagger}(\alpha = 1)_{\text{unbind}}$ was from a uniform distribution with an interval of 0-100 kcal/mol and the starting value $\tau(\alpha = 1)^{\text{sim}}_{\text{unbind}}$ was from a uniform distribution with an interval of 10^2-10^{14} ns.

Estimating Binding Free-Energy From Selectively Scaled Molecular Dynamics (ssMD) Simulations Using Free Energy Extrapolation (FEE)

To estimate binding free energies (ΔG^{sim}) from ssMD simulation data, we employed a free energy extrapolation (FEE) approach. To achieve this, we noted that an extrapolation technique (similar to that embodied by Eq. 2) could be used to recover the configurational distribution, $p(\mathbf{r})$, on the unscaled potential from a set of configurational distribution, $p(\mathbf{r}, \alpha)$, determined from a set of ssMD simulations. Briefly, the $p(\mathbf{r})$ on the unscaled potential and $p(\mathbf{r}, \alpha)$ on a selectively scaled potential in which only a subset of the energy terms are scaled by a factor, α, is given by,

$$p(\mathbf{r}) = \exp\left(-\frac{V_0(\mathbf{r}) + V_1(\mathbf{r})}{k_B T}\right) \tag{4}$$

and

$$p(\mathbf{r}, \alpha) = \exp\left(-\frac{V_0(\mathbf{r}) + \alpha V_1(\mathbf{r})}{k_B T}\right) \tag{5}$$

where $\mathbf{r}$ is the configurational vector, V_0 is the sum of energy terms that remain unscaled, and V_1 is the sum of the energy terms that are selectively scaled by α. From Eq. 4 and 5, we arrive at the familiar energy reweighting formula

$$p(\mathbf{r}, \alpha) = p(\mathbf{r}) \exp\left(-\frac{V_1(\mathbf{r})}{k_B T}(\alpha - 1)\right). \tag{6}$$

As was done for $\tau(\alpha)$ (Eq. 2), $p(\mathbf{r}, \alpha)$ can be calculated from a series of ssMD simulations with different values of α and fitted to Eq. 6, where $p(\mathbf{r})$ and V_1 are treated as fitting parameters. Eq. 6 is in some sense an extrapolation-based realization of recently described population-based reweighting method (21) and is also related to the Hamiltonian mapping method (44). Using Eq 6, $p(\mathbf{r})$ could be determined and used, for example, to estimate the binding free-energy of the CDK2 inhibitors we examined in this study (see below).

To first test the validity of Eq. 6, we applied this FEE equation to a 1D potential given by:

$$V(x) = x^4 - 4x^2 + 2\alpha x \tag{7}$$

in which we selectively scaled the third term ($2x$) by α; this term modulates the asymmetry of the potential: if $2\alpha = 0$ the potential is symmetrical and if $2\alpha \neq 0$, $V(x)$ is asymmetrical. Using the analytical expressions for $p(x, \alpha)$ (Eq. 5) at α set to 0.50, 0.40, 0.20, 0.10, 0.00 respectively, we used Eq. 6 to estimate $p(x, \alpha = 1)$, that is the configurational distribution on the unscaled potential. To determine $p(x, \alpha = 1)$, we linearized Eq. 6 by taking the log of both sides, which, for this test, results in the expression:

$$\begin{aligned} \log(p(x, \alpha)) &= \log(p(x, \alpha = 1)) - \frac{V_1(x)}{k_B T}(\alpha - 1) \\ &= \underbrace{\log(p(x, \alpha = 1)) + \frac{V_1(x)}{k_B T}}_{\text{intercept}} - \alpha \underbrace{\frac{V_1(x)}{k_B T}}_{\text{slope}}. \end{aligned} \tag{8}$$

Eq. 8 was then fitted using linear regression using the analytical expressions for $\log(p(x, \alpha))$ at α set to 0.50, 0.40, 0.20, 0.10, 0.00. From the linear fit, $\log(p(x, \alpha = 1))$ was determined and then compared to the exact profile.

To estimate the binding free energies (ΔG^{sim}) of CDK2 inhibitor **1**, **2**, and **3**, using Eq. 6, that is, our FEE approach, we replaced the multi-dimensional configuration distribution, $p(\mathbf{r})$, in Eq. 6 with a 1-dimensional configurational distribution, $p(Q)$, where Q is an order parameter defined as the native contacts between CDK2 and the inhibitors. In this case, the V_1 corresponds to an "effective" energy. For inhibitors **1**, **2**, and **3**, the configurational distributions along Q, $p(Q,\alpha)$, was computed from the corresponding sets of ssMD simulations carried out at $\alpha = 1/\alpha'$, where α' is equal to 3.00, 3.25, 3.50, and 4.00, respectively. To estimate the binding free energies of CDK2 inhibitors on the unscaled potential ($\Delta G(Q,\alpha=1)^{\text{sim}}$), we directly extrapolated the free energy profiles along Q, $G(Q,\alpha)^{\text{sim}}$, using the expression

$$\begin{aligned} -k_B T \log(p(Q,\alpha)) &= -k_B T \log(p(Q,\alpha=1)) - V_1(Q)(\alpha-1) \\ G(Q,\alpha)^{\text{sim}} &= \underbrace{G(Q,\alpha=1)^{\text{sim}} + V_1(Q)}_{\text{intercept}} - \alpha \underbrace{V_1(Q)}_{\text{slope}} \end{aligned} \tag{9}$$

where $G(Q,\alpha)^{\text{sim}}$ is the free energy along Q determined from the ssMD simulation data and $G(Q,\alpha=1)^{\text{sim}}$ corresponding to the free energy along Q on the unscaled potential. Eq. 9 is obtained by simply replacing x in Eq. 8 with Q and then multiplying both sides by $-k_B T$. By fitting Eq. 9 at a series of Q values, the extrapolated free energy profiles, $G(Q,\alpha=1)^{\text{sim}}$, was determined from the $G(Q,\alpha)^{\text{sim}}$ profiles of each CDK2 inhibitor. The binding free energy, ΔG^{sim}, for the three inhibitors was calculated and compared to their experimental values.

