

**Ligand Gaussian accelerated Molecular Dynamics 3 (LiGaMD3): Improved Calculations of
Binding Thermodynamics and Kinetics of Both Small Molecules and Flexible Peptides**

Jinan Wang and Yinglong Miao*

Computational Medicine Program and Department of Pharmacology, University
of North Carolina – Chapel Hill, Chapel Hill, North Carolina, USA 27599

*To whom correspondence should be addressed: Yinglong_Miao@med.unc.edu

Abstract

Binding thermodynamics and kinetics play critical roles in drug design. However, it has proven challenging to efficiently predict ligand binding thermodynamics and kinetics of small molecules and flexible peptides using conventional Molecular Dynamics (cMD), due to limited simulation timescales. Based on our previously developed Ligand Gaussian accelerated Molecular Dynamics (LiGaMD) method, we present a new approach, termed “LiGaMD3”, in which we introduce triple boosts into three individual energy terms that play important roles in small-molecule/peptide dissociation, rebinding and system conformational changes to improve the sampling efficiency of small-molecule/peptide interactions with target proteins. To validate the performance of LiGaMD3, MDM2 bound by a small molecule (Nutlin 3) and two highly flexible peptides (PMI and P53) were chosen as model systems. LiGaMD3 could efficiently capture repetitive small-molecule/peptide dissociation and binding events within 2 microsecond simulations. The predicted binding kinetic constant rates and free energies from LiGaMD3 agreed with available experimental values and previous simulation results. Therefore, LiGaMD3 provides a more general and efficient approach to capture dissociation and binding of both small-molecule ligand and flexible peptides, allowing for accurate prediction of their binding thermodynamics and kinetics.

Keywords: Ligand binding kinetics, Ligand binding free energy, Peptide binding, Enhanced sampling, Ligand Gaussian accelerated Molecular Dynamics.

Introduction

Both small molecules and peptides are important sources of novel drugs targeting many important biological processes¹. It is critical to understand the binding mechanisms of small molecules and peptides to their target proteins, which not only deepens our understanding of fundamental biological processes but also facilitates the development of more potent and selective drugs for treating human diseases^{2, 3}. Several experimental techniques⁴⁻⁶ have been developed to explore the binding interactions between the protein and small molecule or peptide. For example, structural biology techniques including X-ray crystallography and cryo-electron microscopy (cryo-EM) have been widely used to determine the complex structures of protein-small molecule and protein-peptide complexes⁵. Recently, significant advancements in Deep Learning methodologies such as AlphaFold⁷ and RoseTTAFold All-Atom (RFAA)⁸ have led to accurate prediction of protein-small molecule or protein-peptide complex structures. However, such techniques provide only static snapshots of protein-small molecule or protein-peptide interactions. It is still challenging to capture small-molecule/peptide binding and dissociation processes and determine potential intermediate states of small-molecule/peptide binding to their target proteins, which are also important for drug design⁹.

Recently, drug binding kinetics has been recognized to be valuable for drug design¹⁰⁻¹⁵. The drug dissociation rate (k_{off}) appears to correlate with drug efficacy better than the binding free energy¹⁰⁻¹⁵. However, drug binding kinetic rates have proven more challenging to predict, due to the slow processes of drug dissociation and binding^{11, 16}. With remarkable advancements in computer hardware and methodological developments, conventional Molecular Dynamics (cMD) simulations are now able to capture spontaneous small-molecule/peptide binding to their target proteins and predict corresponding association rates (k_{on})¹⁷⁻²⁰. However, it remains a difficult

challenge in applying cMD to capture repetitive small-molecule/peptide dissociation and rebinding processes within accessible timescales, thereby hindering the accurate prediction of small-molecule/peptide binding kinetic rates¹⁷. Shan et al.¹⁷ successfully captured spontaneous binding of the Dasatinib drug to its target Src kinase and accurately predicted the ligand association rate based on tens-of-microsecond cMD simulations. Tens-of-microsecond cMD simulations²⁰ have successfully captured repetitive binding and dissociation events of six small-molecule fragments with very weak millimolar binding affinities to the protein FKBP, allowing accurate prediction of fragment binding free energies. However, no dissociation events have been observed for typical ligand molecules in the cMD simulations. Furthermore, capturing peptide dissociation and rebinding processes poses an even more challenging task for cMD, given that peptides are known to induce significant conformational changes upon binding²¹ and the timescales for the peptide dissociation are even longer²². For example, cMD simulations with elevated temperature conducted for 200 μ s using the Anton specialized supercomputer have captured 70 binding and unbinding events between an intrinsically disordered protein fragment of the measles virus nucleoprotein and the X domain of the measles virus phosphoprotein complex, shedding light on the detailed understanding of the peptide's "folding-upon-binding" mechanism²³. Despite these advancements, it is still rather challenging for cMD to effectively simulate binding and dissociation of typical small molecule or peptide to their target proteins.

Enhanced sampling methods^{24, 25} have been developed to extend the accessible timescales of MD simulations. These methods include Metadynamics²⁶⁻³¹, Steered MD³²⁻³⁴, Umbrella Sampling^{32, 35-37}, Replica Exchange MD³⁸⁻⁴¹, Random Acceleration Molecular Dynamics (RAMD)⁴²⁻⁴⁴, Scaled MD^{45, 46}, accelerated MD (aMD)⁴⁷, Gaussian accelerated MD (GaMD)^{48, 49}, Markov State Model (MSM)⁵⁰⁻⁵², Weighted Ensemble^{53, 54}, and so on. Metadynamics^{21, 55} simulations

utilizing carefully chosen collective variables (CVs) have successfully predicted the peptide binding and dissociation rates in the systems. Particular, 27 μ s Metadynamics simulations of the peptide P53 binding to the MDM2⁵⁵ predicted values of (k_{on} , k_{off}) at ($0.43 \pm 0.22 \times 10^7 \text{M}^{-1}\text{s}^{-1}$, $0.7 \pm 0.4 \text{s}^{-1}$), showing good agreement with the corresponding experimental values of ($0.92 \times 10^7 \text{M}^{-1}\text{s}^{-1}$, 2.06s^{-1}). Similarly, Weighted Ensemble⁵⁶ of a total amount of $\sim 120 \mu$ s cMD simulations with implicit solvent model for the P53-MDM2 system predicted a highly consistent binding kinetic rate (k_{on}) of 7s^{-1} . MSM⁵⁷ analysis based on a total of 831 μ s cMD simulations for peptide P53 binding to MDM2 accurately predicted values of k_{on} and k_{off} at $0.019 \times 10^7 \text{M}^{-1}\text{s}^{-1}$ and 2.5s^{-1} , respectively. Another MSM built on hundreds-of-microsecond cMD and Hamiltonian replica exchange simulations has been implemented to characterize binding and dissociation of the PMI peptide to the MDM2²². The predicted values of (k_{on} , k_{off}) were ($300 \times 10^7 \text{M}^{-1}\text{s}^{-1}$, $0.125\text{-}1.13 \text{s}^{-1}$), being comparable to the corresponding experimental values of ($52.7 \times 10^7 \text{M}^{-1}\text{s}^{-1}$, 0.037s^{-1}). Nevertheless, MSM and Weighted Ensemble require expensive and exceedingly long simulations. GaMD was developed to provide both unconstrained enhanced sampling and free energy calculations of large biomolecules^{48, 49}. It works by applying a harmonic boost potential to reduce system energy barriers. The boost potential exhibits a near Gaussian distribution, which enables accurate reweighting of the free energy profiles through cumulant expansion to the second order^{48, 49}. Recently, novel Ligand GaMD (LiGaMD)⁵⁸ and LiGaMD2⁵⁹ approaches have been developed to more efficiently sample small-molecule dissociation and rebinding processes, offering accurate prediction of ligand binding thermodynamics and kinetics. In LiGaMD, a selective boost is specifically applied at the ligand's non-bonded interaction potential energy⁵⁸. In LiGaMD2, the selective boost extended to the essential potential energy of both the ligand and surrounding residues in the protein pocket, which significantly improved the sampling of ligand dissociation

and rebinding in a closed binding pocket⁵⁹. An increasing amount of studies^{41, 52, 60-63} demonstrated the pivotal role of non-bonded interaction potentials in ligand binding, along with the crucial structural flexibility of proteins and peptides^{52, 64-69}. Building upon the successes of LiGaMD and LiGaMD2, which primarily focused on small-molecules, here we introduce a more general approach, termed LiGaMD3, for binding simulations of both small-molecules and flexible peptides. In LiGaMD3, three distinct boosts are applied: one on the non-bonded interaction energy of the substrate, the second one on the remaining non-bonded potential energy of the system, and a third one on the system bonded potential energy. These boosts are designed to accelerate the substrate dissociation, facilitate substrate rebinding, and promote the system conformational changes, respectively. MDM2⁷⁰, a well-known oncology protein involved in regulating diverse cellular signaling pathways, serves as an ideal model system for investigating the binding and dissociation of both small molecules and highly flexible peptides. Notably, this system has been extensively explored through experimental studies and simulations as mentioned above, highlighting its critical role in drug discovery. Therefore, MDM2 bound by small-molecule drugs and peptides were chosen as model systems in this study. Through two microsecond LiGaMD3 simulations, we successfully captured repetitive ligand and peptide binding and dissociation processes across all MDM2 systems. LiGaMD3 facilitated highly efficient and accurate predictions of ligand and peptide binding thermodynamics and kinetics, being consistent with experimental binding free energies and kinetic rates.

Methods

LiGaMD3: Triple boost for ligand dissociation and rebinding

We consider a system of small-molecule/peptide L binding to a protein P in a biological environment E . The system comprises of N atoms with their coordinates $r \equiv \{\vec{r}_1, \dots, \vec{r}_N\}$ and momenta $p \equiv \{\vec{p}_1, \dots, \vec{p}_N\}$. The system Hamiltonian can be expressed as:

$$H(r, p) = K(p) + V(r), \quad (1)$$

where $K(p)$ and $V(r)$ are the system kinetic and total potential energies, respectively. We decompose the potential energy into the following terms:

$$V(r) = V_{P,b}(r_P) + V_{L,b}(r_L) + V_{E,b}(r_E) + V_{PP,nb}(r_P) + V_{PL,nb}(r_{PL}) + \\ V_{PE,nb}(r_{PE}) + V_{LL,nb}(r_L) + V_{LE,nb}(r_{LE}) + V_{EE,nb}(r_E) \quad (2)$$

where $V_{P,b}$, $V_{L,b}$ and $V_{E,b}$ are the bonded potential energies of protein P, small-molecule/peptide L and environment E , respectively. $V_{P,nb}$, $V_{LL,nb}$ and $V_{EE,nb}$ are the self non-bonded potential energies in the protein P, small-molecule/peptide L and environment E , respectively. $V_{PL,nb}$, $V_{PE,nb}$ and $V_{LE,nb}$ are the non-bonded interaction energies between P-L, P-E and L-E, respectively. According to classical molecular mechanics force fields^{71, 72}, the non-bonded potential energies are usually calculated as:

$$V_{nb} = V_{elec} + V_{vdW}, \quad (3)$$

where V_{elec} and V_{vdW} are the system electrostatic and van der Waals potential energies. The bonded potential energies are usually calculated as

$$V_b = V_{bond} + V_{angle} + V_{dihedral} \quad (4)$$

where V_{bond} , V_{angle} and $V_{dihedral}$ are the system bond, angle and dihedral potential energies. In LiGaMD3, the essential non-bonded interaction potential energy of the ligand is defined as:

$$V_L(r) = V_{LL,nb}(r_L) + V_{PL,nb}(r_{PL}). \quad (5)$$

We add a boost potential selectively to the $V_L(r)$ according to the GaMD algorithm:

$$\Delta V_L(r) = \begin{cases} \frac{1}{2} k_L (E_L - V_L(r))^2, & V_L(r) < E_L \\ 0, & V_L(r) \geq E_L, \end{cases} \quad (6)$$

where E_L is the threshold energy for applying boost potential and k_L is the harmonic constant. The LiGaMD3 simulation parameters are derived similarly as in the previous GaMD⁴⁸, LiGaMD⁵⁸, and LiGaMD2⁵⁹. The subscript L is omitted for simplicity hereafter. When E is set to the lower bound as the system maximum potential energy ($E=V_{max}$), the effective harmonic force constant k_0 can be calculated as:

$$k_0 = \min(1.0, k'_0) = \min\left(1.0, \frac{\sigma_0}{\sigma_V} \frac{V_{max}-V_{min}}{V_{max}-V_{avg}}\right), \quad (7)$$

where V_{max} , V_{min} , V_{avg} and σ_V are the maximum, minimum, average and standard deviation of the boosted system potential energy, and σ_0 is the user-specified upper limit of the standard deviation of ΔV (e.g., $10k_B T$) for proper reweighting. The harmonic constant k is calculated as $k = k_0 \cdot \frac{1}{V_{max}-V_{min}}$ with $0 < k_0 \leq 1$. Alternatively, when the threshold energy E is set to its upper bound

$E = V_{min} + \frac{1}{k}$, k_0 is set to:

$$k_0 = k_0'' \equiv \left(1 - \frac{\sigma_0}{\sigma_V}\right) \frac{V_{max}-V_{min}}{V_{avg}-V_{min}}, \quad (8)$$

if k_0'' is found to be between 0 and 1. Otherwise, k_0 is calculated using Eqn. (7).

