Article

Metadynamics-Based Approaches for Modeling the Hypoxia-Inducible Factor 2α Ligand Binding Process

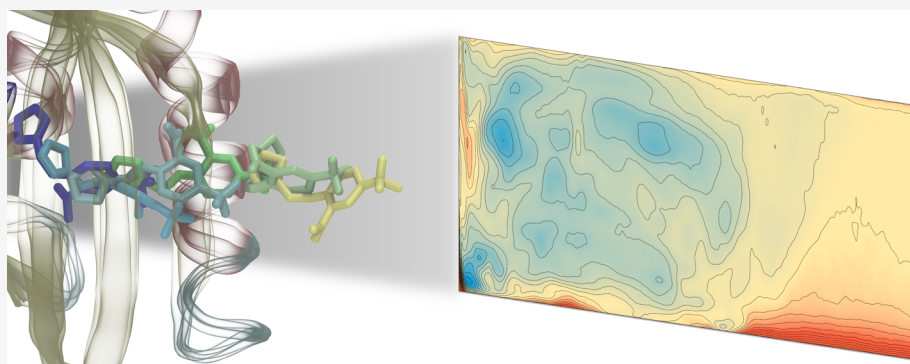Lara Callea, Laura Bonati, and Stefano Motta*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Several methods based on enhanced-sampling molecular dynamics have been proposed for studying ligand binding processes. Here, we developed a protocol that combines the advantages of steered molecular dynamics (SMD) and metadynamics. While SMD is proposed for investigating possible unbinding pathways of the ligand and identifying the preferred one, metadynamics, with the path collective variable (PCV) formalism, is suggested to explore the binding processes along the pathway defined on the basis of SMD, by using only two CVs. We applied our approach to the study of binding of two known ligands to the hypoxia-inducible factor 2α, where the buried binding cavity makes simulation of the process a challenging task. Our approach allowed identification of the preferred entrance pathway for each ligand, highlighted the features of the bound and intermediate states in the free-energy surface, and provided a binding affinity scale in agreement with experimental data. Therefore, it seems to be a suitable tool for elucidating ligand binding processes of similar complex systems.

■ INTRODUCTION

Understanding the thermodynamic principles behind the mechanism of ligand-protein binding is very important for the development of a successful drug design campaign. Experimental techniques are able to estimate binding thermodynamic and kinetic properties but cannot provide the atomistic insight that forms the basis of rational approaches to drug design. In this context, *in silico* methods are becoming increasingly effective in complementing experiments and providing atomic-level descriptions of ligand binding. Docking methods are widely used to rapidly screen libraries containing thousands or even millions of compounds[1] but suffer from several limitations, one above all the lack of protein flexibility. In cases where the induced-fit effects of the ligand are important, a number of alternative methods have been proposed to account for protein flexibility in ligand binding.[2,3] For example, flexibility can be introduced by docking ligands to an ensemble of different protein conformations (ensemble docking).[4–7] This method only partially introduces protein flexibility, and results strongly depend on the type of ensemble used. Recently, molecular dynamics (MD) simulations have been used to study processes happening on timescales that range from nanoseconds to milliseconds and beyond,[8] making them attractive for the study of ligand binding. However, computation of key thermodynamic quantities requires the observation of multiple binding events to obtain reliable statistics on the process, thus increasing the computation time. Typically, enhanced-sampling techniques are used to speed up the simulation of the binding/unbinding events.[9,10] Most of these techniques make use of a bias potential that forces the system to sample higher-energy regions, speeding up the crossing of energy barriers.

Among the methods for studying ligand binding based on enhanced-sampling MD,[11–17] in this work, we focused on steered MD[18] (SMD) and metadynamics[19] (MetaD).

SMD was inspired by single-molecule pulling experiments and applies a moving restraint bias that pulls the system along a selected variable. Despite its wide applications to the study of (un)folding mechanisms of proteins[20,21] and transportation of ions and other molecules across membrane channels,[22,23] SMD has also emerged as a method for studying ligand (un)-binding,[24−27] given that it is particularly well-designed for the investigation of entry and exit pathways. Its points of strength are the easy setup and the shortness of simulations.[28] On the other hand, SMD still suffers from several limitations, in particular regarding the calculation of the potential of the mean force (PMF).[28] During the pulling, indeed, a part of the work is spent as dissipative work, and convergence can be difficult to reach. In theory, the Jarzynski equality may account for the dissipative part of the work; however, when the range of work obtained in multiple replicas is broad, simulations with the lowest work contribute the most to the calculation of the average work.[29] These limitations may be overcome by performing a large number of replicas and reducing the pulling speed, but for some complex systems, this is often not enough.

MetaD is a method based on the introduction of a history-dependent bias potential applied to a small number of suitably chosen collective variables (CVs).[19,30,31] Within the CV subspace, the potential is built up by adding Gaussians along the sampled trajectory to discourage the system from revisiting already sampled configurations. The Gaussian height, the Gaussian width, and the deposition time are crucial parameters to obtain a converged free-energy surface (FES).[30] However, the choice of the CVs is the most critical aspect in MetaD, and results can be seriously affected by the omission of important degrees of freedom (hysteresis).[30] Given that the computational cost to reconstruct the free-energy surface exponentially grows with the number of CVs, Branduardi et al.[32] developed the path collective variable (PCV) method, which allows exploration of complex multidimensional processes along a predefined pathway described by a single CV. An additional CV, which describes the distance from the reference path, usually completes the set of CVs necessary to efficiently sample the process of interest. In the past few years, MetaD in its various forms has been successfully applied to several ligand-protein binding studies.[33−41] One of the main advancements in this field was the development of the so-called "funnel metadynamics" where a funnel-shaped potential limits the space available to the ligand once it has undocked.[42,43]

Here, we investigate the ligand binding process to the hypoxia-inducible factor 2α (HIF-2α), a pharmaceutically relevant system widely recognized as a target for cancer therapy.[44] HIF-2α mediates the physiological responses to hypoxia through heterodimerization with the aryl hydrocarbon receptor nuclear translocator (ARNT).[45,46] Both the HIF-2α and ARNT belong to the mammalian basic helix−loop−helix-PER-ARNT-SIM (bHLH-PAS) family of proteins, where members modulate transcriptional responses to environmental and cellular signals and are involved in a variety of physiological processes and diseases in humans.[44,47] Members of the bHLH-PAS family present an N-terminal bHLH region for DNA binding, two PAS domains (PAS-A and PAS-B) with the role of both sensing external signals and recognizing the dimerization partner and a transactivation domain. For a long time, only another bHLH-PAS protein, the aryl hydrocarbon receptor (AhR), was known to be activated by binding to a wide range of ligands within its PAS-B cavity.[48,49] More recently, following the discovery of a buried cavity within the

HIF-2α PAS-B domain,[50] several artificial small molecules were identified as HIF-2α ligands and potential inhibitors of the HIF-2α:ARNT dimerization.[51−56] The structural determination of the HIF-2α:ARNT dimer encompassing the whole bHLH-PAS region, in the unbound, DNA-bound, and inhibitor-bound forms,[45] recently allowed us to investigate the inhibition mechanism of the 0X3 antagonist and to shed light on pharmacophoric features required for the development of new inhibitors.[57]

In this work, we combined SMD and PCV MetaD simulations to investigate the binding process of two known ligands to the HIF-2α PAS-B domain. We were aimed at both investigating the ligand entrance pathway into the binding cavity and assessing the validity of the selected methods for such a complex system. In fact, the buried nature of the cavity makes it difficult to imagine the entry or exit route of the ligand, and although a previous MD investigation identified probable pathways for water exchange with the bulk solvent,[53] access of larger organic molecules to the cavity has never been studied. Moreover, it is conceivable that ligand entrance into this cavity may involve significant protein conformational rearrangements. The above features of the system make simulation of the ligand binding process a nontrivial task and required the development of specific methodological approaches. In light of the obtained results, these methods appear to be suitable also for the elucidation of other ligand binding processes with similar characteristics.

