Insights into Ligand−Protein Binding from Local Mechanical Response

Jagdish Suresh Patel,[†] Davide Branduardi,*,[†] Matteo Masetti,[†,‡] Walter Rocchia,[†] and Andrea Cavalli*,[†,‡]

[†]Department of Drug Discovery and Development, Italian Institute of Technology, Via Morego 30, I-16163 Genova, Italy

[‡]Department of Pharmaceutical Sciences, Alma Mater Studiorum, Bologna University, Via Belmeloro 6, I-40126 Bologna, Italy

(S) Supporting Information

ABSTRACT: Computational studies of ligand−protein binding are crucial for properly designing novel compounds of potential pharmacological interest. In this respect, researchers are increasingly interested in steered molecular dynamics for ligand−protein binding and unbinding studies. In particular, it has been suggested that analyzing the work profiles along the ligand−protein undocking paths could be fruitful. Here, we propose that small portions of work profiles, termed "local mechanical responses" of the system to a steering force, could serve as a universal measure for capturing relevant information about the system under investigation. Specifically, we first collected a high number of steering trajectories using two biological systems of increasing complexity (i.e., alanine dipeptide and (R)-roscovitine/CDK5 complex). Then, we devised a novel postprocessing tool to be applied to the local mechanical responses, to extract structural information related to the biological processes under investigation. Despite the out-of-equilibrium character of the trajectories, the analysis carried out on the work profiles provided pivotal information about the investigated biological processes. This could eventually be applied to drug design.

INTRODUCTION

Computational drug design has two major goals: (i) the accurate estimation of ligand−protein binding free energy; and (ii) the disclosure of the structural determinants responsible for ligand−protein recognition and binding. In this scenario, molecular dynamics (MD)-based enhanced sampling methods play an increasingly relevant role. As far as the ligand−protein binding free energy (ΔG_b) is concerned, the most widely used strategies are based on alchemical transformations[1] (double decoupling[2,3] and related schemes), where the ΔG_b is estimated using a thermodynamic cycle.[1,4−7] Notwithstanding the impressive results achieved, alchemical transformations[8,9] do not explicitly account for the dynamical events occurring upon ligand−protein recognition and binding. Notably, such dynamical events can be of paramount importance in drug discovery by providing fundamental drug design information.[10] Moreover, simulations of the unbinding process are relevant for the binding kinetics, and thus for drug residence time within the target.[11,12] Very recently, the first promising attempt to obtain the kinetics from straightforward simulations appeared[13] and required a collective computational effort through distributed networks. Therefore, for the time being, umbrella sampling,[14,15] metadynamics,[16,17] and steered MD[18] are the methods of choice to disclose the structural determinants relevant to a ligand binding to a protein on the exit pathways. In particular, metadynamics has proven to be rather effective,[17,19] but its computational cost is not a priori predictable even in an exploratory regime (i.e., when convergence of the free energy is not required). Steered MD is also becoming very popular for studying biophysical processes.[20−26] This is partly due to its conceptual simplicity and integration in several currently available MD codes. In steered MD, a certain transition (such as ligand−protein unbinding) is obtained via a tunable restraining potential, which forces the system to move away from its initial configuration (e.g., a bound state for ligand−protein complexes) to a given position during an MD run.[18] In the case of an unbinding process, the target position may be defined as an unbound state where the ligand−protein interactions may be considered negligible. In steered MD, the simulation time required to complete the transition is an input parameter, which can be reduced to an almost arbitrary small number, making the technique particularly appealing in the drug discovery process. Moreover, some recent attempts to obtain ΔG_b of binding have been reported,[25] but these are limited to very simple model ligands. Concerning more realistic cases, Colizzi et al.[24] have demonstrated that steered MD can be applied to drug design-related problems without requiring an accurate estimation of the ΔG_b. In particular, Colizzi et al. could discern active from inactive enzyme inhibitors by a simple visual inspection of the force profiles required for pulling ligands out of the protein binding site. Increasing the number of steered MD pulling trajectories provides a natural extension to improve the overall reliability of this approach,[26−28] and this is becoming accessible due to the ongoing increase in CPU performance. Although much recent effort has been devoted to analyzing configurations generated by MD runs,[29−31] it is not easy to extract relevant structural information from an ensemble of steered MD work profiles for a complex system. This is due to the lack of specific and effective analysis tools.

Here, we report on a novel postprocessing strategy aimed at analyzing steered MD trajectories and extracting the structural features that are relevant to the biological process under investigation.
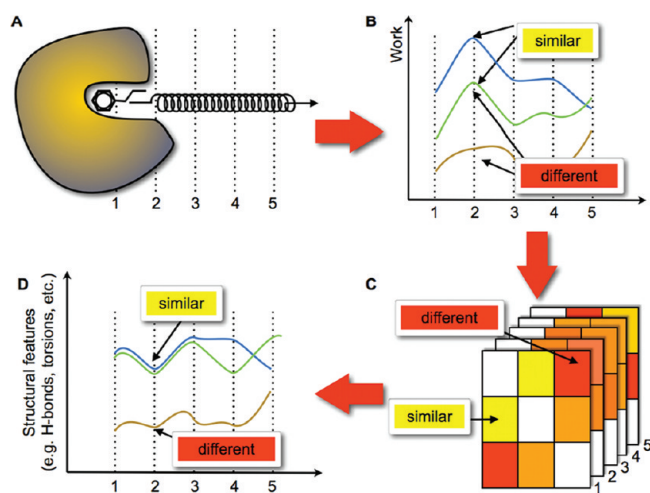
Figure 1. A sketch of the general workflow used in this study. Steered MD simulations are carried out (A), work profiles are subsequently generated (B), and a set of matrices that record the similarity among profiles are produced (C). The matrices are then postprocessed, and the similarities in the work curves are correlated with structural elements (D).