In addition to selectively boosting the essential non-bonded interaction potential energy of the ligand, another boost potential is applied on the remaining non-bonded potential energy of the system ($V_D(r)$) to facilitate ligand rebinding:

$$V_D(r) = V_{nb} - V_L(r) = V_{PP,nb}(r_P) + V_{PE,nb}(r_{PE}) + V_{LE,nb}(r_{LE}) + V_{EE,nb}, \quad (9)$$

$$\Delta V_D(r) = \begin{cases} \frac{1}{2} k_D (E_D - V_D(r))^2, & V_D(r) < E_D \\ 0, & V_D(r) \geq E_D \end{cases} \quad (10)$$

where V_D is the total system potential energy other than the essential non-bonded ligand potential energy ligand, E_D is the corresponding threshold energy for applying the second boost potential and k_D is the harmonic constant. The third boost potential is applied to the total bonded potential energy of the system as:

$$V_B(r) = V_{P,b}(r_P) + V_{L,b}(r_L) + V_{E,b}(r_E) \quad (11)$$

$$\Delta V_B(r) = \begin{cases} \frac{1}{2} k_B (E_B - V_B(r))^2, & V_B(r) < E_B \\ 0, & V_B(r) \geq E_B \end{cases} \quad (12)$$

This leads to LiGaMD3 with a triple-boost potential $\Delta V(r) = \Delta V_L(r) + \Delta V_D(r) + \Delta V_B(r)$.

Energetic Reweighting of LiGaMD3

To calculate potential of mean force (PMF)⁷³ from LiGaMD3 simulations, the probability distribution along a reaction coordinate is written as $p^*(A)$. Given the boost potential $\Delta V(r)$ of each frame, $p^*(A)$ can be reweighted to recover the canonical ensemble distribution, $p(A)$, as:

$$p(A_j) = p^*(A_j) \frac{\langle e^{\beta \Delta V(\vec{r})} \rangle_j}{\sum_{i=1}^M \langle p^*(A_i) e^{\beta \Delta V(\vec{r})} \rangle_i}, \quad j = 1, \dots, M, \quad (13)$$

where M is the number of bins, $\beta = k_B T$ and $\langle e^{\beta \Delta V(\vec{r})} \rangle_j$ is the ensemble-averaged Boltzmann factor of $\Delta V(\vec{r})$ for simulation frames found in the j^{th} bin. The ensemble-averaged reweighting factor can be approximated using cumulant expansion:

$$\langle e^{\beta \Delta V(\vec{r})} \rangle = \exp \left\{ \sum_{k=1}^{\infty} \frac{\beta^k}{k!} C_k \right\}, \quad (14)$$

where the first two cumulants are given by

$$\begin{aligned} C_1 &= \langle \Delta V \rangle, \\ C_2 &= \langle \Delta V^2 \rangle - \langle \Delta V \rangle^2 = \sigma_v^2. \end{aligned} \quad (15)$$

The boost potential obtained from LiGaMD3 simulations usually follows near-Gaussian distribution. Cumulant expansion to the second order thus provides a good approximation for computing the reweighting factor^{24, 74}. The reweighted free energy $F(A) = -k_B T \ln p(A)$ is calculated as:

$$F(A) = F^*(A) - \sum_{k=1}^2 \frac{\beta^k}{k!} C_k + F_c, \quad (16)$$

where $F^*(A) = -k_B T \ln p^*(A)$ is the modified free energy obtained from LiGaMD2 simulation and F_c is a constant.

Ligand binding kinetics obtained from reweighting of LiGaMD3 Simulations

Reweighting of ligand binding kinetics from LiGaMD3 simulations followed a similar protocol using Kramers' rate theory that has been recently implemented in kinetics reweighting of the GaMD^{58, 75-77}. Provided sufficient sampling of repetitive ligand dissociation and binding in the simulations, we record the time periods and calculate their averages for the ligand found in the bound (τ_B) and unbound (τ_U) states from the simulation trajectories. The τ_B corresponds to residence time in drug design¹⁴. Then the ligand dissociation and binding rate constants (k_{off} and k_{on}) were calculated as:

$$k_{\text{off}} = \frac{1}{\tau_B}. \quad (17)$$

$$k_{\text{on}} = \frac{1}{\tau_U [L]}, \quad (18)$$

where $[L]$ is the ligand concentration in the simulation system.

For a particle climbing over potential energy barriers, Kramers showed that the reaction rate depends on temperature and viscosity of the host medium. The reaction rates were derived for both limiting cases of small and large viscosity. In the context of biomolecular simulations in aqueous medium, it is relevant for us to focus on the large viscosity limiting case. Biomolecules move in the high friction (“overdamping”) regime and energy barriers are much greater than $k_B T$ (k_B is the Boltzmann’s constant and T is temperature). In this case, the reaction rate is calculated as⁷⁷:

$$k_R \cong \frac{w_m w_b}{2\pi\xi} e^{-\Delta F/k_B T}, \quad (19)$$

where w_m and w_b are frequencies of the approximated harmonic oscillators (also referred to as curvatures of free energy surface^{78, 79}) near the energy minimum and barrier, respectively, ξ is the frictional rate constant and ΔF is the free energy barrier of transition.

Without the loss of generality, we consider a 1D potential of mean force (PMF) free energy profile of a reaction coordinate $F(A)$. Near minimum at A_m , the free energy can be approximated by a harmonic oscillator of frequency w_m , i.e., $F(A) = \frac{1}{2}(2\pi w_m)^2(A - A_m)^2$. Near barrier at A_b , the free energy is approximated as $F(A) = F_b - \frac{1}{2}(2\pi w_b)^2(A - A_b)^2$, where F_b is the free energy at A_b and w_b is the frequency of the approximated harmonic oscillator. Then we can calculate w_m and w_b as:

$$w = \sqrt{\frac{|F''(A)|}{2\pi}}, \quad (20)$$

where $F''(A)$ is the second-order derivative of the PMF profile.

The friction constant ξ or diffusion coefficient D with $\xi = k_B T / D$ can be estimated as follows. First, we calculate a survival function $S(t)$ as the probability that the system remains in an energy well longer than time t . In a direct approach, we count the events that the system visits the energy well throughout a simulation. We record and measure the time intervals of each visiting event until the system escapes over an energy barrier. Then we have a time series T_i , where $i=1, 2, \dots, N$, and N is the total number barrier transitions observed in the simulation. The time series is subsequently ordered such that $\hat{T}_1 \leq \hat{T}_2 \leq \dots \leq \hat{T}_N$. With that, the survival function is estimated as $S(\hat{T}_i) \approx 1 - i/N$, which is the probability that the system is trapped in the energy well for time longer than $\hat{T}_i$. Alternatively, we can numerically calculate the time-dependent probability density of reaction coordinate A , $\rho(A, t)$ by solving the Smoluchowski equation along 1D PMF profile of the reaction coordinate:

$$\frac{\partial \rho(A, t)}{\partial t} = D \frac{\partial}{\partial A} \left[e^{-F(A)/k_B T} \frac{\partial}{\partial A} (e^{F(A)/k_B T} \rho) \right] \quad (21)$$

Then the survival function is calculated as $S(t) = \int_t^\infty \int_{A_{b1}}^{A_{b2}} \rho(A, t) dA dt$, where A_{b1} and A_{b2} are two boundaries of the energy well. The initial condition is often set as the Boltzmann distribution of reaction coordinate A in the energy well, i.e., $\rho(A, 0) = e^{-F(A)/k_B T}$.

Second, using the above survival functions, we estimate the effective kinetic rates as the negative of the slopes in linear fitting of the $\ln[S(t)]$ versus t , i.e., $k = -d\ln[S(t)]/dt$. This is based on the assumption that the survival function exhibits exponential decay as observed in earlier studies. Finally, the apparent diffusion coefficient D is obtained by dividing the kinetic rate calculated directly using the transition time series collected from the simulation by that using the probability density solution of the Smoluchowski equation⁸⁰. In order to reweight ligand kinetics from the LiGaMD3 simulations using the Kramers' rate theory, the free energy barriers of ligand

binding and dissociation are calculated from the original (reweighted, ΔF) and modified (no reweighting, ΔF^*) PMF profiles, similarly for curvatures of the reweighted (w) and modified (w^* , no reweighting) PMF profiles near the ligand bound (“B”) and unbound (“U”) low-energy wells and the energy barrier (“Br”), and the ratio of apparent diffusion coefficients from simulations without reweighting (modified, D^*) and with reweighting (D). The resulting numbers are then plugged into Eq. (17) to estimate accelerations of the ligand binding and dissociation rates during LiGaMD3 simulations⁷⁷, which allows us to recover the original kinetic rate constants.

System Setup

The complex structures of MDM2 bound by the Nutlin 3 drug, P53 and PMI were obtained from the 5C5A, 1YCR and 3EQS PDB files, respectively. The AMBER ff14SB force field⁸¹ was used for the protein and peptide. The GAFF2 force field⁸² with AM1-BCC charge was used for the Nutlin 3 small molecule. Each system was solvated in a periodic box of TIP3P water molecules with a distance of 18 Å from the solute to the box edge using *tLeap*. Therefore, the ligand/peptide concentration was 0.00335 M in the simulation system. Appropriate number of Na⁺/Cl⁻ ions were added to achieve system neutrality.

Simulation Protocol

Each system was energy minimized and gradually heated to 300 K in 1 ns with the Langevin thermostat and harmonic restraints of 1.0 kcal/mol/Å² on all non-hydrogen atoms of the protein and the ligand using the AMBER23 software⁸³. The simulation system was firstly energy minimized with 1.0 kcal/mol/Å² constraints on the heavy atoms of the proteins, including the steepest descent minimization for 50,000 steps and conjugate gradient minimization for 50,000 steps. The system was then heated from 0 K to 300 K for 200 ps. It was further equilibrated using

the NVT ensemble at 300 K for 200 ps and the NPT ensemble at 300 K and 1 bar for 1 ns with 1 kcal/mol/Å² constraints on the heavy atoms of the protein, followed by 2 ns short cMD without any constraint. The LiGaMD3 simulations proceeded with 2 ns short cMD to collect the potential statistics, 50.0 ns LiGaMD3 equilibration after adding the boost potential and then three independent 2,000 ns production runs. It provided more powerful sampling to set the threshold energy for applying the boost potential to the upper bound (i.e., $E = V_{\min} + 1/k$) in our previous study ligand dissociation and binding using LiGaMD⁷⁶. Therefore, the threshold energy for applying the ligand essential non-bonded potential (first boost) and the remaining non-bonded potential energy of the system (second boost) were set to the upper bound in the LiGaMD3 simulations. The threshold energy for applying the third boost to the system bonded energy potential was set to the lower bound. In order to observe ligand/peptide dissociation during LiGaMD3 production simulations while keeping the boost potential as low as possible for accurate energetic reweighting, the (σ_{0P} , σ_{0D} , σ_{0B}) parameters were set to (2.0 kcal/mol, 6.0 kcal/mol, 6.0 kcal/mol) for the LiGaMD3 simulations of the MDM2 bound by the Nutlin 3, PMI and P53. LiGaMD3 production simulation frames were saved every 0.4 ps for analysis. In the LiGaMD simulations performed for comparison, the (σ_{0P} , σ_{0D}) parameters were set to (4.8 kcal/mol, 6.0 kcal/mol) for simulations of the MDM2-Nutlin 3 system and (8.0 kcal/mol, 6.0 kcal/mol) for MDM2-PMI system, respectively. The threshold energies for applying the boosts in the LiGaMD simulations were set to the upper bound. Example simulation files of the LiGaMD3 simulations of the MDM2-Nutlin 3 system are included in the Supporting Information.