Results

To test our selectively scaled MD (ssMD) method, we studied the unbinding of three inhibitors of CDK2, an important cancer target (Figure 2). Inhibitors **1** (Figure 2A), **2** (Figure 2B), and **3** (Figure 2C) are three closely related inhibitors in which a hydrogen in **1** is replaced with a bromo group in **2** and a tri-fluoromethyl-group in **3**. Inhibitors **1**, **2**, and **3** exhibit experimental ligand residence times ($\tau^{\text{exp}}_{\text{unbind}}$) of ∼6.0, ∼763.2, and ∼7476.6 s, respectively (31). Given the long $\tau^{\text{exp}}_{\text{unbind}}$ of these inhibitors, these CDK2-inhibitor systems provided us with an ideal test set to assess the extent to which our ssMD approach could be used to accelerate simulations of dissociative processes such as ligand unbinding, as well as to assess the extent to which we could estimate the kinetics and thermodynamics of ligand unbinding from the resulting simulation data.

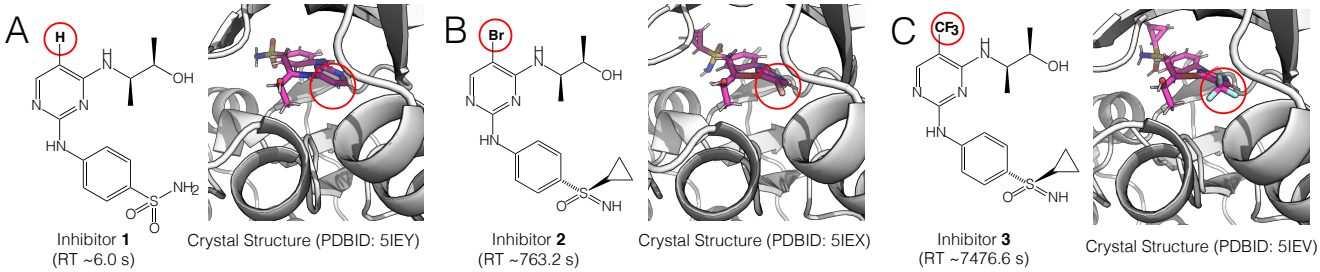

Figure 2. Shown are the three related CDK2 inhibitors simulated in this study: (A) Inhibitor **1** (B) Inhibitor **2**, and (C) Inhibitor **3**, respectively. Indicated are the PDBIDs for the crystal structures of each CDK2-inhibitor complex and their corresponding experimental residence times (RT).

α'	no. of simulations (inhibitor **1**, **2**, **3**)	simulation length (ns) (inhibitor **1**, **2**, **3**)	no. of unbinding events (inhibitor **1**, **2**, **3**)	wall time per simulation (hrs.)
1.00	20, 20, 20	250.0, 250.0, 250.0	5, 2, 0	∼115.0
3.00	50, 50, 50	50.0, 50.0, 50.0	49, 50, 43	∼28.0
3.25	50, 50, 50	50.0, 50.0, 50.0	50, 50, 49	∼28.0
3.50	50, 50, 50	20.0, 20.0, 20.0	50, 50, 46	∼11.2
4.00	50, 50, 50	20.0, 20.0, 20.0	50, 50, 49	∼11.2

Table 1. Comparison between conventional (unscaled) MD (α' = 1.00) and ssMD simulations carried out with α' set to 3.00, 3.25, 3.50, and 4.00, respectively. Listed are results for all three cyclin-dependent kinase 2 (CDK2)-inhibitor complexes (Figure 2).

Selectively Scaled MD (ssMD) Simulations of the Cyclin-Dependent Kinase 2 (CDK2)-Inhibitor Complexes. To confirm that the unbinding of inhibitors **1**, **2**, and **3** from CDK2 were indeed rare dissociative events, we first carried out conventional, unscaled MD simulations of CDK2 bound to these inhibitors (Figure 2). Briefly, for each of the CDK2-inhibitor systems, we carried out 20 independent simulations on the unscaled potential. During these sets of independent simulations, each of which were 250 ns long, we found that inhibitors typically remained bound to CDK2; for inhibitors **1**, **2**, and **3**, we observed 5, 2, and 0 unbinding events over the set of 20 simulations, respectively (Table 1). As such, on the timescale of these simulations, the unbinding of these CDK2 inhibitors were indeed rare events. These results confirmed that these three CDK2-inhibitor complexes were appropriate model systems for assessing the extent to which we could accelerate dissociation events using our ssMD approach.

Next, we carried out ssMD simulations of the three CDK2-inhibitor complexes. Briefly, in these simulations, we selectively scaled-up, and is so doing, strengthened the inhibitor-water interactions (see methods) by setting α' in Eq. 1 to 3.00, 3.25, 3.50, and 4.00, respectively. As highlighted in Table 1, we were able to observe significantly more unbinding events during our ssMD simulations than our unscaled simulations (Table 1). For all values of α', the fraction of unbinding events observed during each set of independent simulations was more than 85%, for all three inhibitors. For instance, with α' set to 3.00, we observed 49, 50, and 43 unbinding events during the 50 independent simulations for inhibitors **1**, **2**, and **3**, respectively. Similar results were obtained for α' = 3.25, 3.50, and 4.00.