We used the work profile calculated on the system in a limited region of space, hereafter referred to as local mechanical response (LMR), as a measure of the ligand—protein interaction strength. The LMR profiles were analyzed via multidimensional scaling (MDS),[32] which allowed us to extract structural similarity/dissimilarity over a large set of steered MD-derived LMRs. We also report on using this postprocessing tool to correlate work profiles with the relevant structural changes occurring during a certain process (see Figure 1 for a schematic representation of the overall process). In particular, we show that: (i) reasonably fast pulling regimes can provide relevant information on structurally different reactive pathways; (ii) by analyzing LMR patterns, it is possible to point out the structural elements important for the unbinding process; (iii) LMR emerges as a simple and system-independent observable that has general applicability for ligand—protein binding studies. To illustrate the usefulness of this novel approach, two case studies of increasing complexity are here investigated.

■ METHODS

Work Estimate in Steered MD Simulations. In the thermodynamic integration formalism,[33] the free energy difference associated with the continuous change of the system from an initial state described with an Hamiltonian $H(\lambda_0)$ to a final one $H(\lambda_1)$ obtained by changing the parameter λ is given by

$$\Delta G = \int_{\lambda_0}^{\lambda_1} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle d\lambda \qquad (1)$$

One possible way to achieve the transition is to add to the standard Hamiltonian a harmonic potential $U_{bias}(t,\mathbf{r})$ acting on a descriptor $s(\mathbf{r})$ (e.g., the ligand—protein distance or the mean square deviation with respect to a given structure), which holds the following time dependency:

$$U(t,\mathbf{r}) = k(s(\mathbf{r}) - s_0 - vt)^2 \qquad (2)$$

where s_0 is the value of the descriptor in the initial state $H(\lambda_0)$, t is the time, and k is a numerical constant. Thus, whenever time is considered in place of the parameter λ, the partial derivative of the Hamiltonian turns out to be the instantaneous value, and the integral in eq 1 (which can be easily calculated via quadrature or trapezoidal rule) corresponds to the work ΔW exerted on the system:

$$\Delta W = \int_{t_0}^{t_1} \left\langle \frac{\partial H(t)}{\partial t} \right\rangle dt = -2kv \int_{t_0}^{t_1} (s(\mathbf{r}) - s_0 - vt) dt \qquad (3)$$

After a predetermined amount of time, the center of the harmonic constraint will be located in its final position:

$$s_1 = s_0 + vt_1 \qquad (4)$$

Therefore, whenever the spring constant k chosen is large enough (stiff-spring regime), it is reasonable to assume that, at the final time t_1, the system has approximately reached the point s_1. Moreover, when the $U_{bias}(t,\mathbf{r})$ is applied in a quasistatic regime (slow growth), the calculated work is equivalent to the free energy estimate obtained by thermodynamic integration:[28]

$$\Delta W(t_1, t_0) \cong \Delta G(s_1, s_0) \qquad (5)$$

This is a translation of the classical result, which states that the work exerted by an external potential to move a system quasistatically from an initial to a final state is equal to the free energy difference between these two states.

From a practical standpoint, Crooks theorem[27,34] or Jarzynski[27] equality are better suited to evaluate free energies from out-of-equilibrium trajectories, since the quasistatic limit is practically never reached. Thus, in such cases the work profile turns out to be rather different from the actual free energy landscape. Here, we explore the hypothesis that, even in a non-quasistatic regime, the work profiles retain some information about the structural events associated with the mechanical response induced by the steering procedure. Hereafter, this will be referred to as the "mechanical response" profile.

To better understand the effects of the application of a steering potential, it is useful to consider the limiting case where the transition is carried out in a single step. During such a time span (typically 2 fs for a classical MD simulation), the system can be considered frozen, and therefore the estimated work will be