■ METHODS

System Preparation and Molecular Dynamics Simulation. Crystal structures of HIF-2α in its bound state with the THS-020 ligand (PDB ID: 3H82[53]) were obtained from the Protein Data Bank (PDB).[58] The PAS-B of the ARNT protein partner, included in the X-ray deposition, was removed. This does not induce perturbations in the structure of the HIF-2α PAS-B, as shown by the RMSD plot (Figure S1) that highlights the stability of the HIF-2α domain during the MD simulation. The KG-721 bound form was obtained with molecular docking calculations (see the next subsection). The protein was prepared with the Protein Preparation Wizard[59] included in Maestro: hydrogen atoms were added, all water molecules were removed, C- and N-terminal cappings were added, disulfide bonds were assigned, and residue protonation states were determined by PROPKA[60] at pH = 7.0. The ligands were prepared using the LigPrep[61] tool included in Maestro in order to optimize the structures. The partial charges of ligands were calculated using the RESP[62] method at the AM1-BCC[63] level of theory in Antechamber,[64] while a GAFF[65] parametrization was used to achieve the complete topological description of each ligand. The unbiased MD simulations were performed using GROMACS 2018.6.[66] The protein was solvated in an orthorhombic box with TIP3P[67] water molecules and neutralized with Na$^+$/Cl$^-$ ions. The minimal distance between the protein and the box boundaries was set to 20 Å. The Amber ff14SB force field[68] was used for the protein, and a multistage equilibration protocol was applied: the system was first subjected to 2000 steps of the steepest descent energy minimization, with positional restraints (239 kcal mol^{-1} nm^{-2}) for the backbone and the ligand. Subsequently, a 200 ps NVT MD simulation was used to heat the system from 0 to 100 K with restraints lowered to 96 kcal mol^{-1} nm^{-2}; then, the system was heated up to 300 K in 400 ps during an NPT simulation with further lowered restraints (48 kcal mol^{-1}

nm^{-2}). Finally, the system was equilibrated during an NPT simulation for 2 ns with backbone restraints lowered to 12 kcal mol^{-1} nm^{-2}. In the NVT simulations, temperature was controlled using the Berendsen thermostat[69] with a coupling constant of 0.2 ps, while in the NPT simulations, the V-rescale thermostat[70] (coupling constant of 0.1 ps) was used and the pressure was set to 1 bar with the Parrinello−Rahman barostat[71] (coupling constant of 2 ps). A time step of 2.0 fs was used, together with the LINCS[72] algorithm to constrain all the bonds. The particle mesh Ewald method[73] was used to treat the long-range electrostatic interactions with the cutoff distance set at 11 Å. Short-range repulsive and attractive dispersion interactions were simultaneously described by a Lennard-Jones potential, with a cutoff at 11 Å. Finally, a 20 ns production run was performed without the constraints.

Molecular Docking of the KG-721 Ligand. Conformational analysis of the ligand structure was performed using Macromodel[74] with the OPLS_2005[75] force field. The obtained global minimum was used as the starting point for molecular docking calculations using Glide[76] XP[77] (Extra Precision). In particular, Glide uses a flexible ligand-rigid protein approach, in which a series of hierarchical filters are applied to find the possible positions and conformations of the ligand in the binding cavity (poses). The properties of the protein are represented on a grid that provides gradually more accurate scores. The initial screenings are deterministically performed over the complete phase space of the ligand to identify the most promising poses. From the selected poses, the ligand is then refined in the torsional space in the receptor field. To take into account the flexibility of the protein, the ensemble-docking approach was used, which involves ligand docking to multiple receptor conformations. These can be derived either experimentally or computationally (e.g., by MD simulations).[5] The conformational ensemble here selected consisted of the crystallographic structures of the HIF-2α PAS-B in complex with artificial ligands available in PDB (3F1O,[50] 3H82,[53] 3H7W,[53] 4GS9,[51] and 4GHI[52]). The results showed that the best XP score is the one related to the KG-721 ligand in the 4GHI structure.

Steered MD Simulations (SMD). In SMD simulations,[26,27,78] a time-dependent external force is applied to the ligand to aid its unbinding from the protein. In particular, the transition between the bound and the unbound states is achieved by adding a harmonic time-dependent potential, acting on a descriptor (or collective variable), to the standard Hamiltonian. During the transition, the external work performed on the system can be calculated using the Jarzynski equation.[79] All the SMD simulations were performed using the PLUMED 2.4.6[80,81] plugin integrated in GROMACS 2018.6.[66] We chose the ligand-protein distance as the pulling variable. This was defined as the distance between the center of mass of selected atoms at the bottom of the binding cavity (different for the two pathways, see Figure S2) and the center of mass of the ligand heavy atoms. The spring constant was set to the value of 10.0 kcal/mol·Å^2, and the ligand was pulled from the initial value of CV to 35 Å in 25 ns with a resulting pulling velocity of 0.984 Å/ns. We ran 50 independent replicas, and the time length for each simulation was 25 ns, which ensured the achievement of a complete solvation of the ligand in the unbound state. The starting point of each replica was derived from an ensemble of states extrapolated at regular time intervals of 0.2 ns from the last 10 ns of the unbiased simulation.

Metadynamics (MetaD) and Path Collective Variables (PCVs). The central idea of the metadynamics method[19,30] is to bias the system along a set of CVs using a history-dependent potential. To achieve this, a Gaussian-shaped potential is added to bias the system at the current position of the CVs, at regular time intervals. This allows the system to escape from any local minimum and to visit new regions in the CVs space. In metadynamics, to push the system to visit even high free-energy regions, the Gaussian-shaped potential has constant height. On the contrary, in the well-tempered metadynamics[82] approach, used in this work, the height of the Gaussian is decreased with the amount of bias already deposited according to

$$w = w_0 e^{-\left[\frac{V(s,t)}{\Delta T}\right]} \tau_G$$

where w_0 is an initial Gaussian height, ΔT an input parameter with the dimension of a temperature, k_B is the Boltzmann constant, and τ_G is the time interval at which Gaussians are deposited.[82]

The path CV formalism[32,83] has been widely used to investigate biological processes, to compute their free-energy surfaces, and to characterize their kinetic behavior.[39,34] In this work, PCVs were used to study the transition between the bound and the unbound states in the unbinding process of some HIF2-α ligands. We described the transition pathway with a set of frames derived from the SMD simulations: 12 frames were used for the THS-020 ligand and 11 frames for the KG-721 ligand. For the first part of the path, the frames (from 1 to 7 for THS-020 and from 1 to 6 for KG-721) were obtained from the SMD simulations with the lowest value of external work performed on the system. Frames were selected to be equally spaced (2 Å). For the second part of the path, the frames were obtained with linear interpolation (see the Supporting Information, Supplementary Text for the details). Following the procedure proposed by Branduardi et al.,[32] we introduced two collective variables: $s(R)$, the progress along the reference path, and $z(R)$, the distance orthogonal to the reference path. The λ value was set to 33.0 nm^2. The distance between the instantaneous conformational state during the simulation and the reference coordinates in the path was evaluated by the RMSD metric.[84] In particular, the RMSD along the entry/exit pathway was calculated between a selection of protein atoms and all the ligand heavy atoms (see Figure S3). In all simulations, the Gaussian-shaped potentials were deposited every 500 simulation steps, the initial height was set to 1 kJ/mol, and the decay corresponding to a bias factor of 10 was chosen. The Gaussian widths (σ) for the $s(R)$ and $z(R)$ variables were set to 0.05 and 0.007, respectively. Widths were set so that they are about 1/3 of the CV standard deviations observed in the unbiased MD simulation. The two variables, $s(R)$ and $z(R)$, were constrained to be less than 12 and 0.2 nm^2, respectively.