Simulation Analysis

The VMD⁸⁴ and CPPTRAJ⁸⁵ tools were used for simulation analysis. The number of ligand dissociation and binding events (N_D and N_B) and the ligand/peptide binding and unbinding time

periods (τ_B and τ_U) were recorded from individual simulations (**Tables 1 & S1**). With high fluctuations, τ_B and τ_U were recorded for only the time periods longer than 1 ns. The 1D and 2D free energy profiles, as well as the ligand binding free energy, were calculated through energetic reweighting of the LiGaMD3 simulations. The center-of-mass distance between the small-molecule/peptide and the protein pocket (defined by protein residues within 5 Å of the ligand, denoted as $d_{\text{MDM2-substrate}}$) and small-molecule heavy atom or peptide backbone RMSDs relative to the PDB structures with the protein aligned were chosen as the reaction coordinates. The bin size was set to 1.0 Å. The cutoff for the number of simulation frames in one bin was set to 500. The ligand binding free energies (ΔG) were calculated using the binding kinetic rates as $\Delta G = -RT \ln(k_{\text{off}}/k_{\text{on}})$. The ligand dissociation and binding rate constants (k_{on} and k_{off}) were calculated from the LiGaMD3 and LiGaMD simulations with their accelerations analyzed using the Kramers' rate theory (**Table S2**).

Table 1. Summary of LiGaMD3 simulations performed on small molecule and peptide binding to the MDM2. ΔV is the total boost potential. N_D and N_B are the number of observed ligand dissociation and binding events, respectively. ΔG_{sim} and ΔG_{exp} are the ligand-MDM2 binding free energies obtained from LiGaMD3 simulations and experiments, respectively. ^a The simulation binding free energy is estimated using $\Delta G_{\text{sim}} = -RT \ln(k_{\text{off}}/k_{\text{on}})$.

System	Method	ID	N _B	N _D	ΔV (kcal/mol)	ΔG_{sim}^a (kcal/mol)	ΔG_{exp} (kcal/mol)	
MDM2-Nutline3	LiGaMD	Sim1	5	6	94.92±4.01	-10.18±2.22	-10.96	
		Sim2	3	3				
		Sim3	3	4				
MDM2-Nutline3	LiGaMD3	Sim1	6	5	12.94±4.05	-11.02±0.59		
		Sim2	7	7				
		Sim3	5	5				
MDM2-PMI	LiGaMD3	Sim1	4	4	45.11±7.04	-11.86±1.16	-12.02	
		Sim2	6	5				
		Sim3	6	6				
MDM2-P53	LiGaMD3	Sim1	5	6	45.59±7.03	-10.59±0.11	-9.27	
		Sim2	4	5				
		Sim3	5	6				

Results

Microsecond LiGaMD3 simulations captured repetitive small-molecule and peptide dissociation and rebinding to the MDM2

Both LiGaMD and LiGaMD3 have effectively captured the binding and dissociation processes of the Nutlin 3 small molecule to the MDM2 protein across all three independent 2,000 ns simulations (**Figs. 1B & 1C**). However, LiGaMD encountered difficulty in capturing the rebinding of the PMI peptide to the protein, as no frames with peptide RMSD < 5 Å were observed in all three 2,000 ns simulations (**Fig. 1F**). In contrast, LiGaMD3 demonstrated consistent performance in successfully capturing the repetitive binding and dissociation of the PMI peptide in the MDM2 (**Fig. 1E**). Moreover, an additional system wherein MDM2 is bound by the peptide P53 was included to further evaluate the performance of LiGaMD3 (**Fig. 2**). LiGaMD3 could capture multiple times of P53 binding and dissociation in 2000ns simulations (**Fig. 2**).

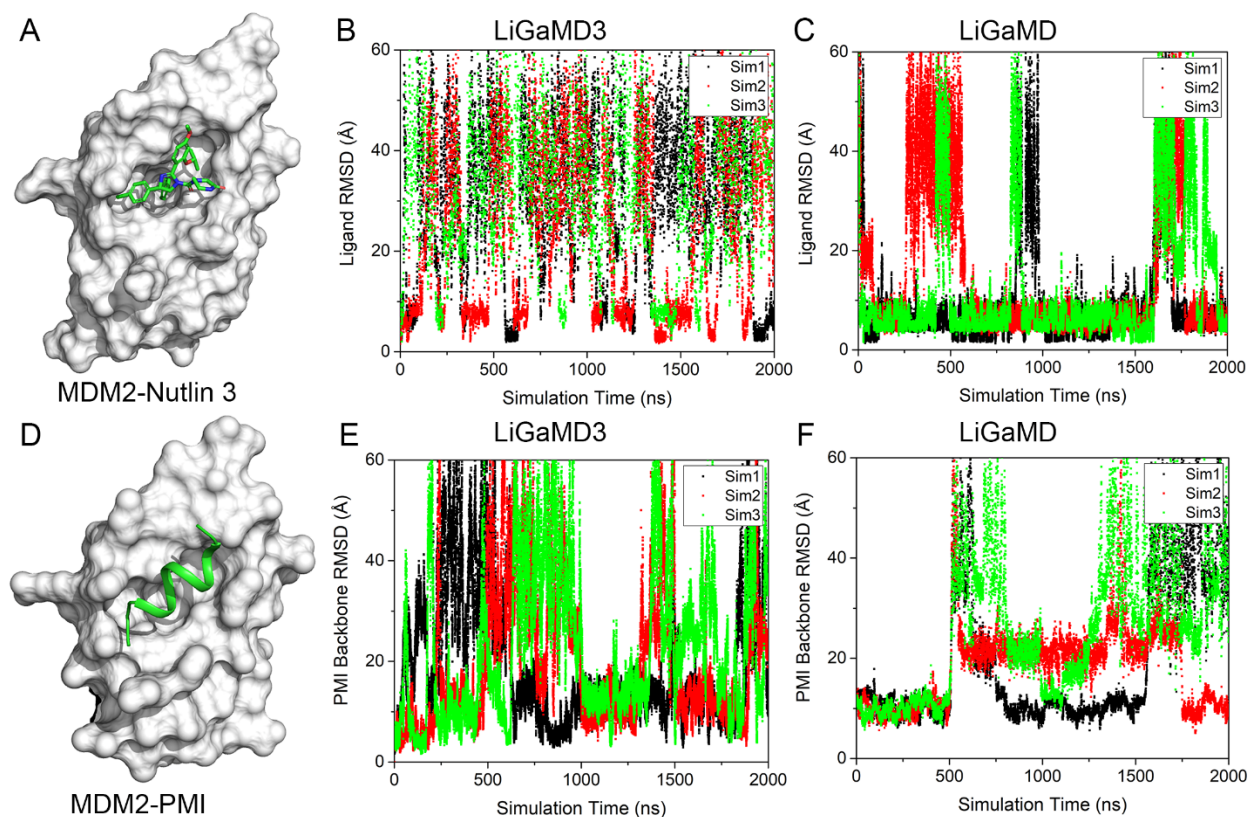


Figure 1. Comparison of LiGaMD and LiGaMD3 simulations on the MDM2 protein bound by the Nutlin 3 small molecule and PMI peptide: Computational models of the MDM2 bound by the Nutlin 3 small molecule (A) and PMI peptide (D); Time courses of ligand root-mean-square deviation (RMSD) relative to the experimental bound structure (PDB ID: 5C5A) in the MDM2-Nutlin 3 system calculated from LiGaMD3 (B) and LiGaMD (C) simulations, respectively. Time courses of peptide RMSD relative to the experimental bound structure (PDB ID: 3EQB) in the MDM2-PMI calculated from LiGaMD3 (E) and LiGaMD2 (F) simulations, respectively.

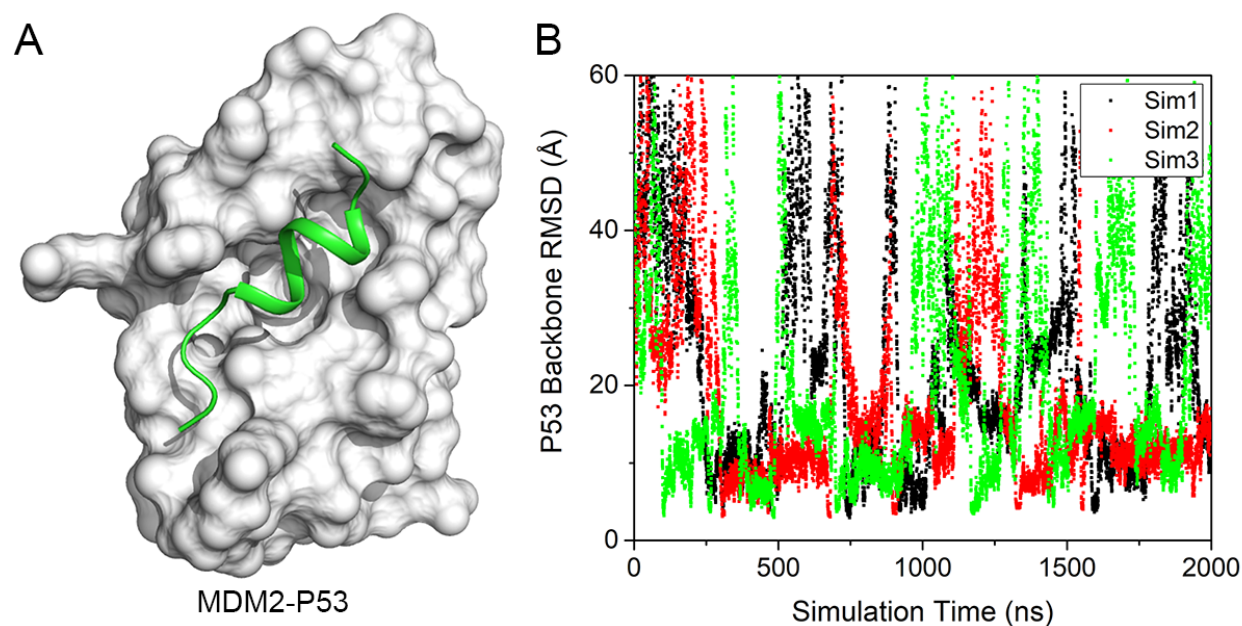


Figure 2. (A) Computational model of the MDM2 bound by the P53 peptide; (B) Time courses of peptide P53 RMSD relative to the experimental bound structure (PDB ID: 1YCR) in the MDM2-P53 system calculated from LiGaMD3 simulations.

The LiGaMD3 simulations of the MDM2-Nutlin 3 system yielded an average boost potential of 12.92 kcal/mol with a standard deviation of 4.05 kcal/mol (**Table 1**). In contrast, achieving ligand dissociation and binding required a significantly larger boost in LiGaMD, with an average boost potential of 94.92 kcal/mol and a standard deviation of 4.01 kcal/mol (**Table 1**). For the MDM2-PMI and MDM2-P53 systems, LiGaMD3 simulations recorded average boost potentials of 45.11 ± 7.04 kcal/mol and 45.59 ± 7.03 kcal/mol, respectively (**Table 1**). Notably, the boost potentials applied in simulations of peptide-protein systems are much higher compared to that of small molecule-MDM2 system. LiGaMD3 required substantially smaller boosts than LiGaMD simulations, indicating its enhanced efficiency in sampling, which was also advantageous for accurate energetic and kinetic reweighting.

RMSDs of the small-molecule/peptide relative to their experimental bound structures with the MDM2 protein aligned were computed (**Figs. 1, 2 & Table 1**) to calculate the number of small-molecule/peptide dissociation (N_D) and binding (N_B) events in each of the 2,000 ns LiGaMD3 simulations. With close examination of the ligand/peptide binding trajectories, RMSD cutoffs of the ligand unbound and bound states were set to $>15 \text{ \AA}$ and $<5.0 \text{ \AA}$, respectively. Due to fluctuations in small-molecule/peptide-protein interactions, we recorded only the corresponding binding and dissociation events that lasted for more than 1.0 ns. In 2,000 ns simulations of the MDM2-Nutlin 3 system, LiGaMD3 consistently captured 5-7 binding and 5-7 dissociation events, whereas LiGaMD captured only 3-5 binding and 3-6 dissociation events (**Fig. 1 & Table 1**). The total number of binding events recorded in LiGaMD3 was 18, compared to 11 in LiGaMD. The total number of dissociation events in LiGaMD3 and LiGaMD was 17 and 13, respectively. Hence, LiGaMD3 demonstrated improved efficiency in capturing both binding and dissociation events compared to LiGaMD. Additionally, no rebinding events were observed in LiGaMD simulations of PMI to MDM2 (**Fig. 1F**), whereas each 2,000 ns LiGaMD3 simulation successfully captured 4-6 binding and 4-6 dissociation events, indicating its superior capability in capturing flexible peptide-protein interactions (**Fig. 1E & Table 1**). Similar numbers of peptide dissociation (5-6) and binding (4-5) events were observed in simulations of the MDM2-P53 system (**Fig. 2 & Table 1**). In summary, LiGaMD3 simulations successfully captured repetitive dissociation and rebinding events of both small-molecules and flexible peptides to the MDM2 on three model systems: Nutlin 3 bound to MDM2 (MDM2-Nutlin 3), PMI bound to MDM2 (MDM2-PMI), and P53 bound to MDM2 (MDM2-P53) (**Figs. 1, 2 & S1**).