$$\Delta W = k(s_1 - s_0)^2 \qquad (6)$$

In this case, the path followed by the system (e.g., ligand and protein in the case of an undocking experiment) becomes irrelevant, and all of the contribution to the work comes from the restraint position. Moreover, being system independent, the estimated work cannot be by any means similar to the free energy of binding, which, in contrast, is strongly system dependent. This conceptual experiment shows that, although the requirement of quasistatic transformation can be somewhat relaxed, the simulations should nevertheless be performed with a steering velocity that can capture the mechanical response of the system subjected to the typical relaxation time of the variables implied in the investigated process. In accordance with this limit, the ligand acts as a probe for ligand—protein interactions, and the trajectories may retain some physicochemical relevance. This pulling velocity range is worth exploiting. This is because it is computationally affordable and preferable to both very fast steering regimes (which are not sensitive to the ligand—protein rearrangement) and very low velocities (which are required whenever accurate
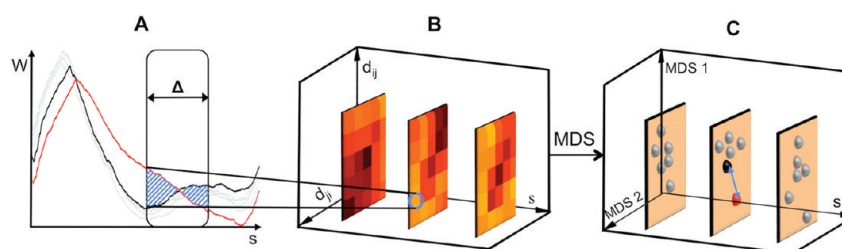
3369

dx.doi.org/10.1021/ct200324j |*J. Chem. Theory Comput.* 2011, 7, 3368–3378

Figure 2. Schematic representation of the postprocessing analysis of the work curves. The first step (A) is the calculation of the distance between mechanical responses recorded on a small portion of the order parameter s (blue-shaded region). The second step (B) is the collection of the distance matrices along the order parameter, and the last step (C) is the retrieval of a fictitious representation via MDS in a reduced dimensionality for each of the matrices. MDS1 and MDS2 are the fictitious coordinates produced by the MDS algorithm.

free energy estimations are needed). The latter case in particular must meet a stricter requirement: Whenever multiple paths are competing, the correct free energy is obtained only when the relative occurrence of different pathways is at convergence. This is particularly unlikely when each path is separated from the others by a large barrier. In the present study, we show that structurally relevant information about the different paths can be gathered even without obtaining full convergence on the statistics among competing pathways. This results in a large computational saving.

MDS Analysis. Since a detailed analysis of a considerable number of trajectories may be cumbersome, we want to test the hypothesis that adopting a postprocessing technique for the mechanical response profiles $W(s)$ may be helpful in capturing the principal structural trends in the collected ensemble of trajectories. The rationale is that structural features are expected to affect the resistance that the system opposes to the external force and, in turn, the mechanical response profiles $W(s)$. Although reasonable, this assumption must be kept in mind while postprocessing two similar work profiles, since this does not exclude the possibility that they represent two different pathways.

In the stiff spring regime we can assume that $s(t_1) \approx s_1$ and can define the n-th mechanical response profile[18,28] as

$$W_n(s_1) \cong W_n(t_1) = -2k\nu \int_{t_0}^{t_1} (s(\mathbf{r}) - s_0 - \nu t)\mathrm{d}t \qquad (7)$$

A portion of the mechanical response within a certain interval of s, of width Δ, is termed here "local mechanical response" (LMR). The distance between two LMRs obtained from two steered MD trajectories within a certain interval of s, of width Δ, may be defined as

$$d_{nm}(s_0, \Delta) = \int_{s_0}^{s_0 + \Delta} [W_n(s) - \langle W_n(s) \rangle_\Delta - W_m(s)$$

$$+ \langle W_m(s) \rangle_\Delta]^2 \mathrm{d}s \qquad (8)$$

with $\langle W_n(s) \rangle_\Delta$ being the average work over the interval:

$$\langle W_n(s) \rangle_\Delta = \frac{\int_{s_0}^{s_0 + \Delta} W_n(s)\mathrm{d}s}{\Delta} \qquad (9)$$

This defines a set of distance matrices along the pulling coordinate that can be exploited to monitor different families of LMRs along the steering pathway. A brief sketch of this step is

represented in Figure 2A and B. One valuable outcome of calculating the local difference between two profiles over blocks of fixed size Δ (instead of calculating it on the work profile over the whole transition) is the crucially important possibility of detecting branching in the mechanical response intersection among different realizations. To avoid irregular behaviors at the borders between adjacent blocks, some degree of overlap between them is allowed. Clearly, the window size Δ must be properly tuned in accordance with the scales of the events involved in the transition. For the ligand unbinding event investigated here, Δ was optimally sized to identify hydrogen (H)-bonding breaking/formation or local conformational rearrangements, while much finer thermal motions were chosen to be averaged out.

For the sake of completeness, we note that the same information could be obtained by averaging the exerted forces rather than using work. Here, we prefer to consider the work profiles because they may provide an interesting view of the dissipative work produced and a useful hint concerning the appropriateness of the pulling parameter as well as the amount of orthogonal degrees of freedom interfering with the pulling direction.

Once the set of matrices $d_{nm}(s, \Delta)$ is obtained, there are two possible ways to postprocess them. One approach relies on standard clustering techniques. Because of the intrinsic need for a clustering threshold value whose choice in the case of the work curve realizations would be arbitrary and far from simple, we decided to adopt a MDS approach[29,32] so as to have a direct grasp over the topology of the LMR pattern. Hence, we term this analysis LMR-MDS.