Inhibitor	$\tau^{\text{sim}}_{\text{unbind}}(\alpha' = 3.00)$ (ns)	$\tau^{\text{sim}}_{\text{unbind}}(\alpha' = 3.25)$ (ns)	$\tau^{\text{sim}}_{\text{unbind}}(\alpha' = 3.50)$ (ns)	$\tau^{\text{sim}}_{\text{unbind}}(\alpha' = 4.00)$ (ns)
1	7.23 ± 5.53	5.01 ± 3.85	2.62 ± 1.29	1.65 ± 0.52
2	7.60 ± 6.34	5.82 ± 5.58	3.71 ± 1.95	2.61 ± 1.15
3	16.92 ± 13.38	11.20 ± 8.82	6.37 ± 3.71	3.44 ± 2.32

Table 2. Mean unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) for inhibitors **1**, **2**, and **3** computed from ssMD simulations carried out with α' set to 3.00, 3.25, 3.50, and 4.00, respectively.

Shown in Table 2 are the observed mean unbinding times, $\tau^{\text{sim}}_{\text{unbind}}$, for the inhibitors at each value of α'. For all the inhibitors, we observed a consistent decrease in $\tau^{\text{sim}}_{\text{unbind}}$ as α' increases from 3.00 to 4.00. When α' was set to 3.00, the mean $\tau^{\text{sim}}_{\text{unbind}}$ for the three inhibitors was 7.23, 7.60, and 16.92 ns, respectively. By comparison, with α' to 4.00, $\tau^{\text{sim}}_{\text{unbind}}$ was 1.65, 2.61, and 3.44 ns, respectively (Table 2). Collectively, these results demonstrate the remarkable acceleration in unbinding we achieved by employing our ssMD simulation approach. We note that we also observed a consistent decrease in the standard deviations of the mean $\tau^{\text{sim}}_{\text{unbind}}$ over the set of independent ssMD simulations as α' increases. The decrease in variance with the increase in α' is most likely a result of the increased smoothing of the effective CDK2-inhibitor interaction potential with α'.

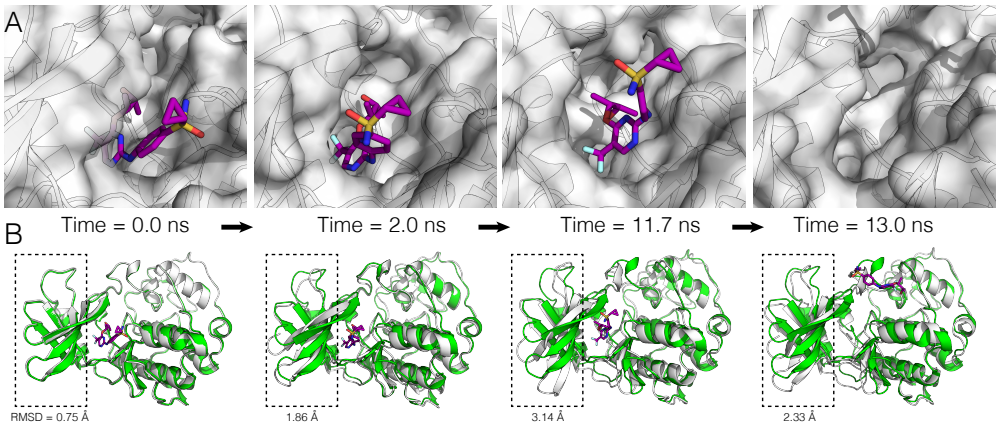

Figure 3. Snapshots from a representative unbinding trajectory of inhibitor **3** that was generated with α' set to 3.00. (A) Closeup surface view of the binding pocket and (B) global view of the complex. The β-rich region comprised of subdomains I-IV (residues 1-100), which undergoes a significant rearrangement during unbinding, is enclosed in the dashed box. Indicated below each dashed box is the RMSD for the residues in subdomains I-IV, which was calculated between the crystal structure (*green*) and the simulation structure (*white*) at t = 0.0, 2.0, 11.7, and 13.0 ns, respectively.

Shown in Figure 3 are snapshots from a representative unbinding trajectory of inhibitor **3**, which were generated with α' set to 3.00. This representative trajectory was identified as the trajectory whose native-contact distances most closely matched the

average native-contact distances over each frame across the 50 independent ssMD trajectories that were generated with α' set to 3.00. Analysis of this trajectory revealed an unbinding mode that was characterized by a progressive opening of the binding pocket (Figure 3A). The opening of the binding pocket appeared to be facilitated by a structural rearrangement in residues 1-100 (comprised of the β-rich sub-domains I-IV) (Figure 3B), which acted as a gate that controlled access to the binding pocket. In this example, unbinding occurred within 13 ns.

As an alternative to scaling up ligand-water interactions, we note that one could also selectively scale-down protein-ligand interactions in an attempt to speed up simulations. Indeed, scaling up the ligand-water interactions by α' can be thought of as effectively scaling down the ligand-protein interactions by $\alpha \equiv 1/\alpha'$. However, in preliminary studies, we discovered that though directly and selectively scaling-down protein-ligand interactions was effective at decoupling the ligand from the protein, in many cases, the ligand became trapped within the binding pocket waiting for the binding pocket to open. Selectively strengthening the ligand-water interactions avoids this problem because as ligand becomes increasingly hydrated, the protein responds to the larger effective size of the ligand by the opening-up the binding pocket, which allows the decoupled ligand to exit.

Estimating Unbinding Times From Selectively Scaled MD (ssMD) Simulations

Next, we explored whether we could generate unbiased estimates of the mean unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) from our ssMD simulation data. As our preliminary simulations on the unscaled potential of the three CDK2 inhibitors revealed, we were unable to generate a statistically significant number of unbinding trajectories (Table 1). From 3 $\times$ 20 independent simulations (each 250 ns long), we only observed 7 unbinding trajectories (Table 1). As such, we were interested in estimating, from our ssMD simulation data, the mean $\tau^{\text{sim}}_{\text{unbind}}$ associated with original unscaled potential.