Extraction of Minima and Cluster Analysis. To characterize the different minima identified on the final free-energy surface (FES), we extracted a group of frames belonging to each minimum hole. To obtain a representative structure of the complex in each minimum, a cluster analysis on the metadynamics trajectory frames with a stride of 10 ps was performed. The GROMOS[85] clustering algorithm was applied, with a 2 Å RMSD cutoff on the heavy atoms of the ligands. The centroid of the most populated cluster was then defined as the representative structure in that minimum.

■ RESULTS AND DISCUSSION

The analysis of the unbinding pathways was performed for two of the HIF-2α ligands identified in the study of Key et al.[53] The THS-020 ligand (Figure 1A) has a good binding affinity
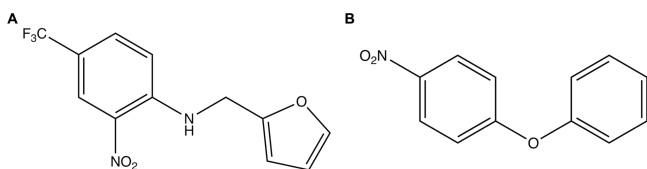

Figure 1. 2D structure of THS-020 (A) and KG-721 (B).

for the protein (ΔG_{exp} = −7.9 ± 0.5 kcal/mol), and the ligand-protein bound structure, determined by X-ray crystallography,[53] is available. The KG-721 ligand (Figure 1B) is a lower-affinity ligand (ΔG_{exp} = −6.9 ± 0.1 kcal/mol) identified in the same work.[53] We choose it among the other HIF-2α ligands,[50−56,86−88] not only for the different binding affinity but also to deal with a molecule not congeneric to THS-020,[53,51] with different physicochemical properties and with lower size. Moreover, for this ligand, no experimental structures of the ligand-protein complex are available, thus offering us the opportunity to study a system where the starting conformation, obtained by docking, could not take into account the induced-fit effects on the protein.

In the following two subsections, we present the application of a specific SMD MetaD protocol on the THS-020. We first used SMD to identify the unbinding pathway, and then, we applied PCV MetaD simulations to characterize the relevant states in the binding/unbinding process and to obtain a reliable estimate of the binding affinity. In the third subsection, we present the results obtained with the same protocol for KG-721.

Identification of Unbinding Pathways for the THS-020 Ligand by the SMD Method. Starting from the X-ray structure for the HIF-2α PAS-B domain in complex with THS-020, the possible ligand unbinding pathways were investigated using the SMD approach. Steered molecular dynamics is a popular method for studying ligand-protein unbinding events[89−91] and can provide both a qualitative description of the pathways and a quantitative estimate of the free-energy difference between the bound and unbound states.

To calculate the free-energy difference using the Jarzynski equality, it is necessary to have a large number of SMD replicas. To this aim, a 20 ns unbiased MD simulation was performed starting from the X-ray structure, generating an ensemble of 50 slightly different states of the complex (Figure S4a), extracted from the last 10 ns. These were then used as starting points for the 50 SMD simulations. The structural convergence of the unbiased simulation was assessed by calculating the RMSD matrix on the protein Cα atoms and on the ligand heavy atoms (Figure S4b). Each SMD replica was 25 ns long. To verify if the values chosen for the SMD simulation parameters (see the Methods section) are appropriate, the RMSD plot on Cα atoms (Figure S5) and the secondary structure conservation graphs (Figure S6) were calculated for the 50 replicas along path 1. As shown in Figure S6, no significant distortions of the protein structure (except for a slight deformation of the Fα helix upon ligand unbinding) were observed during simulations, thus confirming the validity of the proposed protocol.

Other authors identified two entry/exit pathways for solvent water by MD simulations of the apo HIF-2α PAS-B[53]: path 1 gets through the Fα helix and the Gβ strand, while path 2 gets through the Fα helix, the short Eα helix, and the AB loop (Figure 2). On this basis, for each of these two pathways, we
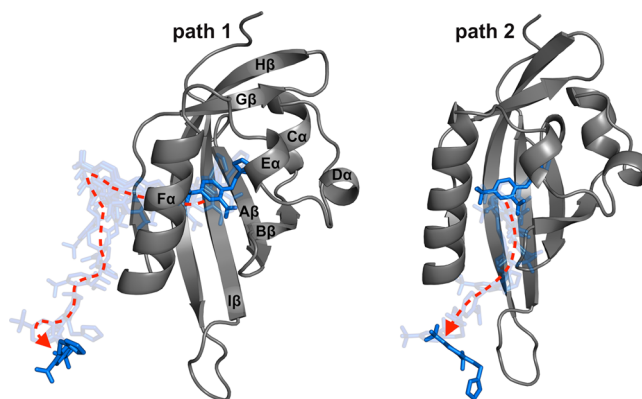

Figure 2. THS-020 unbinding pathways. In pathway 1 (left), the ligand passes through Fα and Gβ while in pathway 2 (right) through Fα, Eα, and the AB loop. The starting protein structure is represented as gray cartoons, the ligand conformations in the first and last frames of the trajectory as blue sticks, and the conformations of the ligand in the intermediate frames as transparent sticks.

calculated a CV allowing us to pull the ligand out of the binding cavity, by selecting an appropriate set of residues at the bottom of the binding cavity (see the Methods section).

The work profiles resulting from the 50 replicas, in Figure S7, show an increase in the work value during the initial part of the unbinding process followed by relatively settled work values, indicating the absence of interaction between the ligand and the protein.

All the curves show a similar profile, but a qualitative comparison of the total work reveals that less work is required for unbinding following pathway 1. Moreover, a broader range of values is observed for the replicas following pathway 2, indicating that higher barriers can occur in some of the replicas associated to this pathway. For a quantitative analysis, the minimum and maximum work values (W_{min} and W_{max}), the minimum value of the maximal force (F_{max}) among the replicas, and the free-energy difference between the unbound and the bound states (ΔF_{unbind}) were extracted for each pathway (Table 1). The results show that the W_{min} necessary

Table 1. Results of SMD Simulations for the THS-020 Ligand

pathway	W_{min} (kcal/mol)	W_{max} (kcal/mol)	F_{max} (pN)	ΔF_{unbind} (kcal/mol)	st. dev.
1	28.19	55.17	1013.77	30.55	16.80
2	38.16	66.42	1089.36	40.25	13.10

to pull the ligand outside the cavity along path 1 is about 10 kcal/mol less than that required for path 2; a similar trend is observed for the values of W_{max} and ΔF_{unbind}. A difference of 75.59 pN between the F_{max} values in the two paths is observed, which confirms a clear preference of path 1 over path 2.