MDS is a standard pattern-recognition technique[32] that can detect the intrinsic dimensionality from a distance matrix by searching for the lowest dimensional possible manifold that can preserve the reciprocal neighbor distances and produce a representation that resembles the original topology in the original high-dimensional space.

The simple MDS algorithm used here initially picks a random entry $d_{nm}(s, \Delta)$ (i.e., the distance between the n and m work profiles over an interval of width Δ centered in a given point s along the steering coordinate) in the high M-dimensional original distance matrix $d(s, \Delta)$ and creates two fictitious points in a low D-dimensional Euclidean space at the same distance $d_{nm}^{\mathrm{MDS}}(s, \Delta)$. Then, another entry on the same row n but at a different column l is picked ($d_{nl}(s, \Delta)$) and used to constrain the distance from n when a new point representing l is placed in the D-dimensional space. Since this operation is not univocal, a Monte Carlo (MC) procedure is used to satisfy as much as possible all the distances with respect to the already projected elements. The procedure is iterated until the entire original distance matrix $d(s, \Delta)$ is spanned. At the end, the positioning of

3370

dx.doi.org/10.1021/ct200324j |*J. Chem. Theory Comput.* 2011, 7, 3368–3378

the last points is subject to a larger number of constraints than the initial ones. To reduce the residual strain, we therefore applied a MC annealing procedure, adopting the merit function $E_{MC}(s,\Delta)$, also called "stress" function in MDS terminology, to all the points in the D-dimensional representation. The stress function $E_{MC}(s,\Delta)$ is the squared difference in the representation of the two matrices, the M-dimensional and its fictitious D-dimensional representation:

$$E_{MC}(s,\Delta) = \sum_{n<m}^{N} [d_{nm}(s,\Delta) - d_{nm}^{MDS}(s,\Delta)]^2 \quad (10)$$

In order to evaluate the appropriateness of the choice of D, the preservation of an order relation within the distance matrices is verified. In the negative case, the dimension is increased, and the entire procedure is repeated.

The postprocessing procedure consists of applying the above procedure for each interval Δ along the steering coordinate. The final outcome is thus a set of reduced representations for this degree of freedom (see Figure 2C).

At the end of the procedure, one gets an intuitive representation of the different families of steering processes that may occur. By inspecting the structural differences between representative members of different families, we obtain a picture of the different processes occurring at the molecular level during ligand unbinding. This final step is represented in Figure 1C and D.

We note that a similar postprocessing tool can be applied to any set of distance matrices. Therefore, when studying the alanine dipeptide, we also applied MDS to the Cartesian coordinates of the atoms. In this case, each entry of the matrix $d_{nm}^{RMSD}(s,\Delta)$ was obtained using the root mean square deviation (RMSD) of the heavy atoms of two structures along the steering path, obtained after optimal alignment through the Kearsley algorithm.[35] We term this analysis "Cartesian-MDS" to differentiate it from the work-based LMR-MDS. Their comparison (see Results Section) is instrumental in verifying the connection between the local mechanical response and the structural changes of the system.

It is worth noting that the MDS implementation here described was intentionally rather unsophisticated, and for this specific problem that included a data set with relatively modest size, our MDS approach could be basically equivalent to other more advanced methods, like classical Torgerson multidimensional scaling. The latter approach has to be highly recommended when a larger data set has to be analyzed.[32]

Simulation Details. All the MD simulations in the present work were carried out with NAMD2.7 code.[36] The simulations in the NVT ensemble were performed using the Langevin thermostat,[37] and additional steering forces were introduced via the PLUMED[38] plugin integrated in the NAMD code.

For alanine dipeptide (see Figure 3A for a molecular sketch), we used CHARMM27 force field,[39] a time step of 0.2 fs without constraining the covalent-bond length involving hydrogen atoms so as to maximize the number of degrees of freedom involved. A Langevin thermostat at 300 K with relaxation time of 8 ps was used to maintain the average temperature during out-of-equilibrium runs. We performed steered MD simulation using mean square displacement (MSD) of the heavy atoms with respect to C_{7ax} configuration as a pulling coordinate. Optimal alignment was obtained using the Kearsley algorithm.[35] A number of tests were performed to choose the speed for pulling and the magnitude of the spring constant. A limited dissipative work was obtained with a value of 2000 kcal/($Å^4 \cdot$mol) for the spring constant and 0.005 $Å^2$/ps
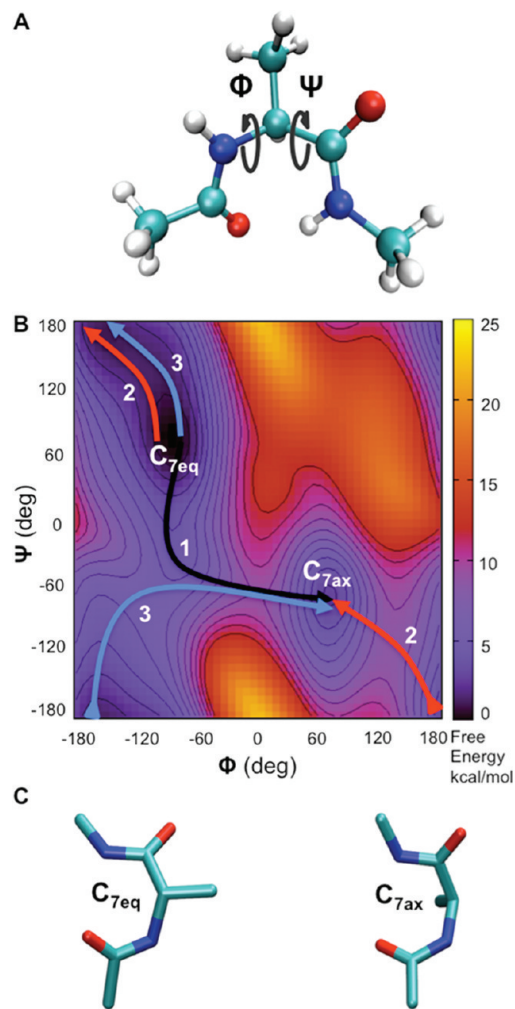