Ligand and peptide binding kinetic rates and free energies calculated from LiGaMD3

agreed with experimental data.

The successful simulations of repetitive small-molecule and flexible peptide binding and dissociation in LiGaMD3 allowed us to predict the small-molecule and peptide binding kinetic rate constants (**Fig. S2 & Table 2**). We recorded the time periods for the small-molecule and peptide found in the bound (τ_B) and unbound (τ_U) states throughout the LiGaMD and LiGaMD3 simulations. Without reweighting, the binding rate constant (k_{on}^*) and dissociation rate constant (k_{off}^*) for Nutlin 3 were directly calculated from the LiGaMD trajectories as $1.69 \pm 0.28 \times 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $3.68 \pm 2.10 \times 10^6 \text{ s}^{-1}$ (**Table 2**). In comparison, in LiGaMD3, these rate constants were calculated as $1.16 \pm 0.28 \times 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $2.77 \pm 0.75 \times 10^7 \text{ s}^{-1}$, respectively (**Table 2**). The peptide binding rate constants (k_{on}^*) were directly calculated from the LiGaMD3 trajectories as $8.57 \pm 1.37 \times 10^8 \text{ M}^{-1} \cdot \text{s}^{-1}$ and $1.02 \pm 0.29 \times 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ for the MDM2-PMI and MDM2-P53 systems, respectively (**Table 2**).

Table 2 Comparison of kinetic rates obtained from experimental data and LiGaMD3 simulations for ligand binding to MDM2. k_{on} and k_{off} are the kinetic dissociation and binding rate constants, respectively, from experimental data or LiGaMD3 simulations with reweighting using Kramers' rate theory. k_{on}^* and k_{off}^* are the accelerated kinetic dissociation and binding rate constants calculated directly from LiGaMD3 simulations without reweighting. $\Delta \log k_{on} = \log k_{on}^{sim} - \log k_{on}^{exp}$ $\Delta \log k_{off} = \log k_{off}^{sim} - \log k_{off}^{exp}$

System	Method	$k_{on} (\text{M}^{-1} \cdot \text{s}^{-1})$	$\Delta \log(k_{on})$	$k_{off} (\text{s}^{-1})$	$\Delta \log(k_{off})$	$k_{on}^* (\text{M}^{-1} \cdot \text{s}^{-1})$	$k_{off}^* (\text{s}^{-1})$
MDM2-Nutlin	Experiment	3.3×10^7	-	0.48	-	-	-
	LiGaMD	$5.26 \pm 0.65 \times 10^9$	1.82	69.51 ± 58.37	1.36	$1.69 \pm 0.28 \times 10^9$	$3.68 \pm 2.10 \times 10^6$
	LiGaMD3	$8.29 \pm 4.80 \times 10^8$	1.37	15.45 ± 4.69	1.39	$1.16 \pm 0.31 \times 10^9$	$2.77 \pm 0.75 \times 10^7$
MDM2-PMI	Experiment	5.27×10^8	-	0.037	-	-	-
	LiGaMD3	$5.76 \pm 4.80 \times 10^8$	0.04	2.66 ± 1.73	1.85	$8.51 \pm 1.37 \times 10^8$	$3.41 \pm 0.96 \times 10^7$
MDM2-P53	Experiment	9.2×10^6	-	2.06	-	-	-
	LiGaMD3	$8.03 \pm 7.42 \times 10^8$	1.94	28.0 ± 19.2	1.13	$1.02 \pm 0.29 \times 10^9$	$1.95 \pm 0.73 \times 10^7$

Next, we performed reweighting on the LiGaMD and LiGaMD3 simulations of ligand-MDM2 systems to calculate acceleration factors for the small-molecule/peptide binding and dissociation processes (**Table S2**) and to recover the original kinetic rate constants using the Kramers' rate theory (**Table 2**). In the LiGaMD simulations, the dissociation free energy barrier (ΔF_{off}) significantly decreased from 9.10 ± 1.26 kcal/mol in the reweighted PMF profiles to 2.77 ± 1.76 kcal/mol in the modified PMF profiles for the system of MDM2-Nutlin 3 (**Fig. S1** and **Table S2**). Similarly, for the MDM2-Nutlin 3, MDM2-PMI, and MDM2-P53 systems, the dissociation free energy barrier (ΔF_{off}) significantly decreased from 9.65 ± 0.76 , 9.05 ± 0.23 , 7.23 ± 0.40 kcal/mol in the reweighted PMF profiles to 0.79 ± 0.10 , 0.89 ± 0.07 , 0.67 ± 0.18 kcal/mol in the modified PMF profiles in LiGaMD3 simulations, respectively (**Table S1** and **Fig. S1**). Curvatures of the reweighted (w) and modified (w^* , no reweighting) free energy profiles were calculated near the ligand Bound ("B") and Unbound ("U") low-energy wells and the energy barrier ("Br"), as well as the ratio of apparent diffusion coefficients calculated from LiGaMD and LiGaMD3 simulations with reweighting (D) and without reweighting (modified, D^*) (**Table S2**). According to the Kramers' rate theory, the association and dissociation of the Nutline 3 small molecule in LiGaMD were accelerated by 0.32 and 5.30×10^4 times. In contrast, in LiGaMD3, the association and dissociation of the Nutlin 3 were accelerated by 1.40 and 1.79×10^6 times, respectively. Moreover, the association of the peptide in the LiGaMD3 was accelerated by 1.47 and 1.27 times for the MDM2-PMI and MDM2-P53 systems, respectively. While the peptide dissociation was significantly accelerated by 1.28×10^7 and 6.96×10^5 times for the MDM2-PMI and MDM2-P53 systems, respectively. Therefore, the reweighted k_{on} in the MDM2-Nutlin 3 system with LiGaMD and LiGaMD3 were calculated as $5.26 \pm 0.65 \times 10^9$ M⁻¹·s⁻¹ and $8.29 \pm 4.80 \times 10^8$ M⁻¹·s⁻¹, respectively, being in consistent with the experimental values of 3.3×10^7 M⁻¹·s⁻¹. Similarly, for the MDM2-PMI

and MDM2-P53 systems, the reweighted k_{on} values were predicted as $5.76 \pm 4.80 \times 10^8$ and $8.03 \pm 7.42 \times 10^8$, respectively, being consistent with the corresponding experimental values⁷⁶ of 5.27×10^8 and $9.20 \times 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$ (**Table 2**). The reweighted k_{off} values for the Nutlin 3 in the MDM2-Nutline 3 with LiGaMD and LiGaMD3 were calculated as $69.51 \pm 58.37 \text{ s}^{-1}$ and $15.45 \pm 4.69 \text{ s}^{-1}$, being in accordance with the experimental of 0.48 s^{-1} . For the peptide in the MDM2-PMI and MDM2-P53 systems, the reweighted peptide k_{off} were calculated from LiGaMD3 simulations as 2.66 ± 1.73 , $28.0 \pm 19.2 \text{ s}^{-1}$, in agreement with the corresponding experimental values⁷⁶ of 0.037 and 2.06 s^{-1} , respectively.

Based on the ligand binding kinetic rates (k_{on} and k_{off}), we calculated the ligand binding free energies as $\Delta G = -RT \ln(k_{off}/k_{on})$. The resulting binding free energies in the MDM2-Nutlin 3 system with LiGaMD and LiGaMD3 were $-10.18 \pm 2.22 \text{ kcal/mol}$ and $-11.02 \pm 0.59 \text{ kcal/mol}$, respectively, demonstrating high consistency with the experimental value of -10.96 kcal/mol . In the MDM2-PMI and MDM2-P53 systems (**Table 1**), the calculated peptide binding free energy values were $-11.86 \pm 1.16 \text{ kcal/mol}$ and $-10.59 \pm 0.11 \text{ kcal/mol}$, exhibiting strong agreement with the corresponding experimental values of -12.02 kcal/mol and -9.27 kcal/mol respectively. The root-mean square error (RMSE) of binding free energy for the three systems was only 0.94 kcal/mol . Hence, LiGaMD3 simulations achieved both efficient sampling and accurate small-molecule/peptide binding thermodynamics and kinetics calculations.

Multiple ligand binding and dissociation pathways were identified from LiGaMD3 simulations.

We closely examined the LiGaMD3 trajectories to explore the pathways involved in the small-molecule/peptide binding and dissociation of the MDM2. Two primary pathways were identified for the binding and dissociation, denoted as “pathway 1” (residues 65-95, including motifs $\beta 3$, $\beta 1'$, $\alpha 1'$ and $\beta 2'$) and “pathway 2” (residues 97-106, $\alpha 2'$ helix) (**Figs. 3 and S2**). These pathways were consistently observed in the binding and dissociation of Nutlin 3, PMI and P53. Binding of Nutlin 3 via pathways 1 and 2 were observed 13 times and 5 times, respectively (**Fig. 3B**). The same pathways 1 and 2 were identified in the simulations of the MDM2-PMI and MDM2-P53 systems. Peptide binding in the MDM2-PMI and MDM2-P53 systems occurred along pathways 1 and 2 for 9 and 7 times, respectively (**Fig. 3B**). Similarly, peptide P53 binding events along pathways 1 and 2 were 8 and 6, respectively (**Fig. 3B**). The same pathways were identified for the dissociation of the MDM2-Nutlin 3, MDM2-PMI and MDM2-P53 systems (**Fig. 3C**). Dissociation of Nutlin-3 via pathways 1 and 2 were observed 11 and 6 times, respectively (**Fig. 3C**). Peptide dissociation in the MDM2-PMI system along pathways 1 and 2 occurred 9 and 9 times, respectively (**Fig. 3C**). Similarly, peptide P53 dissociation along pathways 1 and 2 occurred 6 and 8 times, respectively (**Fig. 3C**). Note that the number of binding/dissociation events along different pathways were counted directly from the simulations without reweighting. Nevertheless, with GaMD enhanced sampling formula, the overall shape of the system free energy profiles should be maintained even with the boost potential⁴⁹. This suggests that the absolute number of the binding and dissociation events may not be meaningful, but the relative preference of the pathways derived from the LiGaMD3 simulations could be still used. Moreover, one can use the reweighted free energy profiles (e.g., those in **Fig. 4**) to examine the energetic preference of the ligand pathways.

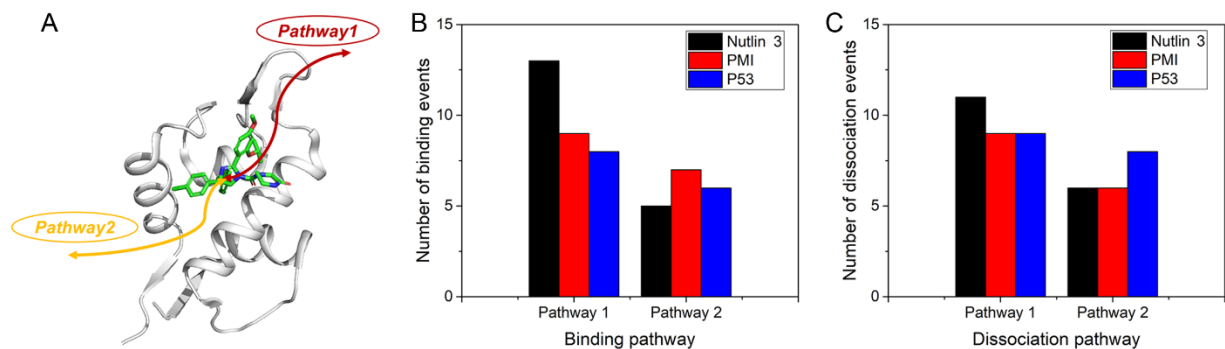


Figure 3. Pathways of ligand/peptide binding and dissociation in the MDM2 protein. (A) Cartoon representation of the protein. Binding and dissociation pathways are denoted by the arrow lines. Number of binding (B) and dissociation (C) events through the different pathways captured by the LiGaMD3 simulations.