Therefore, steered MD allowed us to compare the two unbinding pathways of THS-020 and to select the preferred one by identifying a higher-energy barrier along pathway 2. However, SMD provided a value for the ΔF_{unbind}, estimated by
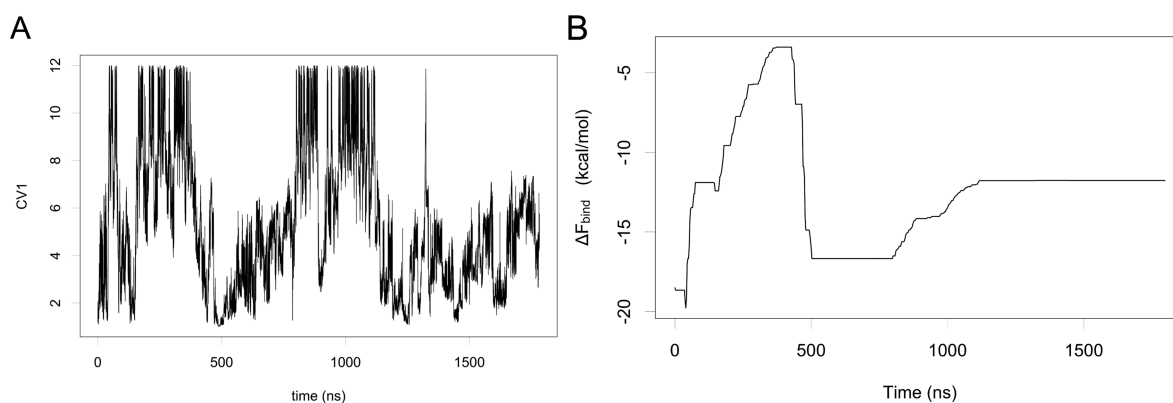
Figure 3. (A) Plot of the CV1 ($s(R)$) against simulation time during THS-020 MetaD simulation: the lowest values of $s(R)$ correspond to the bound state while the highest to the unbound ones; (B) one-dimensional projection of the binding free-energy values associated to path 1 during the metadynamics simulation.
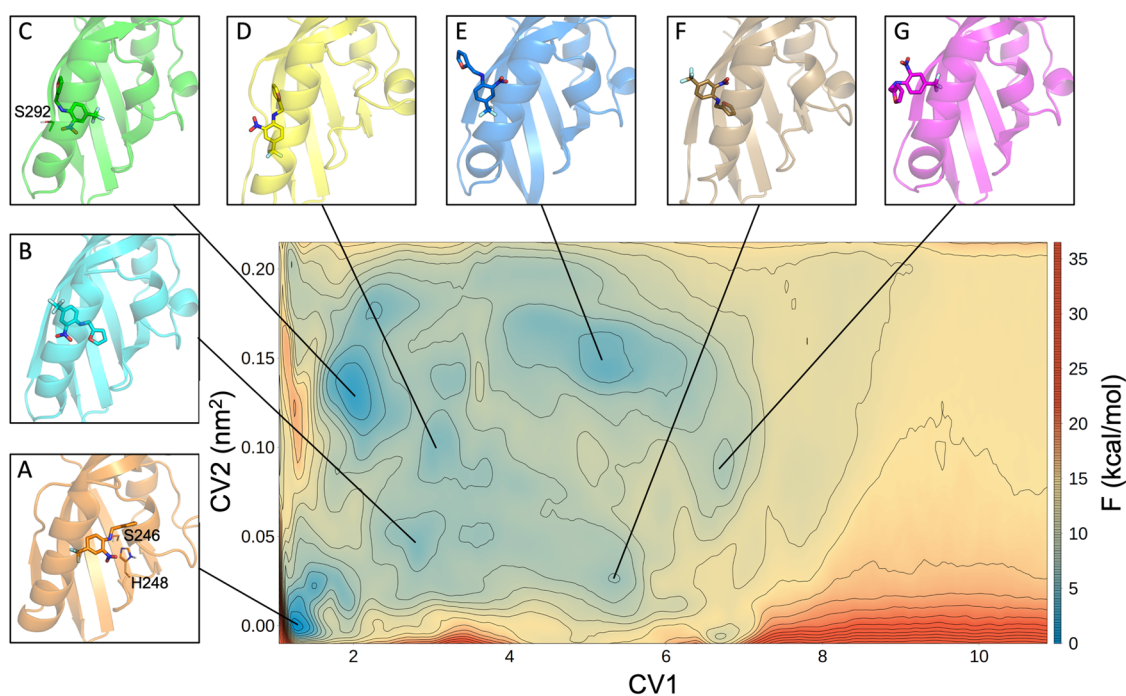

Figure 4. Free-energy surface obtained from the PCV approach for the binding/unbinding of the THS-020 ligand. The isolines are drawn using a 1.5 kcal/mol spacing. The 3D structures of the centroids of the main minima are reported with different colors: the protein is represented as cartoons and the ligand as sticks. The black lines indicate the corresponding minima in the FES.

means of the Jarzynski equality, that was about 4 times higher than the experimental value. It is indeed known that these nonequilibrium simulations generally undersample the relevant protein-ligand states across the unbinding pathway, leading to errors in the computed binding free energy.[92] Moreover, replicas with less work done on the system have an enormous weight compared to all the other trajectories, which makes the method extremely sensitive to insufficient sampling.[93] For this reason, we then applied the PCV MetaD approach that was recently proposed as a valuable method for computing absolute binding free energies in ligand binding.[34,39,35]

Metadynamics Simulations and Free-Energy Profiles with PCVs for THS-020. For a detailed mechanistic interpretation of the ligand binding/unbinding process, we used well-tempered metadynamics simulations with the PCV approach[83,35,94] (see the Methods section). This allowed us to characterize the relevant states along the preferred path

obtained with SMD simulations (the one with lower values of total work obtained from the SMD simulation, path 1), as well as to estimate the binding free-energy value. The key points of this method are the choice of appropriate CVs and the construction of a set of equally spaced frames along the CVs in terms of RMSD between adjacent snapshots. This frameset represents a reference path for investigating the process. As CVs, we used the progress along the path, $s(R)$, and the distance orthogonal to the reference path, $z(R)$. We want to underline the importance of the path construction phase, especially in a case with a buried binding site like the one presented in this work. Here, we decided to include both ligand and protein atoms in the frameset that represents the path to better describe the protein conformational changes during the process (mouth opening through side-chain conformational changes and small backbone adjustment). Only protein atoms involved in the conformational changes,

highlighted by SMD simulations, were included. Moreover, a hybrid approach that combines frames from SMD simulations and linear interpolation was used for the inner and outer parts of the path, respectively. Details about the construction of the reference path can be found in the Supporting Information, Supplementary Text. The reference path obtained with this approach is represented in Figure S8. The RMSD matrix of the frameset (Figure S8) is a symmetric matrix with a typical gull wing shape, indicating that the frames are correctly equally spaced.

We collected a total of 1.8 μs of metadynamics simulation in which we observed several binding/unbinding events, as shown in Figure 3A. The binding free energy (ΔF_{bind}), calculated as the free-energy difference between the deepest minima in the bound state and the flat plateau in the unbound state, turns out to be equal to -11.8 kcal/mol. The free-energy profile during the simulation, shown in Figure 3B, indicates that the simulation reaches a constant value of ΔF_{bind} after about 1200 ns. The convergence was also monitored by plotting the hill heights as a function of the simulation time (Figure S9).

The free-energy surface (FES) for the binding/unbinding process, as a function of CV1 ($s(R)$) and CV2 ($z(R)$) in 1.8 μs of metadynamics simulation, is shown in Figure 4 together with the relevant minima found along the pathway (labeled as A to G). The coordinates and the binding free-energy values of each minimum are reported in Table S1. A cluster analysis of the conformations belonging to each minimum hole was then performed (see the Methods section for details). In each minimum, the first cluster is the most populated (Figure S10), and its centroid was used as the representative structure for that minimum. Looking at the FES (Figure 4), three different regions can be identified following CV1: the bound state, with $s(R)$ values between 1 and 3 (minima A−C), the intermediate states, with $s(R)$ values between 3 and 7 (minima D−G), and the unbound state, with $s(R)$ values from 7 onward. In the deepest minimum (A), the ligand is oriented in the same way as in the X-ray structure (ligand RMSD = 1.36 Å). This geometry is stabilized by a hydrogen bond between the NH group of the ligand and the H248 residue as well as by a transient hydrogen bond between the oxygen atom of furan and the S246 residue.