Figure 3. Structural and free energy features of alanine dipeptide. (A) Ball and stick representation of alanine dipeptide along with the Φ and Ψ dihedrals used in the Ramachandran plot are shown. (B) Free energy landscape of alanine dipeptide as a function of the Φ and Ψ dihedrals produced via umbrella sampling. The isoline separation is of 1.0 kcal/mol. Both C_{7eq} and C_{7ax} minima are connected by three different pathways denoted with three different colors (black, red, and blue). (C) Stick representation of the two metastable conformers C_{7eq} and C_{7ax} (heavy atoms only).

for the pulling speed (Supporting Information, Figure S1). Unless specified, all the molecular representations were generated using VMD.[40]

For (R)-roscovitine/CDK5 complex, the starting geometry used in the simulations was obtained after removing the p25 activator from the X-ray structure retrieved from the Protein Data Bank (PDB code: 1UNL).[41] The Amber parm99SB[42] force field was used for the protein, while the (R)-roscovitine was treated with the general Amber force field for organic molecules[43] and the charges were derived according to the restrained electrostatic potential (RESP) procedure.[44] Prior to the steered MD simulations, the system was minimized and equilibrated in a box with 10 371 TIP3P[45] water molecules and pressurized for 2 ns in the NPT ensemble using a Langevin thermostat[37] and a Langevin piston barostat.[46] Long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method.[47] Short-range nonbonded interactions were calculated using a cutoff

radius of 12 Å for both Coulomb and van der Waals potentials. A 2 fs time step was used, and covalent-bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm.[48]

Undocking experiments were performed by means of steered MD. The pulling variable was the distance between the center of mass of the residues belonging to the binding site of the protein and the center of mass of the ligand (see Supporting Information, Figure S2 for a pictorial view of the binding pocket). The pulling parameters (spring constant, pulling velocity, and maximum extension) were determined by performing different trial simulations in different conditions and by comparing the calculated work values. In particular, 12 steered MD runs were performed using a spring constant of 10, 100, 1000 kcal/(mol·Å^2) and pulling velocities of 0.5, 0.2, 0.1, and 0.01 Å/ps. In the end, a spring constant of 100 kcal/(mol·Å^2) along with a pulling velocity of 0.01 Å/ps were chosen, since they provided the lowest work values in stiff spring regime. The target distance for the steered MD simulations was chosen to be 22 Å. This is because preliminary runs showed that, at this distance, the ligand was completely detached from the target (Supporting Information, Figure S3). Moreover, this pulling velocity was shown not to irreversibly disrupt the secondary structure (Supporting Information, Figure S4).

■ RESULTS AND DISCUSSION

The content of this section is here briefly outlined. First, the postprocessing technique was tested on a set of steered MD runs using alanine dipeptide in vacuum. This is a widely used test case for benchmarking various MD-based simulation techniques. Then, the procedure was applied to the pharmacologically relevant complex (R)-roscovitine/CDK5, which is currently being investigated in the search for novel drug candidates to treat Alzheimer's disease.

Alanine Dipeptide. Our prototypical case study is the alanine dipeptide in vacuum (see Figure 3A), a system widely used as a benchmark for enhanced sampling schemes.[15,16,49−52] Alanine dipeptide, in a lower dimension, recapitulates most of the features that are relevant to the free energy landscape of ligand−protein recognition and binding. In particular, it displays two minima: a wider basin that can resemble a ligand in the bulk of the solvent, and a much narrower basin that can be similar to a ligand−protein complex. These two minima are connected by multiple reactive pathways with sizable free energy barriers (see Figure 3B). The two main metastable states are usually denoted as C_{7eq} and C_{7ax} (see Figure 3C), and the transition from one to the other is generally represented in terms of the dihedral angles phi (Φ) and psi (Ψ) (Figure 3A and B). Due to the periodic nature of these descriptors, three distinct pathways connect C_{7eq} to C_{7ax}, and two different free energy barriers can be identified (8.5 and 10.5 kcal/mol; see Figure 3B). Furthermore, alanine dipeptide represents an ideal test case since it allows steered MD to be performed in a fully reversible work regime. This situation is highly desirable as it allows us to minimize the effects of energy dissipation. In contrast, protein−ligand unbinding processes display a large number of degrees of freedom that are only partly orthogonal to the pulling direction, thus showing a much larger relaxation time. As a result, one usually observes a significant increase in the amount of dissipative work, with respect to the alanine dipeptide case study, and the distribution of work values is broader.