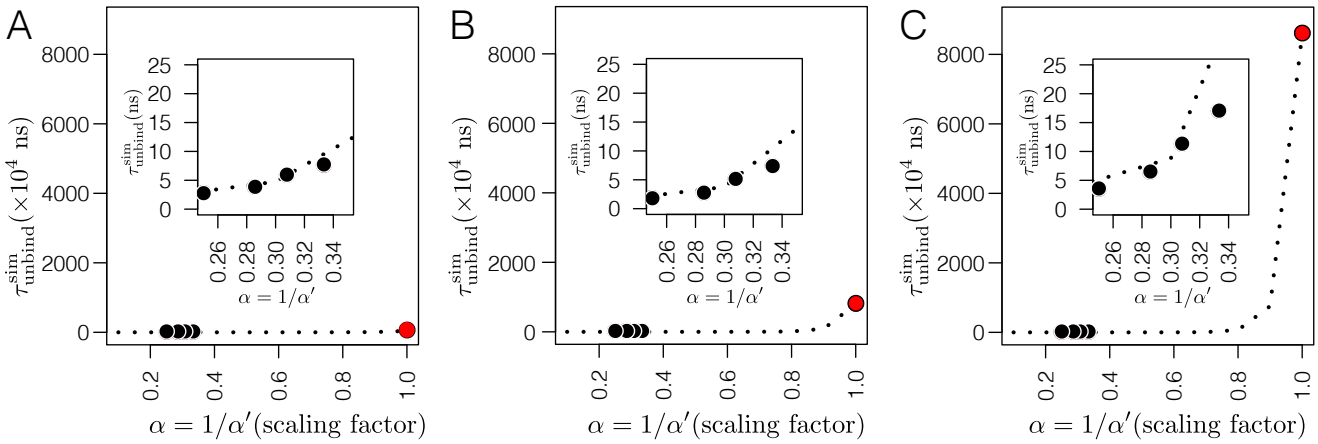

Figure 4. Kramers'-based rate extrapolation (KRE) results for inhibitors **1** (A), **2** (B), and **3** (C). Shown are plots of $\tau^{\text{sim}}_{\text{unbind}}$ versus α (where $\alpha = 1/\alpha'$). The *black circles* correspond to the mean unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$) obtained from our ssMD simulations carried out with α' = 3.00, 3.25, 3.50, and 4.00, respectively. The *red circles* correspond to the estimated $\tau^{\text{sim}}_{\text{unbind}}$ on the unscaled potential (i.e., α = 1 or equivalently α' = 1). Expanded $\tau^{\text{sim}}_{\text{unbind}}$ versus α plots are shown in inset, which highlight the exponential relationship between $\tau^{\text{sim}}_{\text{unbind}}$ and α obtained from our ssMD simulations (*black circles*).

To generate *unbiased* estimates of the mean unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$), that is, estimates of $\tau^{\text{sim}}_{\text{unbind}}$ on the unscaled (α'=1.00) potential, we employed a Kramers'-based rate extrapolation (KRE) method (Eq. 2; see Methods) (27, 28). In brief, from the ssMD simulations carried out with different α's, the corresponding mean unbinding times (i.e., $\tau(\alpha')^{\text{sim}}_{\text{unbind}}$) were fitted to the KRE equation (Eq. 2). The fitted KRE equation was then used to estimate the mean unbinding times on the unscaled potential (i.e., $\tau(1)^{\text{sim}}_{\text{unbind}} \equiv \tau^{\text{sim}}_{\text{unbind}}$). We note that because Eq. 2 is valid for $0 < \alpha < 1$, we carried out fitting and extrapolation with $\alpha \equiv 1/\alpha'$, such that at large values of α', $\alpha \to 0$.

For inhibitors **1**, **2**, and **3**, the predicted unbinding times on the unscaled potential, $\tau^{\text{sim}}_{\text{unbind}}$, were $\sim$40.4 (Figure 4A), $\sim$796.4 (Figure 4B), and $\sim$8588.4 $\times$ 10^4 ns (Figure 4C), respectively (Table 3). These estimates correspond to the mean estimated value over 100 independent fits to the KRE equation. To assess the robustness of the fits, we also determined the statistical errors in our estimates as the standard deviation over 100 independent fits. For inhibitors **1**, **2**, and **3**, the estimated errors were $\pm$0.0000055, $\pm$0.0002162, and $\pm$0.0013671 $\times 10^4$ ns, respectively. As such our estimates of the $\tau^{\text{sim}}_{\text{unbind}}$ for the three CD2K inhibitors were robust.

Inhibitor	$\tau^{\text{sim}}_{\text{unbind}}$ (10^4 ns)	Fitting Error (10^4 ns)	$\tau^{\text{exp}}_{\text{unbind}}$ (s)
1	40.4	$\pm$ 0.0000055	6.0
2	796.4	$\pm$ 0.0002162	763.2
3	8588.4	$\pm$ 0.0013671	7476.6

Table 3. Comparison between our estimates of the unbinding times on unscaled potential ($\tau^{\text{sim}}_{\text{unbind}}$) and the experimentally measured residence times ($\tau^{\text{exp}}_{\text{unbind}}$) for inhibitor **1**, **2**, and **3**, respectively. $\tau^{\text{sim}}_{\text{unbind}}$ correspond to a mean over 100 independent fits to the KRE equation (Eq. 2). The fitting errors in $\tau^{\text{sim}}_{\text{unbind}}$ correspond to the standard deviation in $\tau^{\text{sim}}_{\text{unbind}}$ over the 100 independent fits.