Moving up to higher CV2 values, alternative binding geometries can be detected. In minimum B, the ligand is shifted toward the exit of the cavity, and breaking of the hydrogen bonds that stabilize minimum A causes a lower stability. Instead, in minimum C, the ligand is located on the mouth of the binding cavity and is even turned 180°, with the CF$_3$ group oriented toward the bottom of the cavity. Also, in this minimum, the amino group of the ligand forms a hydrogen bond with the S292 residue. This last conformation represents the second minimum in energy.

Unbinding Pathway for the KG-721 Ligand. Following the encouraging results on THS-020, for which the selected methods were able to identify the experimental binding geometry as the most stable among all the possible bound states, we extended the study to the lower-affinity KG-721 ligand. Given the lack of an experimental structure, we obtained the starting geometry for our calculations by molecular docking, using the ensemble-docking technique[5] (details are reported in the Methods section). This technique has led to improvement of the description of ligand-induced protein conformational changes in many systems,[4−7] but it

may be not sufficient in some particularly challenging cases. These include docking studies of ligands different from the ones cocrystallized in the protein structures of the ensemble (if any), like in our case.

As for the THS-020 ligand, 50 independent replicas of SMD simulations (each of 25 ns) were performed for the two possible pathways. The resulting work profiles (Figure S11) are similar to those obtained for THS-020. However, different from that ligand, they show a similar range of work values for path 1 and path 2 and do not suggest any preference for one path over the other. This is also confirmed by negligible differences between the values of W_{min}, W_{max}, and ΔF_{unbind} (Table 2) in the two paths. This result suggests that for a small

Table 2. Results of SMD Simulations for the KG-721 Ligand

pathway	W_{min} (kcal/mol)	W_{max} (kcal/mol)	F_{max} (pN)	ΔF_{unbind} (kcal/mol)	st. dev.
1	28,46	54,45	851,21	30,55	13,04
2	27,47	59,13	956,29	27,12	16,60

ligand, the two pathways may have a similar probability. This hypothesis is consistent with the findings of Key et al.,[53] who observed a similar percentage of transferring of the small water molecules in the two paths. However, the observed difference of 105.08 pN between the F_{max} values of the two paths of KG-721 suggests that a higher barrier for unbinding exists along path 2, similarly to what was observed for the THS-020 ligand.

Based on the results obtained with the steered MD simulations, 11 frames along path 1 were used to build the reference path for metadynamics simulations. The resulting RMSD matrix of the frameset and a representation of the reference path are shown in Figure S12.

After 3 μs of metadynamics simulation, we reconstructed the free-energy profile (Figure S13a). Starting from 2200 ns, the free-energy difference between the bound and the unbound states fluctuates around a value of -8.0 kcal/mol with a variation of ± 1 kcal/mol. The calculated binding free energy revealed that the KG-721 ligand has a lower binding affinity than THS-020 (-11.8 kcal/mol), in agreement with the experimental data: ΔG_{exp} (KG-721) = -6.9 ± 0.1 kcal/mol and ΔG_{exp} (THS-020) = -7.9 ± 0.5 kcal/mol.

During the simulation, we observed multiple binding and unbinding events, and the hill heights decrease toward 0 (Figure S13b,c).

The final FES obtained for this system is shown in Figure 5. Even in this case, we identified several minima (details in Table S1), and we used the centroid of the most populated cluster (details on cluster compositions in Figure S14) in each minimum, as the representative structure of that minimum.

Again, following the CV1, three regions can be distinguished: the bound state, between 1 and 3, the intermediate states, between 3 and 7, and the unbound state, from 7 and on. The region around the bound state displays a multiplicity of alternative binding geometries and does not allow us to distinguish a favorite bound minimum. Minima from A to D can be associated to alternative bound states in which the ligand rotates within the binding cavity. In particular, in minimum A, the ligand is oriented with NO$_2$ toward the most polar part of the cavity (S292, S304, and Y307 residues, at the entrance of the cavity) and the phenyl ring toward the apolar part (F244, F254, and I261 residues, at the bottom of the cavity), as expected (Figure 6, right panel). Even in minimum
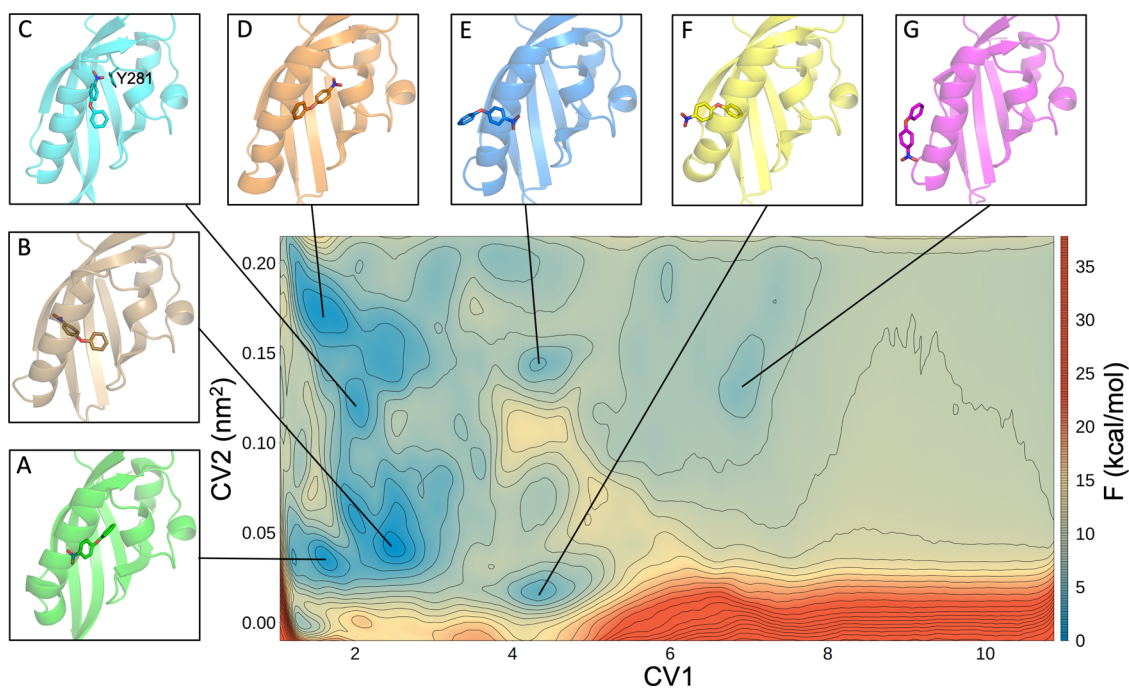
Figure 5. Free-energy surface obtained from the PCV approach for the binding/unbinding of the KG-721 ligand. The isolines are drawn using a 1.5 kcal/mol spacing. The 3D structures of the centroids of the main minima are reported with different colors: the protein is represented as cartoons and the ligand as sticks. The black lines indicate the corresponding minima in the FES.
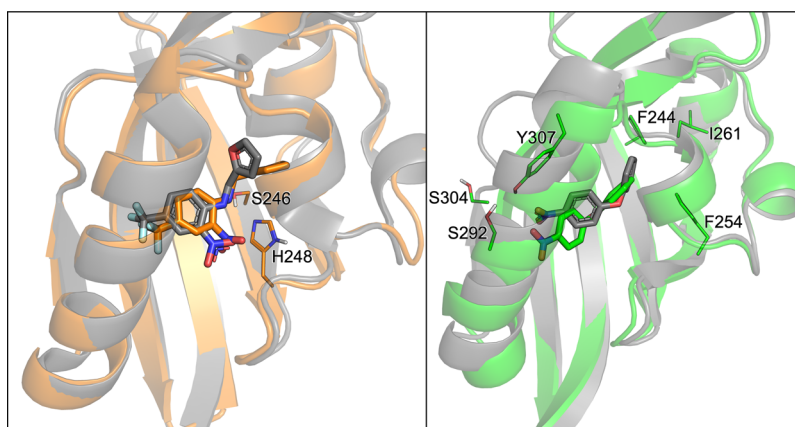

Figure 6. Comparison between minimum A and the starting structure for the two ligands. On the left: for the THS-020 ligand, the overlay of minimum A, in orange, with the crystallographic structure of the complex, in gray. On the right: for the KG-721 ligand, the overlay of minimum A, in green, with the docking pose, in gray.