Alanine dipeptide was initially investigated by simple MD, which allowed us to generate an ensemble of starting structures in
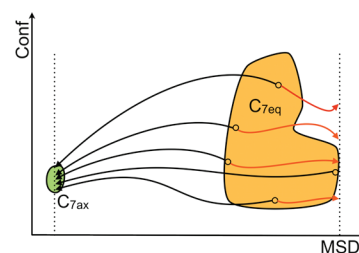

Figure 4. A sketch of the extension procedure adopted for the steered MD experiments in alanine dipeptide. The starting ensemble is represented in orange (C_{7eq}), while the targeted ensemble is in green (C_{7ax}). The starting configurations may assume different positions in the MSD variable. So as not to be limited by the shortest pulling experiment (black arrows), the pulling was extended to the highest MSD value occurring in the equilibrium ensemble (red arrows).

C_{7eq} basin to be used for subsequent steered MD runs. MSD was then used (see Methods Section) as the steering variable s to drive the system from C_{7eq} to C_{7ax} within the targeted MD framework.[53] About 500 simulations were carried out starting from initial configurations to provide a fairly large initial data set of trajectories for building a robust statistics for all three pathways. Since each of the possible paths from C_{7eq} to C_{7ax} may start from a different point in the C_{7eq} basin (Figure 4), the trajectories displayed variable length. To consistently compare them, the steered MD of the shorter trajectories was elongated in a backward direction to cover the same MSD range (overall 2.8 Å^2) as the longest one. These additional steered MD trajectories are represented by red arrows in Figure 4. This allowed us to produce a set of mechanical response profiles spanning homogeneously from 0.0 to 2.8 Å^2, corresponding to the C_{7ax} and C_{7eq} conformation, respectively. The ensuing profiles covered all three possible pathways from C_{7eq} to C_{7ax}. Of these, we considered just 60 work profiles (20 for each pathway) to minimize redundancy within the set. We could thus check the ability of our postprocessing tool to detect the structural features relevant to each reactive pathway and control the reliability of the results. This was possible because structural differences along the C_{7eq}−C_{7ax} transition path are well-known and can easily be traced within the Ramachandran plot.[54]

The work profiles were then used to compute distance matrices along the path (see Methods Section and Figures 1 and 2). The window size Δ was set to 0.80 Å^2 in MSD space. This value was chosen by visually inspecting the mechanical response profiles and selecting the MSD interval in which sizable structural motions occurred. A series of matrices ($d_{nm}(s,\Delta)$) was generated by calculating pair-wise distances between two work profiles (in the LMR space) from a maximum of 2.80 Å^2 to a minimum of 0.04 Å^2 by steps of 0.28 Å^2 (roughly equal to $\Delta/3$). These matrices were then processed by the MDS technique, as reported in the Methods Section. Here, a single dimension was sufficient. This is because additional dimensions did not significantly reduce the difference between the reduced and the full dimensionality distance matrices.

Figure 5A shows the projections of the steered MD trajectories over the Ramachandran plot. One representative path for each family is highlighted with different symbols and color codes. From the plot, it is evident that path 3 (blue path) started very close to path 2 (red path) in the upper left corner (around $\Phi \approx -2.0$ and $\Psi \approx 3.1$ in Figure 5A). By crossing the periodic image, it joined path 1 (black path) at $\Phi \approx -1.0$ and $\Psi \approx 0.0$, finally ending up in C_{7ax}. This sequence of events is fully
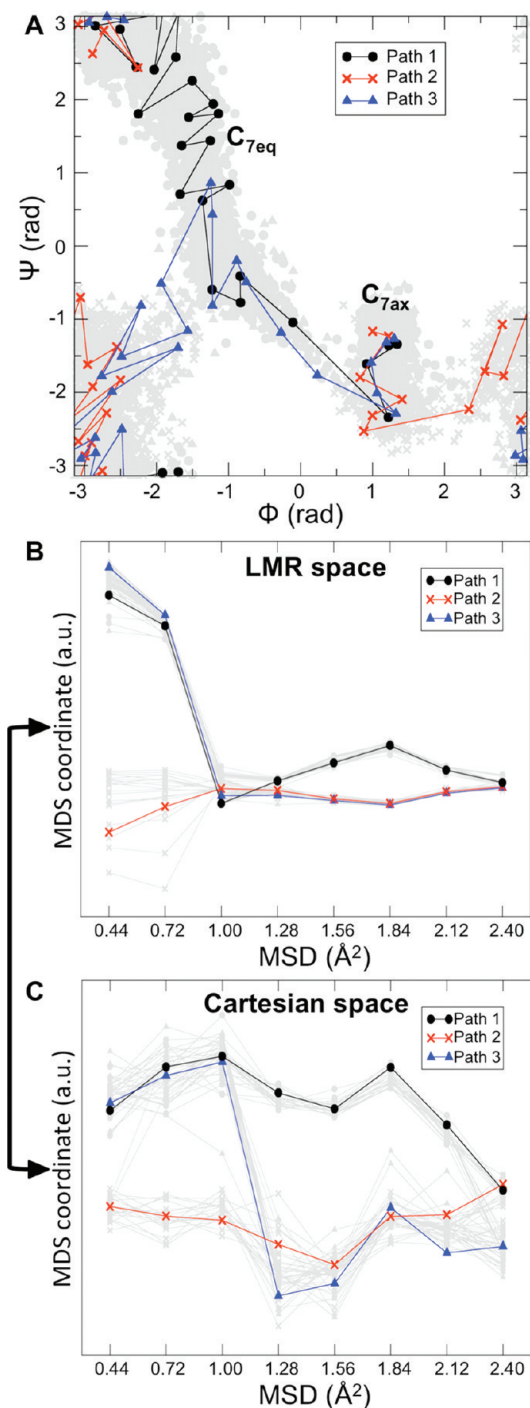
Figure 5. Analysis of alanine dipeptide trajectories. Projections of the steered MD trajectories on the Ramachandran plot (A), LMR-MDS as a function of the steering coordinate (B), and Cartesian-MDS as a function of the steering coordinate (C). Arbitrary units are used for the ordinate in the MDS representation because the coordinates from the MDS representations are fictitious. To guide the eye, in all plots, single representative trajectories for each of the three pathways are shown in black, red, and blue.