To assess the quality of our estimates, we compared the $\tau^{\text{sim}}_{\text{unbind}}$ for these three inhibitors to the known, experimentally measured residence time of each compound on CDK2 (denoted here as, $\tau^{\text{exp}}_{\text{unbind}}$). The $\tau^{\text{exp}}_{\text{unbind}}$ for inhibitors **1**, **2**, and **3** are 6.0, 763.2, and 7476.6 s, respectively (Table 3), which are orders of magnitude longer than our estimates. The discrepancy between the absolute values of $\tau^{\text{sim}}_{\text{unbind}}$ and $\tau^{\text{exp}}_{\text{unbind}}$ could be due to the fact that the $\tau^{\text{exp}}_{\text{unbind}}$ of these inhibitors on CDK2 were measured in the presence of Cyclin-A, a member of the cyclin family that activates the enzymatic activity of its CDK partner, which was absent from our simulation. Assuming that the presence of Cyclin-A in our simulation systems would effectively add "friction" to the internal motions of CDK2 and slowdown its dynamics, our ssMD estimates of the unbinding times ($\tau^{\text{sim}}_{\text{unbind}}$), derived in the absence of Cyclin-A, can viewed as lower-bound estimates of the residence times of these inhibitors on CDK2. We note, however, that though $\tau^{\text{sim}}_{\text{unbind}}$ were significantly shorter than the corresponding $\tau^{\text{exp}}_{\text{unbind}}$, the trend in $\tau^{\text{sim}}_{\text{unbind}}$ agreed with the trend in $\tau^{\text{exp}}_{\text{unbind}}$ (Table 3).

Estimating Binding Free-Energy from Selectively Scaled MD (ssMD) Simulations

Finally, we explored whether we could also generate unbiased estimates of the binding free energy (ΔG^{sim}) of inhibitors **1**, **2**, and **3**, from our ssMD simulations. In principle, a free energy extrapolation (FEE) equation (see Eq. 9) could also be used to generate unbiased estimates of ΔG^{sim} on the unscaled potential.

To test the validity of this FEE approach, we first applied it to a bistable 1D potential (Eq. 7). For this test, we attempted to recover the known configurational distribution, $p(x)$, from the distributions associated with a series of potentials in which a term in the potential was selectively scaled by α (i.e., from $p(x,\alpha)$). Shown in Figure 5A is the comparison between the exact $\log(p(x))$ and the $\log(p^{\text{extra}}(x))$ that was recovered via extrapolation to $\alpha = 1$. The results indicate that by fitting $\log(p(x,\alpha))$ (Eq. 5), at α 0.50, 0.40, 0.20, 0.10, and 0.00, to the FEE equation (Eq. 8), we could recover the reference $\log(p(x))$ corresponding to the unscaled potential (Figure 5A).

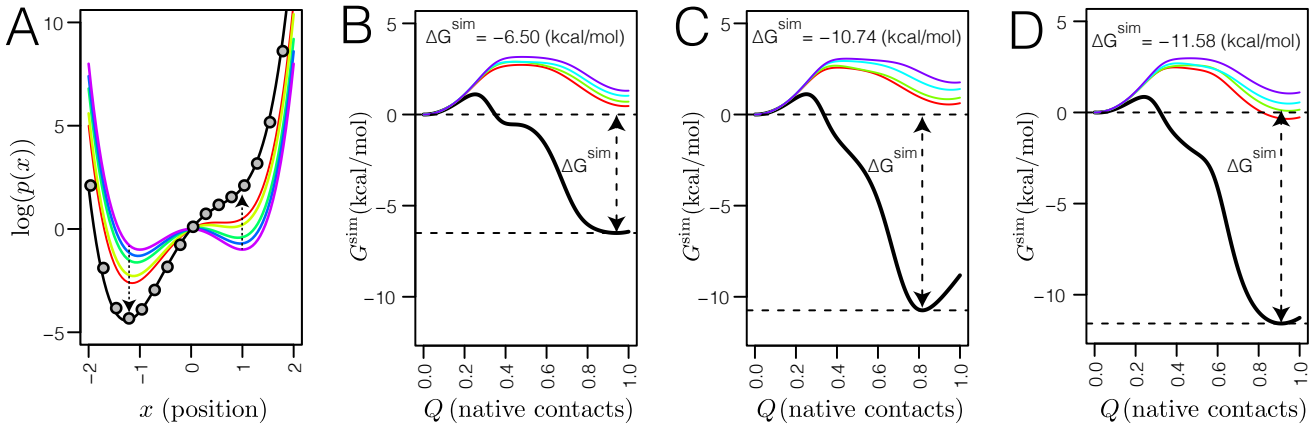

Figure 5. (A) Extrapolation results obtained for the 1-dimensional toy model (see Eq. 7). Shown is the comparison between the exact α=1 $\log(p(x))$ (*black line*) and the α=1 $\log(p^{\text{extra}}(x))$ (*grey circles*) that was obtained by extrapolating the $\log(p(x))$ profiles associated with α equal to 0.50 (*red*), 0.40 (*yellow*), 0.20 (*green*), 0.10 (*blue*), and 0.00 (*magenta*), respectively. (B-D) Free energy extrapolation (FEE) results obtained for inhibitors **1** (B), **2** (C), and **3** (D), respectively. For each, free energy profiles are shown along Q (fraction of native contacts) for $\alpha = 1/\alpha'$ equal to $1/4.50$ (*magenta*), $1/4.00$ (*cyan*), $1/3.50$ (*green*), $1/3.25$ (*red*), and $1/1.00$, respectively. Included in these plots are estimates of the binding free energies of the inhibitors (ΔG^{sim}), which were determined based on these profiles.

We next used this FEE method to generate unbiased estimates of ΔG^{sim} (the binding free energies) for the three CDK2 inhibitors. Shown in Figure 5B-D, are the free energy profiles along the fraction of native contacts, Q, for inhibitors **1**, **2**, and **3**, respectively. In each case, we were able to extrapolate the free energy associated with the unscaled potential from profiles computed using our ssMD simulation data; here again, extrapolation was carried out using data derived from the ssMD simulations carried out with α' set to 3.00, 3.25, 3.50, and 4.00, respectively and with $\alpha \equiv 1/\alpha'$. From the extrapolated

Inhibitor	ΔG^{sim} (kcal/mol)	Fitting Error (kcal/mol)	ΔG^{exp} (kcal/mol)	ΔG^{sim} - ΔG^{exp} (kcal/mol)
1	-6.50	± 0.03	-8.37	1.87
2	-10.74	± 0.04	-10.93	0.19
3	-11.58	± 0.07	-11.81	0.23