B, NO_2 is oriented toward the polar region, but the phenyl ring is slightly bent with respect to the other ring. Moving up to higher CV2 values, minima present different orientations of the ligand inside the cavity: minimum C is stabilized by a hydrogen bond between NO_2 and the Y281 residue; in minimum D, the ligand is rotated 180° with respect to minimum A and does not show stable hydrogen bonds with the protein.

While we previously observed that in the deepest minimum (A), THS-020 well overlaps the experimental binding geometry (Figure 6, left panel), minimum A of KG-721 (Figure 6, right panel) is the most similar to the starting docking pose (where values in the CV subspace are $s(R) = 1.3$ and $z(R) = 0$). Indeed, the RMSD between the centroid of minimum A and the docked pose is 2.45 Å, indicating that the two conformations are quite different.

In the case of KG-721, which is not congeneric with any of the cocrystallized HIF-2α ligands,[50,51,53] the ensemble-docking strategy was not sufficient for the correct definition of the binding mode. In light of our results, we underline the importance of including protein flexibility more completely. Our results indicate that MetaD calculations are not influenced by the inaccurate starting conformation of the complex but lead the system to evolve to a more stable conformation. Therefore, this technique appears a promising tool in cases where structural information for congeneric ligands is not available.

■ CONCLUSIONS

Modeling the pathways for ligand binding to the HIF-2α PAS-B domain represents a nontrivial task due to the buried nature of the binding cavity that suggests that significant protein

conformational changes may occur upon ligand access. The computational protocol here proposed effectively combines two promising methods based on enhanced-sampling MD. Steered MD simulations are used to identify the preferred unbinding pathway among alternative ones and to guide the construction of the reference path for the subsequent step. On the other side, metadynamics, with the path collective variable formalism, is used to obtain a more rigorous characterization of the free-energy surface and to calculate the binding free-energy value.

By applying this approach to elucidate the binding process of two different ligands of HIF-2α, we obtained the correct binding affinity scale, according to the experimental data available, and we identified minima in the FES that clearly depict the bound state(s) and the intermediate states characteristic of each ligand. Moreover, the method was effective in leading the system to evolve to the most stable binding conformation, starting either from an X-ray structure of the ligand-protein complex or from a docking pose. Therefore, it appears a promising tool also in cases where reference structural information is lacking.

Given the recent discovery of HIF-2α as a pharmaceutical target for cancer therapy, the proposed computational approach based on enhanced-sampling MD appears to be an invaluable tool to investigate the binding process of different ligands, thus contributing to the development of successful drug design projects. The results obtained here also encourage us to extend applications to other binding mechanisms of bHLH-PAS proteins, including significant targets, such as the AhR, for which no experimental structural information on the ligand-bound states is available.

■ ASSOCIATED CONTENT

ⓈⒾ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00114.

(Supplementary text) Construction of path CVs, (Table S1) coordinates and binding free-energy values of the relevant minima along the CV1 ($s(R)$) and CV2 ($z(R)$), (Figure S1) plot of the RMSD values computed on Cα atoms for the HIF-2α PAS-B domain during the unbiased MD simulation, (Figure S2) representation of the pulling variable for the two paths, (Figure S3) selection of protein atoms for RMSD calculations, (Figure S4) monitoring of the conformational changes during unbiased MD simulation of the HIF-2α PAS-B with the THS-020 ligand, (Figure S5) plots of the RMSD values computed on Cα atoms for the THS-020 SMD replicas along path 1, (Figure S6) DSSP plots for the HIF-2α secondary structure transitions, (Figure S7) work profiles for the two unbinding pathways of THS-020, (Figure S8) resulting reference path for PCVs for the THS-020 ligand, (Figure S9) plot of the hill heights during the metadynamics simulation for the THS-020 ligand, (Figure S10) cluster analysis on each minimum of the THS-020 FES, (Figure S11) KG-721 unbinding pathways, (Figure S12) resulting reference path for PCVs for the KG-721 ligand, (Figure S13) results of KG-721 MetaD simulation, and (Figure S14) cluster analysis on each minimum of the KG-721 FES (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Stefano Motta − *Department of Earth and Environmental Sciences, University of Milano-Bicocca, 20126 Milan, Italy;* ⓞ orcid.org/0000-0002-0812-6834; Email: stefano.motta@unimib.it

Authors

Lara Callea − *Department of Earth and Environmental Sciences, University of Milano-Bicocca, 20126 Milan, Italy*
Laura Bonati − *Department of Earth and Environmental Sciences, University of Milano-Bicocca, 20126 Milan, Italy;* ⓞ orcid.org/0000-0003-3028-0368

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c00114

Author Contributions

S.M., L.B., and L.C. conceived and designed the study. S.M. and L.C. conceived the computational protocol. L.C. performed the computational studies. L.C. and S.M. analyzed and interpreted the data. All the authors contributed to writing of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ ABBREVIATIONS

MD, molecular dynamics; SMD, steered molecular dynamics; MetaD, metadynamics; PMF, potential of the mean force; CVs, collective variables; FES, free-energy surface; PCVs, path collective variables; HIF-2α, hypoxia-inducible factor 2α; ARNT, aryl hydrocarbon receptor nuclear translocator; bHLH-PAS, basic helix−loop−helix/Per-ARNT-SIM; AhR, aryl hydrocarbon receptor; RMSD, root-mean-square deviation

■ REFERENCES

(1) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov.* **2004**, *3*, 935−949.
(2) Lexa, K. W.; Carlson, H. A. Protein flexibility in docking and surface mapping. *Q. Rev. Biophys.* **2012**, *45*, 301−343.
(3) Buonfiglio, R.; Recanatini, M.; Masetti, M. Protein Flexibility in Drug Discovery: From Theory to Computation. *ChemMedChem* **2015**, *10*, 1141−1148.
(4) Amaro, R. E.; Baron, R.; McCammon, J. A. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 693−705.
(5) Totrov, M.; Abagyan, R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178−184.
(6) Ellingson, S. R.; Miao, Y.; Baudry, J.; Smith, J. C. Multiconformer ensemble docking to difficult protein targets. *J. Phys. Chem. B* **2015**, *119*, 1026−1034.
(7) Motta, S.; Bonati, L. Modeling Binding with Large Conformational Changes: Key Points in Ensemble-Docking Approaches. *J. Chem. Inf. Model.* **2017**, *57*, 1563−1578.
(8) Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.* **2012**, *41*, 429−452.

(9) Bernardi, R. C.; Melo, M. C. R.; Schulten, K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850*, 872−877.

(10) Doshi, U.; Hamelberg, D. Towards fast, rigorous and efficient conformational sampling of biomolecules: Advances in accelerated molecular dynamics. *Biochim. Biophys. Acta - Gen. Subj.* **2015**, *1850*, 878−888.