represented in the LMR-MDS of Figure 5B: path 2 and path 3 initially overlapped, and at MSD ≈ 1.00 Å^2, path 3 joined path 1.

We performed the same analysis using a Cartesian-MDS. It is important to note here that the comparison between Cartesian

and LMR spaces is instrumental in verifying the connection between structural features and mechanical response. Additionally, we can test the ability of the MDS algorithm to reproduce multidimensional topologies without introducing aberrations in the context of molecular simulations.

In the Cartesian-MDS (Figure 5C), we observed a very similar pattern of events and a remarkably similar distinction in the three pathways. This strengthened our hypothesis that local mechanical response analysis can point to different structural features along several dynamical paths. Indeed, LMR-MDS enhances the distance in those points where the difference in terms of mean force is higher. In our example, two different downhill access routes to C$_{7eq}$ in the mean force space (Figure 5B) emerged as two distinct routes, even though their structures were similar at low MSD values, as reported in Figure 5C. Furthermore, small distances in LMR could recognize two paths as identical, even though they may display different structural features (around 1.0 Å^2 in the MSD space).

In summary, although LMR-MDS does not produce an identical representation to that of Cartesian-MDS, it can convey the same information in terms of pathway topology. This implies that the raw information from the mechanical response contains relevant details of the structural rearrangement acting on the system. This could be of paramount importance for drug design. By reverting this procedure, one can postprocess the raw information contained in the work profiles so as to highlight different classes of unbinding pathways, which should point to different structural features.

In fact, in contrast with the alanine dipeptide case study, ligand–protein systems usually show a complex and dynamical network of interactions. This is because the atoms involved in the exiting pathway change along the route. As such, the structural clustering of the ligand alone may not provide enough information about the unbinding mechanism (see Supporting Information, Figure S5 for a Cartesian-MDS for ligand–protein unbinding case). Therefore, we suggest that LMR (an atom-independent observable) could be exploited to point to structural differences between trajectories for complex undocking studies.

(*R*)-roscovitine/CDK5 Complex. Our protocol was then applied to the cyclin-dependent kinase 5 (CDK5) in complex with the inhibitor (*R*)-roscovitine. CDK5 is an important kinase protein that regulates various processes in developing adult neurons. It is associated with several neural functions.[55−58] Increased CDK5 activity has been implicated in certain neural disorders,[59−61] including Alzheimer's disease.[59] For these reasons, we chose the (*R*)-roscovitine/CDK5 complex as a pharmaceutically relevant system for the application of our postprocessing technique.