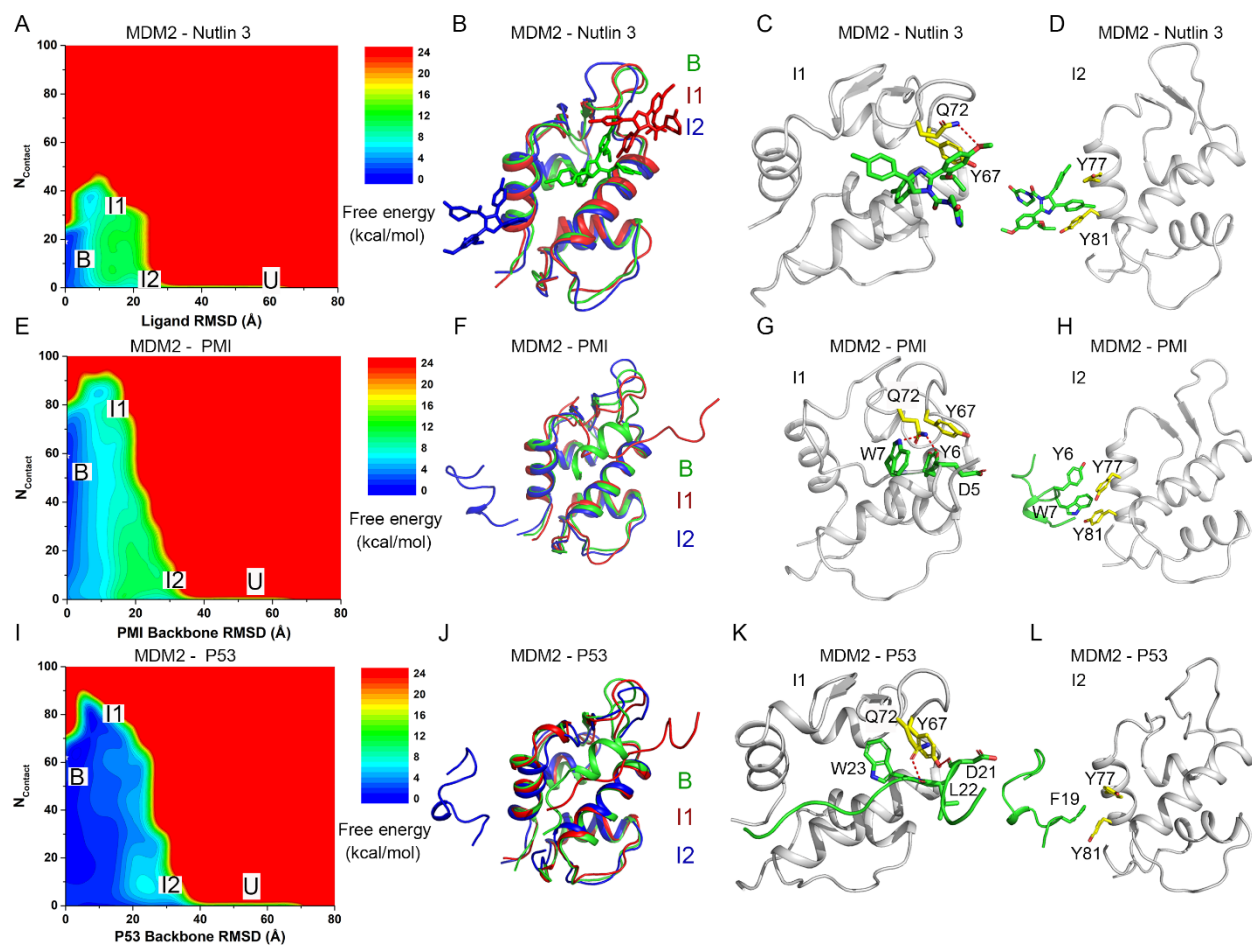


Figure 4. 2D Potential of Mean Force (PMF) free energy profiles and low-energy conformational states of ligand/peptide binding to the MDM2: **(A)** 2D PMF profile regarding the ligand heavy atom RMSD and the number of contacts between the ligand and residues 65-95 of MDM2 in the LiGaMD3 simulations of Nutlin 3 binding to the MDM2 protein; **(B)** Low-energy “Intermediate” conformations “I1” (blue) and “I2” (red) as identified from the 2D PMF profiles of Nutlin 3 binding to MDM2 protein; **(C-D)** Important ligand-MDM2 interactions in the low-energy conformations “I1” (**C**) and I2 (**D**). **(E)** 2D PMF profile regarding the peptide backbone RMSD and the number of contacts between the PMI and residues 65-95 of the MDM2 in the LiGaMD3 simulations of PMI binding to the MDM2 protein; **(F)** Low-energy “Intermediate” conformations “I1” (blue) and “I2” (red) as identified from the 2D PMF profiles of PMI binding to MDM2 protein; **(G-H)** Important PMI-MDM2 interactions in the low-energy conformations “I1” (**G**) and I2 (**H**). **(I)** 2D PMF profile regarding the peptide backbone RMSD and the number of contacts between the P53 and residues 65-95 of the MDM2 in the LiGaMD3 simulations of P53 binding to the MDM2 protein; **(J)** Low-energy “Intermediate” conformations “I1” (blue) and “I2” (red) as identified from the 2D PMF profiles of P53 binding to MDM2 protein; **(K-L)** Important P53-MDM2 interactions in the low-energy conformations “I1” (**K**) and I2 (**L**).

Small-molecule and Peptide binding to the MDM2 involved Induced Fit

After identifying the pathway 1, which involves motifs $\beta 3$ - $\beta 1'$ - $\alpha 1'$ - $\beta 2'$ (residues 65-95), we further investigated the relationship between conformational changes within this region upon small-molecule/peptide binding. Therefore, the ligand RMSD and the number of contacts between the ligand and residues 65-95 in MDM2 (denoted as N_{contact}) were used as reaction coordinates to calculate 2D PMF profiles (**Figs. 4A-4C**). Four low-energy states were identified in the 2D PMF profile of the MDM2-Nutlin 3 system including the Bound (“B”), Intermediate (“I1” and “I2”), and Unbound (“U”) (**Fig. 4A**). The ligand RMSD and N_{contact} of these states centered around (3.0 Å, 24), (9.0 Å, 37), (20.1 Å, 0), and (50.0 Å, 0), respectively (**Fig. 4A**). In the MDM2-PMI system,

four low-energy states were identified: Bound (“B”), Intermediate (“I1” and “I2”), and Unbound (“U”) (**Fig. 4E**), with the PMI peptide RMSD and N_{contact} centered around (3.5 Å, 60), (9.7 Å, 85), (27.5 Å, 0), and (50 Å, 0), respectively (**Fig. 4E**). Similarly, in the MDM2-P53 system, four low-energy states were observed: Bound (“B”), Intermediate (“I1” and “I2”), and Unbound (“U”) states (**Fig. 4I**), with the P53 peptide RMSD and N_{contact} centered around (4.0 Å, 64), (8.5 Å, 78), (30.5 Å, 0), and (60.8 Å, 0), respectively (**Fig. 4I**). Compared to the Bound state, the intermediate “I1” and “I2” states exhibited significant conformational alterations in the MDM2-Nutlin 3, MDM2-PMI, and MDM2-P53 systems (**Figs. 4B, 4F & 4J**). In the intermediate “I1” and “I2” states, motifs $\beta 3$ - $\beta 1'$ - $\alpha 1'$ - $\beta 2'$ (residues 65-95) in the MDM2-Nutlin3 system moved outward significantly compared to the X-ray Bound structures, resulting in the opening of the binding pocket (**Fig. 4B**). Two critical interactions were identified in “I1” state: a hydrogen bond interaction between MDM2:Q72 and Nutlin 3 and aromatic interaction between MDM2:Y67 and Nutlin 3 (**Fig. 4C**). While in the “I2” state, Π - Π interactions were observed between MDM2:Y77 or MDM2:Y81 and Nutlin 3 (**Figs. 4C&4D**). In the MDM2-PMI system, significant conformational changes occurred upon peptide PMI binding, involving particularly motifs $\beta 3$ - $\beta 1'$ - $\alpha 1'$ - $\beta 2'$ (residues 65-95), resulting in distinct open and closed conformations in the “I1” and “I2” states, respectively (**Fig. 4F**). In the “I1” state, hydrogen bonds were formed between MDM2:Q72 and PMI:W7, MDM2:Q72 and PMI:D5 (**Figs. 4G**). In the “I2” state, Π - Π interactions were formed between MDM2:Y77-PMI:Y6, MDM2:Y81 and PMI: W7 (**Figs. 4H**). In the MDM2-P53 system, significant conformational changes occurred upon peptide P53 binding, involving particularly motifs $\beta 3$ - $\beta 1'$ - $\alpha 1'$ - $\beta 2'$ (residues 65-95), resulting in a more closed conformation in the “I1” and “I2” states (**Fig. 4J**). In the “I1” state, hydrogen bonds were formed between MDM2:Q72 and P53:L22, MDM2:Y67 and P53:D21 (**Fig. 4K**). In the “I2” state, Π - Π interactions were formed between

MDM2:Y77 and P53:F19, MDM2:Y81 and P53:F18 (**Figs. 4K&4L**). Hence, residues Q72, Y67, Y77 and Y81 in the MDM2 protein play pivotal roles during ligand binding in the intermediate conformational states.

In order to further explore the mechanism of ligand binding to the MDM2, we computed 2D PMF free energy profiles to characterize conformational changes of both the protein and ligand during binding. The intermediate “I1” and “I2” states showed quite large conformational changes in the motifs $\beta 3\text{-}\beta 1'\text{-}\alpha 1'\text{-}\beta 2'$ (residues 65-95). Therefore, we calculated 2D PMF profiles regarding the RMSD of the ligand and the MDM2 motifs $\beta 3\text{-}\beta 1'\text{-}\alpha 1'\text{-}\beta 2'$ (residues 65-95) RMSD (denoted as Loop RMSD) relative to the experimental bound structures with the protein aligned (**Figs. 5A-5C**). For the MDM2-Nutlin 3 system, three low-energy states were identified from the 2D PMF profile, including the Bound, Intermediate “I2” and Unbound (**Figs. 5A**). The protein motifs $\beta 3\text{-}\beta 1'\text{-}\alpha 1'\text{-}\beta 2'$ at the peptide-binding site adopted the “Open” conformation in the I2 state (**Figs. 5A and 5B**). The peptide and loop RMSDs centered around (4.5 Å, 1.0 Å), (18.0 Å, 2.9 Å) and (59.0 Å, 2.0 Å) in the Bound “B”, Intermediate “I2” and Unbound “U” states, respectively (**Fig. 5A**). For the MDM2-PMI system, three low-energy states were identified from the 2D PMF profile, including the Bound “B”, Intermediate “I1” and Unbound “U”. The peptide and loop RMSDs centered around (5.0 Å, 0.8 Å), (10.2 Å, 3.0 Å) and (58.2 Å, 2.0 Å) in the “B”, “I1” and “U” states, respectively (**Figs. 5B**). Four low-energy conformational states were identified in the MDM2-P53 system. The peptide and loop RMSDs centered around (5.2 Å, 0.8 Å), (10.5 Å, 3.9 Å), (31.0 Å, 4.0 Å) and (61.0 Å, 1.9 Å) in the Bound “B”, Intermediate “I1” and “I2”, and Unbound “U” states, respectively (**Fig. 5C**).