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

(12) Sabbadin, D.; Moro, S. Supervised molecular dynamics (SuMD) as a helpful tool to depict GPCR-ligand recognition pathway in a nanosecond time scale. *J. Chem. Inf. Model.* **2014**, *54*, 372−376.

(13) Dickson, A.; Lotz, S. D. Multiple Ligand Unbinding Pathways and Ligand-Induced Destabilization Revealed by WExplore. *Biophys. J.* **2017**, *112*, 620−629.

(14) Zeller, F.; Luitz, M. P.; Bomblies, R.; Zacharias, M. Multiscale Simulation of Receptor−Drug Association Kinetics: Application to Neuraminidase Inhibitors. *J. Chem. Theory Comput.* **2017**, *13*, 5097−5105.

(15) Spitaleri, A.; Decherchi, S.; Cavalli, A.; Rocchia, W. Fast Dynamic Docking Guided by Adaptive Electrostatic Bias: The MD-Binding Approach. *J. Chem. Theory Comput.* **2018**, *14*, 1727−1736.

(16) Basciu, A.; Malloci, G.; Pietrucci, F.; Bonvin, A. M. J. J.; Vargiu, A. V. Holo-like and Druggable Protein Conformations from Enhanced Sampling of Binding Pocket Volume and Shape. *J. Chem. Inf. Model.* **2019**, *59*, 1515−1528.

(17) Deb, I.; Frank, A. T. Accelerating Rare Dissociative Processes in Biomolecules Using Selectively Scaled MD Simulations. *J. Chem. Theory Comput.* **2019**, *15*, 5817−5828.

(18) Lu, H.; Schulten, K. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 453−463.

(19) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562−12566.

(20) Milles, L. F.; Schulten, K.; Gaub, H. E.; Bernardi, R. C. Molecular mechanism of extreme mechanostability in a pathogen adhesin. *Science* **2018**, *359*, 1527−1533.

(21) Hsin, J.; Strümpfer, J.; Lee, E. H.; Schulten, K. Molecular Origin of the Hierarchical Elasticity of Titin: Simulation, Experiment, and Theory. *Annu. Rev. Biophys.* **2011**, *40*, 187−203.

(22) Giorgino, T.; De Fabritiis, G. A high-throughput steered molecular dynamics study on the free energy profile of ion permeation through gramicidin A. *J. Chem. Theory Comput.* **2011**, *7*, 1943−1950.

(23) Nademi, Y.; Tang, T.; Uludağ, H. Steered molecular dynamics simulations reveal a self-protecting configuration of nanoparticles during membrane penetration. *Nanoscale* **2018**, *10*, 17671−17682.

(24) Jorgensen, W. L. Pulled from a protein's embrace Closing in on evaders. *Nature* **2010**, *466*, 42−43.

(25) Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. *J. Am. Chem. Soc.* **2010**, *132*, 7361−7371.

(26) Do, P. C.; Lee, E. H.; Le, L. Steered Molecular Dynamics Simulation in Rational Drug Design. *J. Chem. Inf. Model.* **2018**, *58*, 1473−1482.

(27) Patel, J. S.; Berteotti, A.; Ronsisvalle, S.; Rocchia, W.; Cavalli, A. Steered molecular dynamics simulations for studying protein-ligand interaction in cyclin-dependent kinase 5. *J. Chem. Inf. Model.* **2014**, *54*, 470−480.

(28) Limongelli, V. Ligand binding free energy and kinetics calculation in 2020. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2020**, *10*, 1−32.

(29) Jarzynski, C. Rare events and the convergence of exponentially averaged work values. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **2006**, *73*, No. 046105.

(30) Laio, A.; Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports Prog. Phys.* **2008**, *71*, 126601.

(31) Bussi, G.; Branduardi, D. Free-Energy Calculations with Metadynamics: Theory and Practice. *Rev. Comput. Chem.* **2015**, *28*, 1−49.

(32) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **2007**, *126*, No. 054103.

(33) Limongelli, V.; Marinelli, L.; Cosconati, S.; La Motta, C.; Sartini, S.; Mugnaini, L.; Da Settimo, F.; Novellino, E.; Parrinello, M. Sampling protein motion and solvent effect during ligand binding. *Proc. Natl. Acad. Sci.* **2012**, *109*, 1467−1472.

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

(35) Bernetti, M.; Masetti, M.; Recanatini, M.; Amaro, R. E.; Cavalli, A. An Integrated Markov State Model and Path Metadynamics Approach To Characterize Drug Binding Processes. *J. Chem. Theory Comput.* **2019**, *15*, 5689−5702.

(36) Cavalli, A.; Spitaleri, A.; Saladino, G.; Gervasio, F. L. Investigating drug-target association and dissociation mechanisms using metadynamics-based algorithms. *Acc. Chem. Res.* **2015**, *48*, 277−285.

(37) Comitani, F.; Limongelli, V.; Molteni, C. The free energy landscape of GABA binding to a pentameric ligand-gated ion channel and its disruption by mutations. *J. Chem. Theory Comput.* **2016**, *12*, 3398−3406.

(38) Limongelli, V.; Bonomi, M.; Marinelli, L.; Gervasio, F. L.; Cavalli, A.; Novellino, E.; Parrinello, M. Molecular basis of cyclooxygenase enzymes (COXs) selective inhibition. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5411−5416.

(39) 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**, *139*, 4780−4788.

(40) Saleh, N.; Saladino, G.; Gervasio, F. L.; Haensele, E.; Banting, L.; Whitley, D. C.; Sopkova-de Oliveira Santos, J.; Bureau, R.; Clark, T. A Three-Site Mechanism for Agonist/Antagonist Selective Binding to Vasopressin Receptors. *Angew. Chemie Int. Ed.* **2016**, *55*, 8008−8012.

(41) Saleh, N.; Ibrahim, P.; Saladino, G.; Gervasio, F. L.; Clark, T. An Efficient Metadynamics-Based Protocol To Model the Binding Affinity and the Transition State Ensemble of G-Protein-Coupled Receptor Ligands. *J. Chem. Inf. Model.* **2017**, *57*, 1210−1217.

(42) Raniolo, S.; Limongelli, V. Ligand binding free-energy calculations with funnel metadynamics. *Nat. Protoc.* **2020**, *15*, 2837−2866.

(43) Limongelli, V.; Bonomi, M.; Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci.* **2013**, *110*, 6358−6363.

(44) Bersten, D. C.; Sullivan, A. E.; Peet, D. J.; Whitelaw, M. L. bHLH-PAS proteins in cancer. *Nat. Rev. Cancer* **2013**, *13*, 827−841.

(45) Wu, D.; Potluri, N.; Lu, J.; Kim, Y.; Rastinejad, F. Structural integration in hypoxia-inducible factors. *Nature* **2015**, *524*, 303−308.

(46) Kewley, R. J.; Whitelaw, M. L.; Chapman-Smith, A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int. J. Biochem. Cell Biol.* **2004**, *36*, 189−204.

(47) McIntosh, B. E.; Hogenesch, J. B.; Bradfield, C. A. Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annu. Rev. Physiol.* **2010**, *72*, 625−645.

(48) Denison, M. S.; Soshilov, A. A.; He, G.; DeGroot, D. E.; Zhao, B. Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sci.* **2011**, *124*, 1−22.

(49) Giani Tagliabue, S.; Faber, S. C.; Motta, S.; Denison, M. S.; Bonati, L. Modeling the binding of diverse ligands within the Ah receptor ligand binding domain. *Sci. Rep.* **2019**, *9*, 10693.

(50) Scheuermann, T. H.; Tomchick, D. R.; Machius, M.; Guo, Y.; Bruick, R. K.; Gardner, K. H. Artificial ligand binding within the HIF2alpha PAS-B domain of the HIF2 transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 450−455.