Table 4. Comparison between the free energy of binding estimated from our ssMD simulations (ΔG^{sim}) and those measured experimentally (ΔG^{exp}) for inhibitors **1**, **2**, and **3**, respectively. The fitting errors in ΔG^{sim} correspond to the residual standard errors.

free energy profiles we obtained estimates of -6.50, -10.74, and -11.58 kcal/mol for the ΔG^{sim} of inhibitors **1**, **2**, and **3**, respectively (Figure 5B-D and Table 4). Like our estimates of $\tau^{\text{sim}}_{\text{unbind}}$, our estimates of ΔG^{sim} were robust. The estimated errors in the ΔG^{sim}, quantified as the residual standard error in the linear fits to our FEE equation, were ± 0.03, ± 0.04, and ± 0.07 kcal/mol for the inhibitors **1**, **2**, and **3**, respectively (Table 4). Comparison between ΔG^{sim} and ΔG^{exp} (the experimental binding free energies) revealed that our ssMD-derived estimates of the binding free energies were within $\sim$2.0 kcal/mol of ΔG^{exp}; ΔG^{exp} were -8.37, -10.93, and -11.81 kcal/mol (Table 4) (31), respectively. Therefore, for these three CDK2 inhibitors, we were also able to make realistic estimates of their binding free energies from our ssMD simulations.

Discussion

Above, we described the implementation and testing of a selectively scaled MD (ssMD) approach for accelerating dissociative processes such as ligand unbinding. Using our ssMD approach, we were able to significantly accelerate ligand unbinding. For instance, over an aggregate simulation time of 5.0×10^3 ns, conventional simulations on the unscaled potential of CDK2 bound to inhibitor **1** yielded only 5 unbinding trajectories. In contrast, by carrying out ssMD of the same complex with α' set to 3.00, we were able to generate a total of 49 unbinding trajectories over an aggregate simulation time of 2.5×10^3 ns. Our ssMD approach, therefore, proved to be an effective method for accelerating ligand unbinding during simulations.

An attractive feature of our ssMD approach is that it does not incur a significant computational cost. To run ssMD simulations, we merely modified, using the NBFIX facility in the CHARMM, the parameter files to reflect the scaling of the well-depth of the atom types in the ligands and atom types in water; the only additional cost from the initial loading of additional NBFIXes in the modified parameter files. By applying the scaling at the level of the parameter file, simulations could be carried out without any special scripting or source code changes. To run the series of ssMD simulations described in the study, we merely generated a set of modified parameter files, one for each the four scaling values used. An added benefit of applying the scaling at the level of the parameter file is that though we used CHARMM to carry out our simulations, the ssMD simulations could be similarly carried out using other MD simulation packages such as NAMD, GROMACS, or OpenMM.

By coupling ssMD simulations with the KRE and FEE methods, we are also able to overcome an outstanding problem in using uniformly scaled MD simulation to study dissociative processes, namely, choosing the right scaling value (α) at which to estimate kinetic and thermodynamic properties. Rather than deriving estimates of the *relative* values of kinetic and thermodynamic parameters of dissociative processes from simulations carried out using a single scaling factor, whose chosen value is difficult if not impossible to justify, our extrapolation methods inherently make use of simulations carried out at a series of scaling factors to estimate their *absolute* values on the original unscaled potential.

Another advantage of combining our ssMD simulations with the KRE and FEE methods is that collectively, they could be used to fully characterize the kinetics and thermodynamics of specific association/dissociation processes. For instance, in the context of ligand association/dissociation, estimates of not just the binding free energy (ΔG) but also of the barrier to binding and the barrier unbinding can be determined from the extrapolated free energy profiles. Moreover, when combined with KRE-derived estimates of the off-rate ($k_{\text{off}} \equiv 1/\tau_{\text{unbind}}$), estimates of the on-rates (k_{on}) can be obtained by making use of the fact that the dissociation constant (K_d) can be expressed as

$$K_d = e^{(\Delta G/RT)} \tag{10}$$

and that

$$K_d = \frac{k_{\text{on}}}{k_{\text{off}}}. \tag{11}$$

Previous use of scaled MD to study ligand unbinding utilized a uniformly scaled potential to accelerate unbinding during simulations (22). Accordingly, in addition to the inter-energy terms between the ligand and protein, the intra-energy terms in the ligand and protein were also scaled, which has the negative side-effect of unfolding the protein components within the complex. As a result, conformational restraints must be employed to maintain the correct fold of the protein components of a complex during uniformly scaled MD simulations. Because we only selectively scaled up the ligand-water interactions, effectively scaling down the ligand-protein interactions, our ssMD approach does not require the use of conformational restraints, which simplifies system setup as well as allows the generation of unbinding trajectories with an entirely flexible protein.

Beyond the uniformly scaled MD approach, our ssMD simulation approach also resembles other recently described enhanced sampling methods, which too are being used to study ligand unbinding. For instance, in τ-random accelerated molecular dynamics (referred to as τ-RAMD), a randomly oriented force is applied to a ligand to accelerate unbinding (45). Independent τ-RAMD simulations are then used to compute apparent unbinding times. In an impressive study, Wade and co-workers applied the τ-RAMD to estimate the unbinding times for 70 Hsp90 inhibitors; after removing outliers, their method was able to achieve a correlation coefficient (R^2) $\sim$0.86 between relative residence time from simulations and experiments (46). In a related approach, Wong, using the Focal Adhesion Kinase as an example, demonstrated that steered MD simulations could be used to identify compounds having prolonged residences times (47). Collectively, this family of enhanced sampling methods seemed posed to facilitate the expanded use of MD simulations within kinetics-based drug discovery.