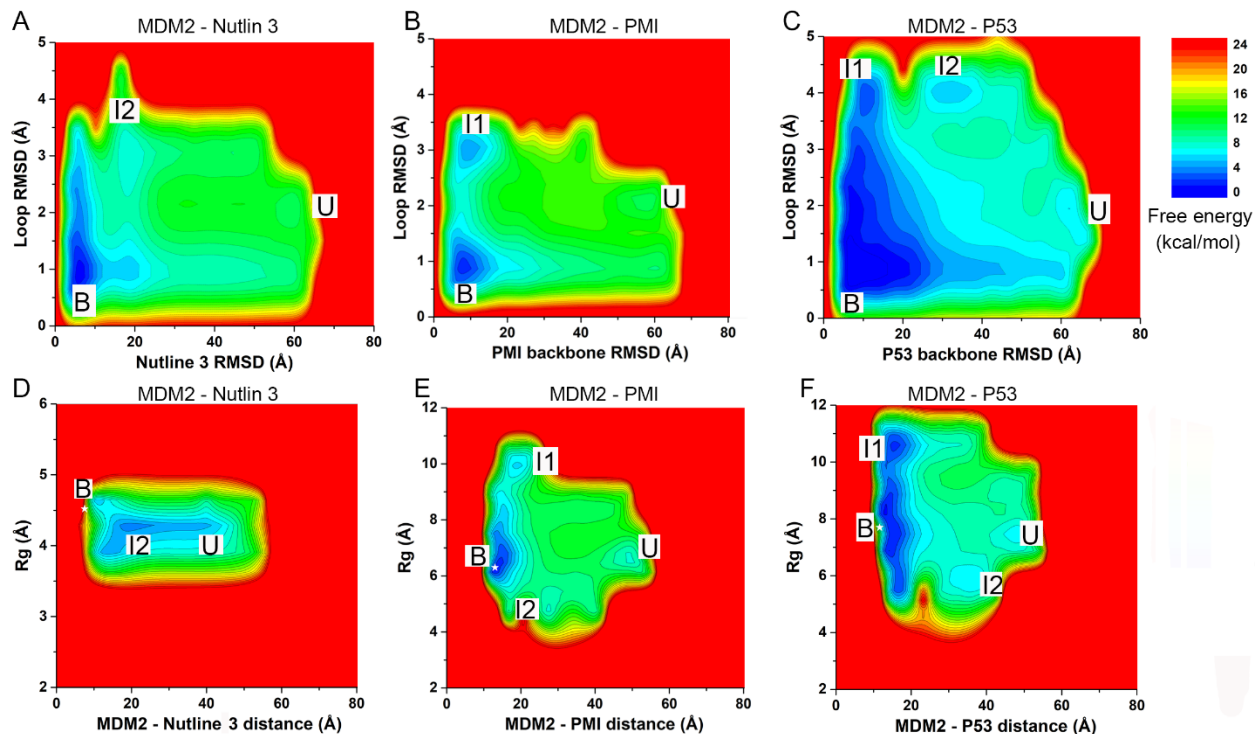


Figure 5. (A-C) 2D PMF profiles regarding the ligand heavy atom RMSD or peptide backbone RMSD and the MDM2 loop (residues 65-95) RMSD relative to their corresponding experimental bound structure in the LiGaMD3 simulations of Nutlin 3(A), PMI peptide (B) and P53 peptide (C) binding to the MDM2 protein; (D-F) 2D PMF profiles regarding the distance between the ligand/peptide and MDM2 binding pocket and the R_g of the substrates in the LiGaMD3 simulations of Nutlin 3 (D), PMI peptide (E) and P53 (F) binding to the MDM2 protein.

In addition, we examined the conformational dynamics exhibited by the small molecule and peptides during their binding processes. In this regard, the ligand radius of gyration (R_g) was calculated and monitored for possible conformational changes. The small-molecule/peptide R_g and the center-of-mass distance between protein pocket and ligand (denoted as $d_{\text{MDM2-substrate}}$) were used as reaction coordinates to calculate 2D PMF profiles. From the reweighted 2D PMF profiles (Fig. 5D-5F), we identified a low-energy “Bound” state in all three systems, for which the $d_{\text{MDM2-substrate}}$ and R_g in the MDM2-Nutlin3, MDM2-PMI and MDM2-P53 systems centered around (8.6

Å, 4.4 Å), (14.0 Å, 6.3 Å) and (13.0 Å, 8.3 Å), respectively. This suggested that successful sampling of complete small-molecule/peptide binding was captured in the LiGaMD3 simulations. Notably, in the intermediate states, the peptides sampled a wider range of R_g and the protein motifs $\beta 3$ - $\beta 1'$ - $\alpha 1'$ - $\beta 2'$ exhibited higher RMSDs compared with the bound states. Therefore, the small-molecule and peptides binding to the MDM2 protein showed predominantly an “induced-fit” mechanism.

Discussions

We have presented a new LiGaMD3 method to improve sampling efficiency and accurately predict thermodynamic and kinetic properties associated with the binding of small molecules and highly flexible peptides. LiGaMD3 works by selectively boosting the essential non-bonded interaction potential energy of the ligand, as well as the remaining non-bonded potential energy and all the bonded potential of the system. Non-bonded potential interactions play a critical role in ligand dissociation and rebinding, while the bonded potentials mainly contribute to conformational changes of the system. Utilizing microsecond timescale simulations, LiGaMD3 effectively captures repetitive dissociation and rebinding processes of both small molecules and peptides in three model systems of MDM2 bound by different small molecules and flexible peptides. These simulations allowed for simultaneous predictions of ligand/peptide binding free energies and kinetic rate constants. However, it remains challenging to achieve highly accurate predictions of ligand binding kinetics. The prediction errors of binding ($\Delta \log k_{on}$) and dissociate ($\Delta \log k_{off}$) rate constants calculated from LiGaMD3 simulations are mostly in the range of -2 to 2 (**Table 1**), which fall within the acceptable accuracy ⁹.

In order to achieve proper acceleration levels, six key parameters need to be fine-tuned in LiGaMD3, including the iEP, iED, iEB, σ_{0P} , σ_{0D} , and σ_{0B} . It generally provides more powerful sampling to set the threshold energy for applying the boost potential to the upper bound (iE=2) than to the lower bound (iE=1). Due to typically high energy barriers for ligand dissociation and binding, the upper-bound mode (iEP=2

and $iED=2$) are usually recommended for the first and second boosts in LiGaMD3 in effective sampling of both ligand dissociation and rebinding. The third boost in LiGaMD3 focuses on conformational changes, which require lower boost in most cases. Consequently, the lower-bound mode ($iEB=1$) is generally recommended for this purpose. The energy barrier for ligand dissociation is typically higher than that for ligand binding. Starting from the bound state, once the applied boost allows the ligand to overcome the dissociation energy barrier, the same boost will likely overcome the energy barrier for ligand rebinding. However, if the boost is too high, it could become challenging for the ligand rebinding. Therefore, one rule for choosing the parameter values is that the acceleration should be sufficient to overcome the dissociation energy barrier but not too high to prevent rebinding. For σ_{0P} , one may start with a large value (e.g., 8.0 kcal/mol) and then decrease it gradually to find one that enables ligand dissociation while maintaining the ligand rebinding ability in LiGaMD3 testing equilibration simulations, for which default values could be used for σ_{0D} , and σ_{0B} (6.0 kcal/mol). Once a final set of the parameters is determined, one can proceed to the LiGaMD3 production simulations.

LiGaMD3 simulations revealed the critical role of nonbonded potentials in governing ligand dissociation and rebinding process, being consistent with previous computational findings^{41, 54, 86}. Non-bonded interactions have been recognized as one of the main factors that govern the ligand binding to its target protein⁶⁶⁻⁶⁸. Furthermore, our simulations identified multiple pathways for ligand binding and dissociation and revealed an “induced-fit” mechanism of ligand binding, being consistent with earlier simulation results^{22, 55, 87}. Compared with the cMD⁸⁸, Metadynamics^{28, 29}, Weighted Ensemble,²⁵ MSM⁵² and Replica Exchange MD simulations⁶⁴, LiGaMD3 offers a more efficient and user-friendly approach. LiGaMD3 also shows advantages over previous LiGaMD, particularly in its ability to accurately capture peptide binding to proteins. While microsecond cMD simulations have proven effective in capturing small molecule and highly flexible peptide binding to target proteins, the slower kinetics of ligand dissociation remain beyond the accessible timescale of cMD. Weighted Ensemble⁵⁴ and MSM⁵² methods have shown promise

in accurately predicting small molecule and peptide binding kinetics, but typically require extensive computational resources, often involving tens-of-microsecond simulations⁵². Metadynamics, with carefully designed CVs, can efficiently capture both ligand binding and unbinding. However, the predefined CVs may impose constraints on binding pathways and conformational space. The approach may also encounter challenges such as the “hidden energy barrier” problem and slow convergence if important CVs are omitted.^{89, 90} Overall, previous methods have been computationally demanding, necessitating significantly longer simulations to adequately characterize ligand binding thermodynamics and kinetics. LiGaMD3 captures the repetitive small-molecule and peptide dissociation and binding events within only microsecond simulations, offering an efficient approach to characterizing ligand binding dynamics and extending the capabilities of the LiGaMD methodology to binding of highly flexible peptides.

Acknowledgements

This work used supercomputing resources with allocation award TG-MCB180049 and BIO210039 through the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296, and project M2874 through the National Energy Research Scientific Computing Center (NERSC), which is a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231, and the Research Computing at the UNC-Chapel Hill. This work was supported in part by the National Institutes of Health (R01GM132572) and National Science Foundation (2121063).

Supporting Information

Example input files of LiGaMD3 simulations of the MDM2-Nutlin 3 system, **Tables S1, S2, Figures S1 and S2** are provided in the Supporting Information. This information is available free of charge via the Internet at <http://pubs.acs.org>.

References

- (1) Nooren, I. M.; Thornton, J. M. Diversity of protein-protein interactions. *EMBO J.* **2003**, *22* (14), 3486-3492.
- (2) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007**, *6* (3), 211-219.
- (3) Klebe, G. Applying thermodynamic profiling in lead finding and optimization. *Nat. Rev. Drug Discovery* **2015**, *14* (2), 95-110.
- (4) Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discovery* **2017**, *16* (12), 829-842.
- (5) Miura, K. An overview of current methods to confirm protein-protein interactions. *Protein Pept. Lett.* **2018**, *25* (8), 728-733.
- (6) Renaud, J.-P.; Chari, A.; Ciferri, C.; Liu, W.-t.; Rémy, H.-W.; Stark, H.; Wiesmann, C. Cryo-EM in drug discovery: achievements, limitations and prospects. *Nat. Rev. Drug Discovery* **2018**, *17* (7), 471-492.
- (7) Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A. J.; Bambrick, J. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493-500.
- (8) Krishna, R.; Wang, J.; Ahern, W.; Sturfels, P.; Venkatesh, P.; Kalvet, I.; Lee, G. R.; Morey-Burrows, F. S.; Anishchenko, I.; Humphreys, I. R.; et al. Generalized Biomolecular Modeling and Design with RoseTTAFold All-Atom. *Science* **2023**, *384*, , eadl2528.
- (9) Wang, J.; Do, H. N.; Koirala, K.; Miao, Y. Predicting biomolecular binding kinetics: A review. *J. Chem. Theory Comput.* **2023**, *19* (8), 2135-2148.
- (10) Schuetz, D. A.; de Witte, W. E. A.; Wong, Y. C.; Knasmueller, B.; Richter, L.; Kokh, D. B.; Sadiq, S. K.; Bosma, R.; Nederpelt, I.; Heitman, L. H. Kinetics for Drug Discovery: an industry-driven effort to target drug residence time. *Drug Discovery Today* **2017**, *22* (6), 896-911.
- (11) Tonge, P. J. Drug–Target Kinetics in Drug Discovery. *ACS Chem. Neurosci.* **2018**, *9* (1), 29-39.
- (12) Ahmad, K.; Rizzi, A.; Capelli, R.; Mandelli, D.; Lyu, W.; Carloni, P. Enhanced-Sampling Simulations for the Estimation of Ligand Binding Kinetics: Current Status and Perspective. *Front. Mol. Biosci.* **2022**, *9*:899805.
- (13) Ijzerman, A. P.; Guo, D. Drug - Target Association Kinetics in Drug Discovery. *Trends Biochem. Sci.* **2019**, *44* (10), 861-871.
- (14) Ferruz, N.; De Fabritiis, G. Binding Kinetics in Drug Discovery. *Mol. Inf.* **2016**, *35* (6-7), 216-226.
- (15) Zhou, Y.; Fu, Y.; Yin, W.; Li, J.; Wang, W.; Bai, F.; Xu, S.; Gong, Q.; Peng, T.; Hong, Y.; et al. Kinetics-Driven Drug Design Strategy for Next-Generation Acetylcholinesterase Inhibitors to Clinical Candidate. *J. Med. Chem.* **2021**, *64* (4), 1844-1855.
- (16) Pawnikar S, B. A., Wang J, Miao Y. . Binding Analysis Using Accelerated Molecular Dynamics Simulations and Future Perspectives. *Adv. Appl. Bioinform. Chem.* **2022**, *15*, 1-19.
- (17) Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. How Does a Drug Molecule Find Its Target Binding Site? *J. Am. Chem. Soc.* **2011**, *133* (24), 9181-9183.
- (18) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (32), 13118-13123.