(51) Rogers, J. L.; Bayeh, L.; Scheuermann, T. H.; Longgood, J.; Key, J.; Naidoo, J.; Melito, L.; Shokri, C.; Frantz, D. E.; Bruick, R. K.; Gardner, K. H.; Macmillan, J. B.; Tambar, U. K. Development of Inhibitors of the PAS - B Domain of the HIF-2 α Transcription Factor. *J. Med. Chem.* **2013**, *56*, 1739−1747.

(52) Scheuermann, T. H.; Li, Q.; Ma, H. W.; Key, J.; Zhang, L.; Chen, R.; Garcia, J. A.; Naidoo, J.; Longgood, J.; Frantz, D. E.; Tambar, U. K.; Gardner, K. H.; Bruick, R. K. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat. Chem. Biol.* **2013**, *9*, 271−276.

(53) Key, J.; Scheuermann, T. H.; Anderson, P. C.; Daggett, V.; Gardner, K. H. Principles of ligand binding within a completely buried cavity in HIF2alpha PAS-B. *J. Am. Chem. Soc.* **2009**, *131*, 17647−17654.

(54) Wallace, E. M.; Rizzi, J. P.; Han, G.; Wehn, P. M.; Cao, Z.; Du, X.; Cheng, T.; Czerwinski, R. M.; Dixon, D. D.; Goggin, B. S.; Grina, J. A.; Halfmann, M. M.; Maddie, M. A.; Olive, S. R.; Schlachter, S. T.; Tan, H.; Wang, B.; Wang, K.; Xie, S.; Xu, R.; Yang, H.; Josey, J. A. A Small-Molecule Antagonist of HIF2α Is Efficacious in Preclinical Models of Renal Cell Carcinoma. *Cancer Res.* **2016**, *76*, 5491−5500.

(55) Chen, W.; Hill, H.; Christie, A.; Kim, M. S.; Holloman, E.; Pavia-Jimenez, A.; Homayoun, F.; Ma, Y.; Patel, N.; Yell, P.; Hao, G.; Yousuf, Q.; Joyce, A.; Pedrosa, I.; Geiger, H.; Zhang, H.; Chang, J.; Gardner, K. H.; Bruick, R. K.; Reeves, C.; Hwang, T. H.; Courtney, K.; Frenkel, E.; Sun, X.; Zojwalla, N.; Wong, T.; Rizzi, J. P.; Wallace, E. M.; Josey, J. A.; Xie, Y.; Xie, X.-J.; Kapur, P.; McKay, R. M.; Brugarolas, J. Targeting renal cell carcinoma with a HIF-2 antagonist. *Nature* **2016**, *539*, 112−117.

(56) Cho, H.; Du, X.; Rizzi, J. P.; Liberzon, E.; Chakraborty, A. A.; Gao, W.; Carvo, I.; Signoretti, S.; Bruick, R. K.; Josey, J. A.; Wallace, E. M.; Kaelin, W. G. On-target efficacy of a HIF-2α antagonist in preclinical kidney cancer models. *Nature* **2016**, *539*, 107−111.

(57) Motta, S.; Minici, C.; Corrada, D.; Bonati, L.; Pandini, A. Ligand-induced perturbation of the HIF-2 α :ARNT dimer dynamics. *PLoS Comput. Biol.* **2018**, *14*, No. e1006021.

(58) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The protein data bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(59) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.

(60) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very fast prediction and rationalization of pKa values for protein-ligand complexes. *Proteins* **2008**, *73*, 765−783.

(61) *Schrödinger Release 2020-1*: LigPrep, Schrödinger, LLC, New York, NY. (2020).

(62) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(63) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(64) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(65) 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*, 1157−1174.

(66) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to super-computers. *SoftwareX* **2015**, *1-2*, 19−25.

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

(68) 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*, 3696−3713.

(69) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(70) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, No. 014101.

(71) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.

(72) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(73) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(74) *Macromodel versione 10.3*, Schrödinger LLC, New York. (2014).

(75) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.

(76) *Glide versione 6.2*, Schrödinger LLC, New York. (2014).

(77) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006**, *49*, 6177−6196.

(78) Patel, J. S.; Branduardi, D.; Masetti, M.; Rocchia, W.; Cavalli, A. Insights into ligand-protein binding from local mechanical response. *J. Chem. Theory Comput.* **2011**, *7*, 3368−3378.

(79) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.

(80) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.

(81) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.

(82) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, No. 020603.

(83) Fidelak, J.; Juraszek, J.; Branduardi, D.; Bianciotto, M.; Gervasio, F. L. Free-energy-based methods for binding profile determination in a congeneric series of CDK2 inhibitors. *J. Phys. Chem. B* **2010**, *114*, 9516−9524.

(84) Hovan, L.; Comitani, F.; Gervasio, F. L. Defining an Optimal Metric for the Path Collective Variables. *J. Chem. Theory Comput.* **2019**, *15*, 25−32.

(85) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Peptide Folding: When Simulation Meets Experiment. *Angew. Chemie Int. Ed.* **1999**, *38*, 236−240.

(86) Scheuermann, T. H.; Stroud, D.; Sleet, C. E.; Bayeh, L.; Shokri, C.; Wang, H.; Caldwell, C. G.; Longgood, J.; MacMillan, J. B.; Bruick, R. K.; Gardner, K. H.; Tambar, U. K. Isoform-Selective and Stereoselective Inhibition of Hypoxia Inducible Factor-2. *J. Med. Chem.* **2015**, *58*, 5930−5941.

(87) Koyasu, S.; Kobayashi, M.; Goto, Y.; Hiraoka, M.; Harada, H. Regulatory mechanisms of hypoxia-inducible factor 1 activity: Two decades of knowledge. *Cancer Sci.* **2018**, *109*, 560−571.

(88) Wu, D.; Su, X.; Lu, J.; Li, S.; Hood, B. L.; Vasile, S.; Potluri, N.; Diao, X.; Kim, Y.; Khorasanizadeh, S.; Rastinejad, F. Bidirectional modulation of HIF-2 activity through chemical ligands. *Nat. Chem. Biol.* **2019**, *15*, 367−376.

(89) Burendahl, S.; Danciulescu, C.; Nilsson, L. Ligand unbinding from the estrogen receptor: A computational study of pathways and ligand specificity. *Proteins Struct. Funct. Bioinforma.* **2009**, *77*, 842−856.

(90) Hu, X.; Hu, S.; Wang, J.; Dong, Y.; Zhang, L.; Dong, Y. Steered molecular dynamics for studying ligand unbinding of ecdysone receptor. *J. Biomol. Struct. Dyn.* **2018**, *36*, 3819−3828.

(91) Shen, J.; Li, W.; Liu, G.; Tang, Y.; Jiang, H. Computational insights into the mechanism of ligand unbinding and selectivity of estrogen receptors. *J. Phys. Chem. B* **2009**, *113*, 10436−10444.

(92) Shirts, M. R.; Mobley, D. L.; Brown, S. P. in *Free. Calc. Struct. drug Des.* (eds. Merz, K. M.; Ringe, D.; Reynolds, C. H.) 61−86 (Cambridge University Press, 2010).

(93) Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. in *Steered Mol. Dyn.* (eds. Deuflhard, P.; Hermans, J.; Leimkuhler, B.; Mark, A. E.; Reich, S.; Skeel, R. D.) 39−65 (Springer, 1999).

(94) Capelli, R.; Carloni, P.; Parrinello, M. Exhaustive Search of Ligand Binding Pathways via Volume-Based Metadynamics. *J. Phys. Chem. Lett.* **2019**, *10*, 3495−3499.