A significant limitation of our method, however, is that though we were able to accelerate ligand unbinding by several orders of magnitude, the individual simulations that required to estimate residence times and free-energy profiles are still computationally expensive simulations to be carried out. Because of this, our ssMD simulations will be best suited to predict the kinetics and thermodynamics of ligands during latter stages of drug-development, where the goal is to rank a few hundred compounds, as opposed to screening an entire compound library typically containing > 1000 compounds. It may be possible, however, to couple ssMD and machine learning to develop predictive models that can learn from ssMD simulations and recapitulate ssMD-derived residence times and binding affinities. Such a predictive model, for example, could be trained to predict ssMD-derived residence times and binding affinities directly from the structures of protein-ligand complexes and incorporated within early-stage virtual screening protocols as an enhanced scoring function to help focus screening on ligands that will exhibit desired kinetic and thermodynamic properties on a given target.

More fundamentally, the accuracy of the underlying force field that is employed limits the utility of all methods that utilize classical molecular dynamics simulations. In the context of protein-ligand unbinding, ligand polarization effects are expected to play a significant role in determining the strength of individual protein-ligand interactions and in so doing, the residence time of the ligand. Indeed, in recent work, in which QM/MM and metadynamics was used to study the binding kinetics of the anticancer drug Imatinib to Src kinase, significant changes in the ligand polarization were observed over the unbinding trajectories which, undoubtedly, will impact the kinetics (and thermodynamics) of ligand unbinding (48).

These limitations notwithstanding, because of the simplicity and computational efficiency of our ssMD approach, it should find utility as a general purpose technique for accelerating events in biomolecular simulations. When combined with the KRE and the FEE approaches described above, it should also be useful in estimating the kinetics and thermodynamics of these processes.

Conclusion

In the preceding, we introduced a selectively scaled MD (or ssMD) simulation method that can be used to rapidly generate trajectories of complex dissociative processes. Also, when coupled with a previously described Kramers'-based rate extrapolation (KRE) and a newly described free energy extrapolation (FEE) scheme, our ssMD simulation data can, in principle, be used to estimate the kinetics and thermodynamics of dissociative processes. Here, we applied our ssMD approach, and both the KRE and the FEE methods, to study the unbinding of three inhibitors of cyclin-dependent kinase 2 (CDK2). For the three inhibitors, we were able to generate unbiased estimates of their unbinding times that agreed well their *relative* experimentally measured residence times. Also, our estimates of the binding free energies of the three CDK2 inhibitors we studied were within 2 kcal/mol of their experimental values. On the basis of these results, we envision that our ssMD approach coupled with the KRE and FEE analysis methods will find utility as a simple and an efficient approaches for modeling and analyzing biologically relevant dissociative events.

References

1. Kamar, R. I., E. J. Banigan, A. Erbas, R. D. Giuntoli, M. O. De La Cruz, R. C. Johnson, and J. F. Marko. 2017. Facilitated dissociation of transcription factors from single DNA binding sites. *Proc. Natl. Acad. Sci.* 114:E3251–E3257.
2. Ullman, K. S., M. A. Powers, and D. J. Forbes. 1997. Nuclear export receptors: from importin to exportin. *Cell*. 90:967–970.

3. McCammon, J. A., B. R. Gelin, and M. Karplus. 1977. Dynamics of folded proteins. *Nature*. 267:585–590.

4. Karplus, M., and J. A. McCammon. 2002. Molecular dynamics simulations of biomolecules. *Nat. Struct. Mol. Biol.* 9:646–652.

5. Dror, R. O., R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw. 2012. Biomolecular simulation: a computational microscope for molecular biology. *Annu. rev. biophys.* 41:429–452.

6. Shaw, D.E., M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao, and M. P. Eastwood. 2008. Anton, a special-purpose machine for molecular dynamics simulation. *commun. ACM*. 51:91–97.

7. Andoh, Y., N. Yoshii, K. Fujimoto, K. Mizutani, H. Kojima, A. Yamada, S. Okazaki, K. Kawaguchi, H. Nagao, K. Iwahashi, and F. Mizutani. 2013. MODYLAS: A highly parallelized general-purpose molecular dynamics simulation program for large-scale systems with long-range forces calculated by fast multipole method (FMM) and highly scalable fine-grained new parallel processing algorithms. *J. Chem. Theory Comp.* 9:3201–3209.

8. Shaw, D. E., J. P. Grossman, J. A. Bank, B. Batson, J. A. Butts, J. C. Chao, M. M. Deneroff, R. O. Dror, A. Even, C. H. Fenton, and A. Forte. 2014. 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. IEEE Press, New Orleans, pp. 41–53.

9. Jung, J., T. Mori, C. Kobayashi, Y. Matsunaga, T. Yoda, M. Feig, and Y. Sugita. 2015. 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.

10. Shan, Y., E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, and D. E. Shaw. 2011. How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* 133:9181–9183.

11. Yu, I., T. Mori, T. Ando, R. Harada, J. Jung, Y. Sugita, and M. Feig. 2016. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *Elife*. 5:e19274.

12. Buch, I., T. Giorgino, and G. De Fabritiis. 2011. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci.* 108:10184–10189.

13. Limongelli, V., M. Bonomi, and M. Parrinello. 2013. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. USA*. 110:6358–6363.

14. Tiwary, P., V. Limongelli, M. Salvalaglio, and M. Parrinello. 2015. Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. USA*. 112:E386–E391.

15. Teo, I., C. G. Mayne, K. Schulten, and T. Lelievre. 2016. Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time. *J. Chem. Theory Comput.* 12:2983–2989.

16. Dickson, A., and S. D. Lotz. 2017. Multiple ligand unbinding pathways and ligand-induced destabilization revealed by WExplore. *Biophys. J.* 112:620–629.

17. Ribeiro, J. M. L., P. Bravo, Y. Wang, and P. Tiwary. 2018. Reweighted autoencoded variational bayes for enhanced sampling (RAVE). *J. Chem. Phys.* 149:072301.

18. Ribeiro, J. M. L., and P. Tiwary. 2018. Achieving reversible ligand-protein unbinding with deep learning and molecular dynamics through RAVE. *bioRxiv*. 400002.