- (19) Shaw, D. E.; Adams, P. J.; Azaria, A.; Bank, J. A.; Batson, B.; Bell, A.; Bergdorf, M.; Bhatt, J.; Butts, J. A.; Correia, T. Anton 3: twenty microseconds of molecular dynamics simulation before lunch. In *SC '21: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; New York, NY, 2021; pp 1-11.
- (20) Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative Characterization of the Binding and Unbinding of Millimolar Drug Fragments with Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2017**, *13* (7), 3372-3377.
- (21) Calderón, J. C.; Plut, E.; Keller, M.; Cabrele, C.; Reiser, O.; Gervasio, F. L.; Clark, T. Extended Metadynamics Protocol for Binding/Unbinding Free Energies of Peptide Ligands to Class A G-Protein-Coupled Receptors. *J. Chem. Inf. Model.* **2024**, *64* (1), 205-218.
- (22) Paul, F.; Wehmeyer, C.; Abualrous, E. T.; Wu, H.; Crabtree, M. D.; Schoneberg, J.; Clarke, J.; Freund, C.; Weikl, T. R.; Noe, F. Protein-peptide association kinetics beyond the seconds timescale from atomistic simulations. *Nat. Commun.* **2017**, *8* (1), 1095.
- (23) Robustelli, P.; Piana, S.; Shaw, D. E. Mechanism of Coupled Folding-upon-Binding of an Intrinsically Disordered Protein. *J. Am. Chem. Soc.* **2020**, *142* (25), 11092-11101.
- (24) Kappel, K.; Miao, Y. L.; McCammon, J. A. Accelerated molecular dynamics simulations of ligand binding to a muscarinic G-protein-coupled receptor. *Q. Rev. Biophys.* **2015**, *48* (4), 479-487, Review.
- (25) Saglam, A. S.; Chong, L. T. Protein-protein binding pathways and calculations of rate constants using fully-continuous, explicit-solvent simulations. *Chem. Sci.* **2019**, *10* (8), 2360-2372.
- (26) Wang, Y.; Valsson, O.; Tiwary, P.; Parrinello, M.; Lindorff-Larsen, K. Frequency adaptive metadynamics for the calculation of rare-event kinetics. *J. Chem. Phys.* **2018**, *149* (7), 072309.
- (27) Wang, Y.; Martins, J. M.; Lindorff-Larsen, K. Biomolecular conformational changes and ligand binding: from kinetics to thermodynamics. *Chem. Sci.* **2017**, *8* (9), 6466-6473, 10.1039/C7SC01627A.
- (28) Banerjee, P.; Bagchi, B. Dynamical control by water at a molecular level in protein dimer association and dissociation. *Proc. Natl. Acad. Sci. USA* **2020**, *117* (5), 2302-2308.
- (29) Antoszewski, A.; Feng, C.-J.; Vani, B. P.; Thiede, E. H.; Hong, L.; Weare, J.; Tokmakoff, A.; Dinner, A. R. Insulin dissociates by diverse mechanisms of coupled unfolding and unbinding. *J. Phys. Chem. B* **2020**, *124* (27), 5571-5587.
- (30) Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. USA* **2015**, *112* (5), E386-E391.
- (31) Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding kinetics of a p38 MAP kinase type II inhibitor from metadynamics simulations. *J. Am. Chem. Soc.* **2017**.
- (32) Kingsley, L. J.; Esquivel-Rodríguez, J.; Yang, Y.; Kihara, D.; Lill, M. A. Ranking protein-protein docking results using steered molecular dynamics and potential of mean force calculations. *J. Comput. Chem.* **2016**, *37* (20), 1861-1865.
- (33) Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* **2003**, *119* (6), 3559-3566.
- (34) Gonzalez, A.; Perez-Acle, T.; Pardo, L.; Deupi, X. Molecular basis of ligand dissociation in beta-adrenergic receptors. *Plos One* **2011**, *6* (9), e23815.
- (35) Gumbart, J. C.; Roux, B.; Chipot, C. Efficient determination of protein-protein standard binding free energies from first principles. *J. Chem. Theory Comput.* **2013**, *9* (8), 3789-3798.

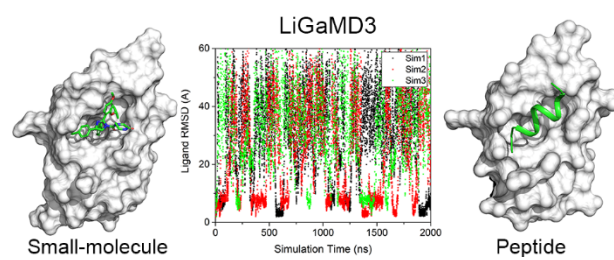
- (36) Joshi, D. C.; Lin, J. H. Delineating Protein–Protein Curvilinear Dissociation Pathways and Energetics with Naïve Multiple-Walker Umbrella Sampling Simulations. *J. Comput. Chem.* **2019**, *40* (17), 1652-1663.
- (37) Reif, M. M.; Zacharias, M. Improving the Potential of Mean Force and Nonequilibrium Pulling Simulations by Simultaneous Alchemical Modifications. *J. Chem. Theory Comput.* **2022**, *18* (6), 3873-3893.
- (38) Siebenmorgen, T.; Zacharias, M. Efficient Refinement and Free Energy Scoring of Predicted Protein–Protein Complexes Using Replica Exchange with Repulsive Scaling. *J. Chem. Inf. Model.* **2020**, *60* (11), 5552-5562.
- (39) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* **1999**, *314* (1-2), 141-151.
- (40) Sugita, Y.; Kamiya, M.; Oshima, H.; Re, S. Replica-exchange methods for biomolecular simulations. In *Biomolecular Simulations*, Springer, 2019; pp 155-177.
- (41) Niitsu, A.; Re, S.; Oshima, H.; Kamiya, M.; Sugita, Y. De novo prediction of binders and non-binders for T4 Lysozyme by gREST simulations. *J. Chem. Inf. Model.* **2019**, *59* (9), 3879-3888.
- (42) Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Ligand unbinding mechanisms and kinetics for T4 lysozyme mutants from τ RAMD simulations. *Curr. Res. Struct. Biol.* **2021**, *3*, 106-111.
- (43) Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H.-P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; et al. Estimation of Drug-Target Residence Times by τ -Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14* (7), 3859-3869.
- (44) Wang, T.; Duan, Y. Chromophore channeling in the G-protein coupled receptor rhodopsin. *J. Am. Chem. Soc.* **2007**, *129* (22), 6970-6971.
- (45) Deb, I.; Frank, A. T. Accelerating Rare Dissociative Processes in Biomolecules Using Selectively Scaled MD Simulations. *J. Chem. Theory Comput.* **2019**, *15* (11), 5817-5828.
- (46) Schuetz, D. A.; Bernetti, M.; Bertazzo, M.; Musil, D.; Eggenweiler, H. M.; Recanatini, M.; Masetti, M.; Ecker, G. F.; Cavalli, A. Predicting Residence Time And Drug Unbinding Pathway Through Scaled Molecular Dynamics. *J. Chem. Inf. Model.* **2019**, *59*(1), 535–549.
- (47) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120* (24), 11919-11929.
- (48) Wang, J.; Arantes, P. R.; Bhattarai, A.; Hsu, R. V.; Pawnikar, S.; Huang, Y. M.; Palermo, G.; Miao, Y. Gaussian accelerated molecular dynamics (GaMD): principles and applications. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2021**, *11* (5), e1521.
- (49) Miao, Y.; Feher, V. A.; McCammon, J. A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comput.* **2015**, *11* (8), 3584-3595.
- (50) Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (25), 10184-10189.
- (51) Plattner, N.; Noé, F. Protein conformational plasticity and complex ligand-binding kinetics explored by atomistic simulations and Markov models. *Nat. Commun.* **2015**, *6* (1), 1-10.
- (52) Mondal, J.; Ahalawat, N.; Pandit, S.; Kay, L. E.; Vallurupalli, P. Atomic resolution mechanism of ligand binding to a solvent inaccessible cavity in T4 lysozyme. *PLoS Comput. Biol.* **2018**, *14* (5), e1006180.

- (53) Lotz, S. D.; Dickson, A. Wepy: A Flexible Software Framework for Simulating Rare Events with Weighted Ensemble Resampling. *ACS Omega* **2020**, *5* (49), 31608-31623.
- (54) Nunes-Alves, A.; Zuckerman, D. M.; Arantes, G. M. Escape of a Small Molecule from Inside T4 Lysozyme by Multiple Pathways. *Biophys. J.* **2018**, *114* (5), 1058-1066.
- (55) Zou, R.; Zhou, Y.; Wang, Y.; Kuang, G.; Ågren, H.; Wu, J.; Tu, Y. Free Energy Profile and Kinetics of Coupled Folding and Binding of the Intrinsically Disordered Protein p53 with MDM2. *J. Chem. Inf. Model.* **2020**, *60* (3), 1551-1558.
- (56) Zwier, M. C.; Pratt, A. J.; Adelman, J. L.; Kaus, J. W.; Zuckerman, D. M.; Chong, L. T. Efficient Atomistic Simulation of Pathways and Calculation of Rate Constants for a Protein–Peptide Binding Process: Application to the MDM2 Protein and an Intrinsically Disordered p53 Peptide. *J. Phys. Chem. Lett.* **2016**, *7* (17), 3440-3445.
- (57) Zhou, G.; Pantelopulos, G. A.; Mukherjee, S.; Voelz, V. A. Bridging Microscopic and Macroscopic Mechanisms of p53-MDM2 Binding with Kinetic Network Models. *Biophys. J.* **2017**, *113* (4), 785-793.
- (58) Miao, Y.; Bhattarai, A.; Wang, J. Ligand Gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *J. Chem. Theory Comput.* **2020**, *16*(9), 5526–5547.
- (59) Wang, J.; Miao, Y. Ligand Gaussian Accelerated Molecular Dynamics 2 (LiGaMD2): Improved Calculations of Ligand Binding Thermodynamics and Kinetics with Closed Protein Pocket. *J. Chem. Theory Comput.* **2023**, *19* (3), 733-745.
- (60) Wang, L.; Berne, B. J.; Friesner, R. A. On achieving high accuracy and reliability in the calculation of relative protein-ligand binding affinities. *Proc. Natl. Acad. Sci. USA* **2012**, *109* (6), 1937-1942.
- (61) Baron, R.; McCammon, J. A. Dynamics, Hydration, and Motional Averaging of a Loop-Gated Artificial Protein Cavity: The W191G Mutant of Cytochrome c Peroxidase in Water as Revealed by Molecular Dynamics Simulations. *Biochemistry* **2007**, *46* (37), 10629-10642.
- (62) Karplus, M.; Gelin, B. R.; McCammon, J. A. Internal dynamics of proteins. Short time and long time motions of aromatic sidechains in PTI. *Biophys. J.* **1980**, *32* (1), 603-618.
- (63) Amaral, M.; Kokh, D. B.; Bomke, J.; Wegener, A.; Buchstaller, H. P.; Eggenweiler, H. M.; Matias, P.; Sirrenberg, C.; Wade, R. C.; Frech, M. Protein conformational flexibility modulates kinetics and thermodynamics of drug binding. *Nat. Commun.* **2017**, *8* (1), 2276.
- (64) Re, S.; Oshima, H.; Kasahara, K.; Kamiya, M.; Sugita, Y. Encounter complexes and hidden poses of kinase-inhibitor binding on the free-energy landscape. *Proc. Natl. Acad. Sci. USA* **2019**, *116* (37), 18404-18409.
- (65) Wilson, C.; Agafonov, R.; Hoemberger, M.; Kutter, S.; Zorba, A.; Halpin, J.; Buosi, V.; Otten, R.; Waterman, D.; Theobald, D. Using ancient protein kinases to unravel a modern cancer drug's mechanism. *Science* **2015**, *347* (6224), 882-886.
- (66) Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. Molecular determinants of drug–receptor binding kinetics. *Drug discovery today* **2013**, *18* (13-14), 667-673.
- (67) Klebe, G. The use of thermodynamic and kinetic data in drug discovery: decisive insight or increasing the puzzlement? *ChemMedChem* **2015**, *10* (2), 229-231.
- (68) Stank, A.; Kokh, D. B.; Fuller, J. C.; Wade, R. C. Protein Binding Pocket Dynamics. *Acc. Chem. Res.* **2016**, *49* (5), 809-815.
- (69) Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discovery* **2003**, *2* (7), 527-541.

- (70) Wang, W.; Qin, J. J.; Rajaei, M.; Li, X.; Yu, X.; Hunt, C.; Zhang, R. Targeting MDM2 for novel molecular therapy: Beyond oncology. *Med. Res. Rev.* **2020**, *40* (3), 856-880.
- (71) Vanommeslaeghe, K.; MacKerell Jr, A. CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850* (5), 861-871.
- (72) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24* (16), 1999-2012.
- (73) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91* (1-3), 275-282.
- (74) Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **2014**, *10* (7), 2677-2689.
- (75) Wang, J.; Miao, Y. Protein-protein interaction-Gaussian accelerated molecular dynamics (PPI-GaMD): Characterization of protein binding thermodynamics and kinetics. *J. Chem. Theory Comput.* **2022**, *18* (3), 1275-1285.
- (76) Wang, J.; Miao, Y. Peptide Gaussian accelerated molecular dynamics (Pep-GaMD): Enhanced sampling and free energy and kinetics calculations of peptide binding. *J. Chem. Phys.* **2020**, *153* (15), 154109.
- (77) Miao, Y. Acceleration of biomolecular kinetics in Gaussian accelerated molecular dynamics. *J. Chem. Phys.* **2018**, *149* (7), 072308.
- (78) Doshi, U.; Hamelberg, D. Extracting realistic kinetics of rare activated processes from accelerated molecular dynamics using Kramers' theory. *J. Chem. Theory Comput.* **2011**, *7* (3), 575-581.
- (79) Frank, A. T.; Andricioaei, I. Reaction coordinate-free approach to recovering kinetics from potential-scaled simulations: application of kramers' rate theory. *J. Phys. Chem. B* **2016**, *120* (33), 8600-8605.
- (80) Hamelberg, D.; Shen, T.; Andrew McCammon, J. Relating kinetic rates and local energetic roughness by accelerated molecular-dynamics simulations. *J. Chem. Phys.* **2005**, *122* (24), 241103.
- (81) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696-3713.
- (82) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25* (9), 1157-1174.
- (83) Case, D. A.; D.S. Cerutti; T.E. Cheatham, I.; T.A. Darden, R. E. D., T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, J.Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. Amber 2022. **2022**.
- (84) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14* (1), 33-38.
- (85) Roe, D. R.; Cheatham, T. E., 3rd. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084-3095.

- (86) Feher, V. A.; Schiffer, J. M.; Mermelstein, D. J.; Mih, N.; Pierce, L. C. T.; McCammon, J. A.; Amaro, R. E. Mechanisms for Benzene Dissociation through the Excited State of T4 Lysozyme L99A Mutant. *Biophys. J.* **2019**, *116* (2), 205-214.
- (87) Zwier, M. C.; Pratt, A. J.; Adelman, J. L.; Kaus, J. W.; Zuckerman, D. M.; Chong, L. T. Efficient Atomistic Simulation of Pathways and Calculation of Rate Constants for a Protein-Peptide Binding Process: Application to the MDM2 Protein and an Intrinsically Disordered p53 Peptide. *J. Phys. Chem. Lett.* **2016**, *7* (17), 3440-3345.
- (88) Ahmad, M.; Gu, W.; Geyer, T.; Helms, V. Adhesive water networks facilitate binding of protein interfaces. *Nat. Commun.* **2011**, *2*, 261, Article.
- (89) Abrams, C.; Bussi, G. Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy-Switz* **2014**, *16* (1), 163-199.
- (90) Miao, Y.; McCammon, J. A. Unconstrained Enhanced Sampling for Free Energy Calculations of Biomolecules: A Review. *Mol. Simul.* **2016**, *42* (13), 1046-1055.

TOC



**Ligand Gaussian accelerated Molecular Dynamics 3 (LiGaMD3): Improved Calculations of
Binding Thermodynamics and Kinetics of Both Small Molecules and Flexible Peptides**

Jinan Wang and Yinglong Miao*

Computational Medicine Program and Department of Pharmacology, University
of North Carolina – Chapel Hill, Chapel Hill, North Carolina, USA 27599

*To whom correspondence should be addressed: Yinglong_Miao@med.unc.edu

Table S1 The ligand bound and unbound time periods (τ_B and τ_U) recorded from LiGaMD3 simulations of the ligand/peptide-MDM2 binding systems.

System	Method	ID	τ_B (ns)	τ_U (ns)
MDM2-Nutlin	LiGaMD	Sim1	74,321,19,172,146	42.4,395,56,97,36,632
		Sim2	187,1026,209	87,317,184
		Sim3	397,334,706	26,66,67,404
MDM2-Nutlin	LiGaMD3	Sim1	14,55.7,23,83,23,115	25,240.9,196,600,623
		Sim2	61.6,53,87,12,65,47,8	272,137,338,20,219,137,164
		Sim3	14,30,30,24,23	189,628,551,119,392
MDM2-PMI	LiGaMD3	Sim1	360.7,191.6,64.29,10.5,32.9,17.1	85.2,17.1,26.3,40.5,43.3,101.5
		Sim2	456.4,47.2,18.4	261.1,110.6,96.3
		Sim3	68.17,147.0,57.6,10.5,9.8,45.0,8.9	129.8,35.4,56.6,45.4,108.2,59.2,221.0
MDM2-P53	LiGaMD3	Sim1	60,17,29,25,30	250,136,249,153.7,534,380
		Sim2	379,41,30,17	306,201,390,197,434
		Sim3	38.5,23.7,88,24,84	100.7,115,118,207,407,787

Table S2 Energy barriers of ligand/peptide-MDM3 dissociation (“off”) and binding (“on”) calculated from the reweighed (ΔF) and modified (no reweighting, ΔF^*) free energy profiles, curvatures of the reweighed (w) and modified (w^*) free energy profiles near the ligand/peptide Bound (“B”), Barrier (“Br”) and Unbound (“U”) states, and the ratio of apparent diffusion coefficients calculated from the LiGaMD3 simulations without reweighting (modified, D^*) and with reweighting (D).

Sim		ΔF (kcal/mol)		ΔF^* (kcal/mol)		w			w^*			D^*/D	
		Off	On	Off	On	B	Br	U	B	Br	U	Off	On
MDM2-Nutlin	LiGaMD	9.10± 1.26	2.77 ±1.76	2.68 ±0.25	0.63 ±0.16	1.54 ±0.36	0.30 ±0.30	0.22 ±0.17	0.78 ±0.028	0.069 ±0.009	0.020 ±0.005	1.43 ±0.44	0.16 ±0.03
	LiGaMD3	9.65 ±0.76	0.79 ±0.10	0.76 ±0.18	0.33 ±0.15	10.28 ± 0.60	0.12 ±0.07	0.012 ±0.006	1.16 ±0.27	0.14 ±0.05	0.032 ±0.009	0.96 ±0.81	0.18 ±0.05
MDM2-PMI	LiGaMD3	9.05± 0.23	0.89 ±0.07	1.38 ±0.12	0.23 ±0.06	0.94±0 .53	0.0097 ±0.0043	0.048 ±0.020	7.21 ±0.30	0.066 ±0.010	0.025 ±0.003	0.13 ±0.03	0.17 ±0.004
MDM2-P53	LiGaMD3	7.23± 0.40	0.67 ±0.18	1.16 ±0.04	0.093 ±0.031	0.43 ±0.20	0.023 ±0.012	0.023 ±0.018	7.04 ±0.10	0.062 ±0.018	0.026 ±0.002	0.25 ±0.08	0.23 ±0.20

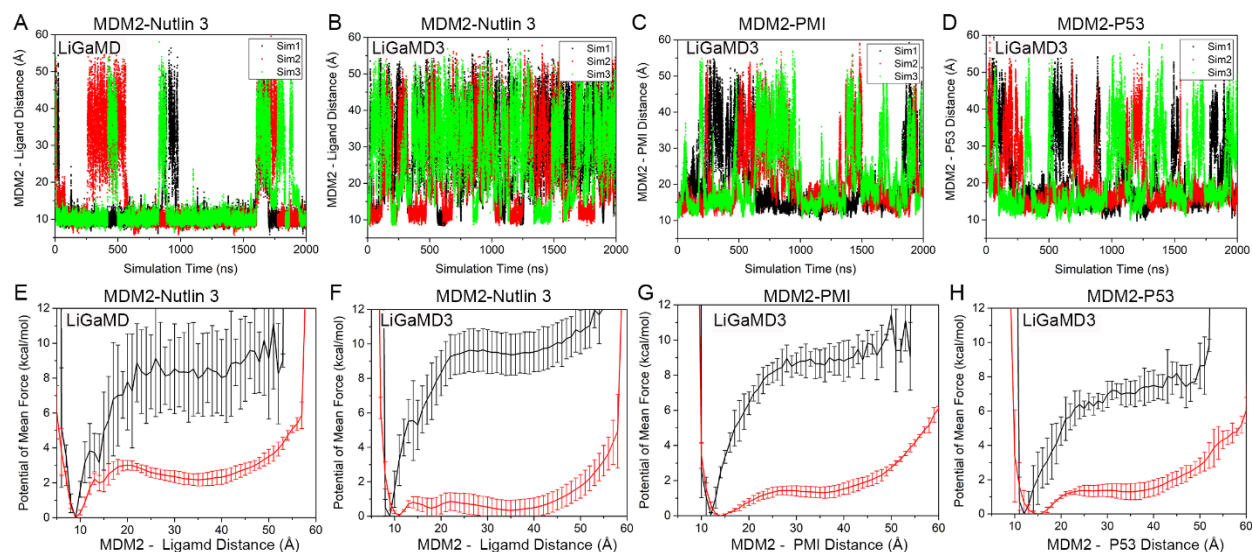


Figure S1. LiGaMD3 simulations captured repetitive dissociation and binding of the Nutlin 3 small molecule and highly flexible PMI and P53 peptides to the MDM2 protein: (A-B) time courses of the distance between the MDM2 and small molecule Nutlin 3 calculated from three independent 2 μ s LiGaMD (A) and LiGaMD3 (B) simulations; (C-D) time courses of the distance between the MDM2 and peptide from three independent 2 μ s LiGaMD3 simulations of (C) PMI and (D) P53 binding to MDM2; (E-F) The corresponding reweighted (black) and non-reweighted (red) PMF profiles of the distance between MDM2 and ligand averaged over three LiGaMD (E) and LiGaMD3 (F) simulations of Nutlin 3 binding to MDM2; (G-H) The corresponding PMF profiles of the MDM2-peptide distances averaged over three LiGaMD3 simulations of (G) PMI and (H) P53 binding to MDM2. Error bars are standard deviations of the free energy values calculated from three LiGaMD and LiGaMD3 simulations.

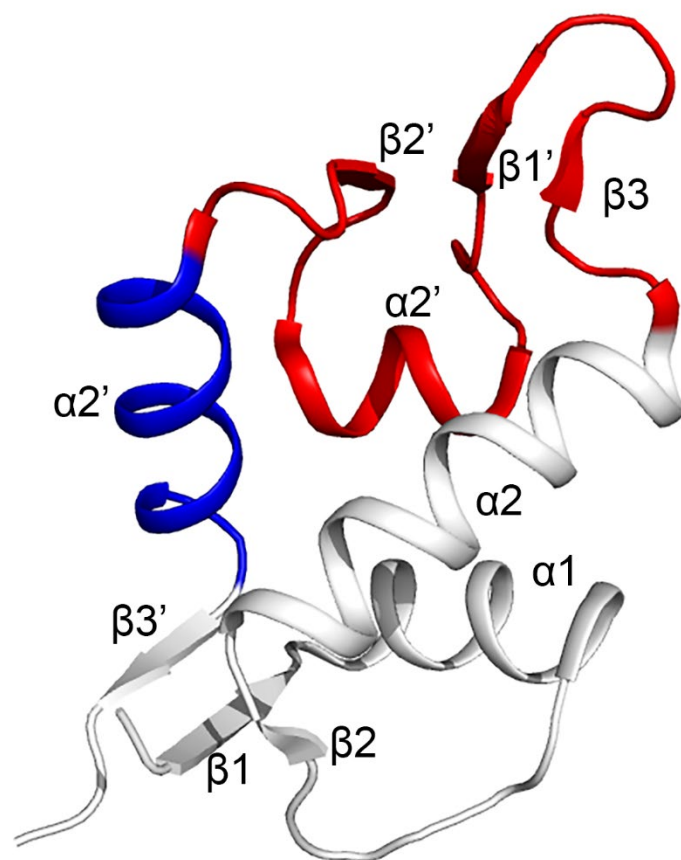


Figure S2. Cartoon representation of the MDM2 with each motif labeled. Residues predominantly involved in “pathway 1” were colored in red, including the $\beta 3$, $\beta 1'$ and $\beta 2'$ strands and the $\alpha 1'$ helix (residues 65-95). Residues mainly involved in “pathway 2” were colored in blue, including the $\alpha 2'$ helix (residues 97-106).