19. Mark, A. E., W. F. van Gunsteren, and H. J. Berendsen. 1991. Calculation of relative free energy via indirect pathways. *J. Chem. Phys.* 94:3808–3816.

20. Tsujishita, H., I. Moriguchi, and S. Hirono. 1993. Potential-scaled molecular dynamics and potential annealing: effective conformational search techniques for biomolecules. *J. Phys. Chem.* 97:4416–4420.

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

22. Mollica, L., S. Decherchi, S. R. Zia, R. Gaspari, A. Cavalli, and W. Rocchia. 2015. Kinetics of protein-ligand unbinding via smoothed potential molecular dynamics simulations. *Sci. Rep.* 5:11539.

23. Mollica, L., I. Theret, M. Antoine, F. Perron-Sierra, Y. Charton, J. M. Fourquez, M. Wierzbicki, J. A. Boutin, G. Ferry, S. Decherchi, and G. Bottegoni. 2016. Molecular dynamics simulations and kinetic measurements to estimate and predict protein–ligand residence times. *J. Med. Chem.* 59:7167–7176.

24. Bernetti, M., E. Rosini, L. Mollica, M. Masetti, L. Pollegioni, M. Recanatini, and A. Cavalli. 2018. Binding residence time through scaled molecular dynamics: A prospective application to hdaao inhibitors. *J. Chem. Inf. Model.* 58:2255–2265.

25. Schuetz, D. A., M. Bernetti, M. Bertazzo, D. Musil H. M. Eggenweiler, M. Recanatini, M. Masetti, G. F. Ecker, and A. Cavalli. 2019. Predicting residence time and drug unbinding pathway through scaled molecular dynamics. *J. Chem. Inf. Model.* 59:535–549.

26. Wereszczynski, J., and J. A. McCammon. 2010. Using selectively applied accelerated molecular dynamics to enhance free energy calculations. *J. Chem. Theory Comp.* 6:3285–3292.

27. Doshi, U., and D. Hamelberg. 2011. Extracting realistic kinetics of rare activated processes from accelerated molecular dynamics using kramers' theory. *J. Chem. Theory Comp.* 7:575–581.

28. Frank, A. T., and I. Andricioaei. 2016. Reaction coordinate-free approach to recovering kinetics from potential-scaled simulations: application of kramers' rate theory. *J. Phys. Chem. B.* 120:8600–8605.

29. Xin, Y., U. Doshi, D. Hamelberg, 2010. Examining the limits of time reweighting and Kramers' rate theory to obtain correct kinetics from accelerated molecular dynamics. *J. Chem. Phys.* 132:224101.

30. Brooks, B., R. Bruccoleri, B. Olafson, and M. Karplus, 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

31. Ayaz, P., D. Andres, D. A. Kwiatkowski, C. C. Kolbe, P. Lienau, G. Siemeister, U. Lücking, and C. M. Stegmann. 2016. Conformational adaption may explain the slow dissociation kinetics of roniciclib (BAY 1000394), a type I CDK inhibitor with kinetic selectivity for CDK2 and CDK9. *ACS Chem. Biol.* 11:1710–1719.

32. Šali, A., and T. L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.

33. Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. 2004. UCSF Chimera — a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25:1605–1612.

34. Jo, S., T. Kim, V. G. Iyer, and W. Im. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29:1859–1865.

35. Lee, J., X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. Qi, and S. Jo. 2015. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* 12:405–413.

36. Best, R. B., X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell Jr.. 2012. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *J. Chem. Theory Comput.* 8:3257–3273.

37. Vanommeslaeghe, K., E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell Jr.. 2010. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* 31:671–690.

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

39. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: An $N\log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

40. Ryckaert, J. P., G. Ciccotti, and H. J. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23:327–341.

41. Kramers, H. 1940. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*. 7:284–304.

42. Nummela, J I, Andricioaei, I 2007 Exact low-force kinetics from high-force single-molecule unfolding events. *Biophys. J.* 93:3373–3381.

43. Bates, D. M., and D. G. Watts. 1988. Nonlinear regression: iterative estimation and linear approximations. In Nonlinear regression analysis and its applications. Wiley and Sons, New York, pp. 33–66.

44. Law, S. M., L. S. Ahlstrom, A. Panahi, and C. L. Brooks III. 2014. Hamiltonian mapping revisited: Calibrating minimalist models to capture molecular recognition by intrinsically disordered proteins. *J. Phys. Chem. Lett.* 5:3441–3444.

45. Lüdemann, S. K., V. Lounnas, and R. C. Wade. 2000. How do substrates enter and products exit the buried active site of cytochrome p450cam? 1. random expulsion molecular dynamics investigation of ligand access channels and mechanisms1. *J. Mol. Biol.* 303:797–811.

46. Kokh, D. B., M. Amaral, J. Bomke, U. Grädler, D. Musil, H. P. Buchstaller, M. K. Dreyer, M. Frech, M. Lowinski, F. Vallée, M. Bianciotto, M. Rak and R. C. Wade. 2018. Estimation of drug-target residence times by τ-random acceleration molecular dynamics simulations. *J. Chem. Theory Comput.* 14:3859–3869.

47. Wong, C. F. 2018. Steered molecular dynamics simulations for uncovering the molecular mechanisms of drug dissociation and for drug screening: A test on the focal adhesion kinase. *J. Comput. Chem.* 39:1307–1318.

48. Haldar, S., F. Comitani, G. Saladino, C. Woods, M. W. van der Kamp, A. J. Mulholland, and F. L. Gervasio. 2018. A multiscale simulation approach to modeling drug–protein binding kinetics. *J. Chem. Theory Comput.* 14:6093–6101.

Author contributions statement

A.T.F. designed the research, I.D. generated the data, and I.D. and A.T.F. analyzed the results. All authors reviewed the manuscript.

Additional information

Competing financial interests No conflict of interest.