Fifty steered MD runs were performed, using as a steering coordinate the distance between the center of mass of the CDK5 binding site (the backbone heavy atoms of the residue represented in cyan in Figure 6) and the center of mass of the ligand (see Methods Section and Figure 6). The data set of 50 work profiles was then postprocessed using the LMR-MDS analysis described above. In the distance range 8−22 Å, a window size Δ of 1.5 Å and a shift factor of 0.5 Å were used to generate the matrices for MDS. As with previous calculations, a monodimensional MDS was sufficient to obtain an acceptable error in the reduced representation. In Figure 7, we show the MDS position of the points as a function of the steering coordinate. All the members in Figure 7 are color-coded according to the difference of calculated work in the considered interval Δ (see the caption of Figure 7 for further details). In particular, all the trajectories
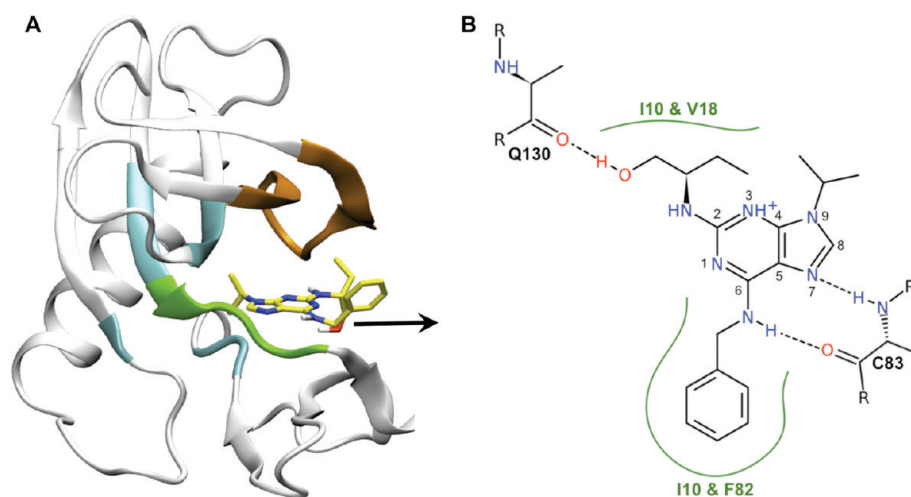
Figure 6. Three-dimensional structure of the CDK5 binding pocket bound with (R)-roscovitine inhibitor. (A) The hinge region is highlighted in green, while the glycine-rich binding loop is highlighted in orange. (R)-roscovitine (stick representation) is highlighted in a C-yellow, O-red, N-blue, and H-white color scheme. The part of the binding pocket highlighted in cyan comprises all the atoms used in defining the pulling variable. The black arrow approximately indicates the steering direction. A 2D sketch of the interactions occurring between (R)-roscovitine and CDK5 at the initial step of the steering procedure is also reported (B) and was produced using PoseView.[65]
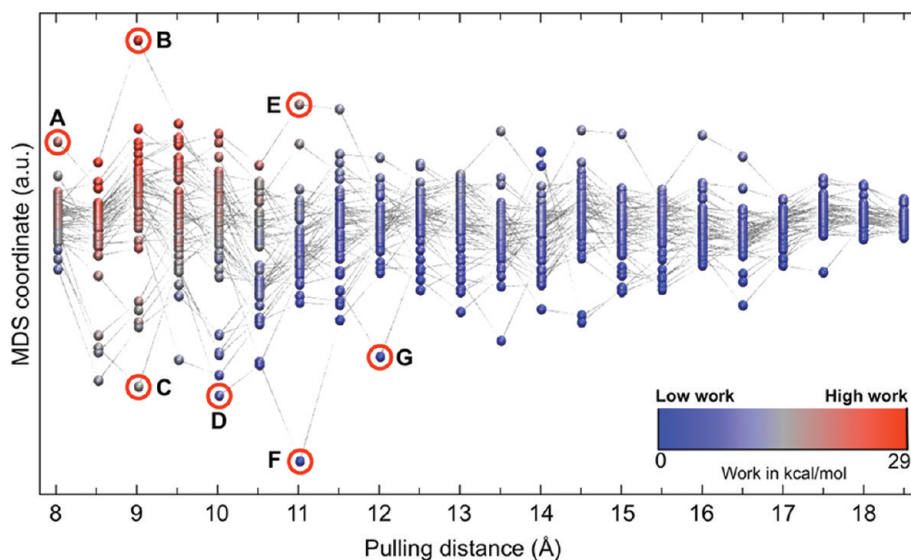

Figure 7. MDS representation of the LMR plotted against the steering coordinate obtained from the (R)-roscovitine/CDK5 system. The work curves are shown according to a BWR color scheme, where the color of each MDS point reflects the work difference within Δ size interval. Labels refer to particular points corresponding to frames selected along the steering variable, which have been analyzed further.

seemed to follow a similar stream, and the work exerted to pull the ligand out of the pocket was primarily spent in the early stages of the unbinding process (red dots at low pulling distance in Figure 7).

In the initial structure of the complex, the NH group of Cys83 donated a H-bond to the purine nitrogen in position seven of the ligand, whereas nitrogen of benzylamino group acted as an H-bond donor to the main chain oxygen of Cys83 (Supporting Information, Figure S2 and Figures 6B and 8A). The chiral hydroxyethyl substituent of (R)-roscovitine interacted with Gln130 via H-bond, while the ethyl group engaged two hydrophobic interactions with Ile10 and Val18. The bulky benzyl substituent protruded into a hydrophobic pocket lined by Ile10 and Phe82 and toward the bulk of the solvent.[41]

During the steered MD simulations, we first observed that the interactions of the ligand with Cys83 could be lost by rotating the purine ring toward the glycine-rich loop so as to disrupt the H-bond with the protein backbone in the initial stage of unbinding. By color coding the ligand configuration according to the LMR values (see Figure 8C), we realized that this particular rotation was associated with a lower LMR, as compared to those trajectories that retained these tight interactions in the early unbinding stage. This rupture was compensated for by the formation of H-bonds between the hydroxyethylamine group of (R)-roscovitine and the backbone oxygens of different residues (i.e., Ile10, Glu12, Gln130, and Asp86), depending on the generated unbinding trajectories. Moreover, the oxygen of the hydroxyethyl group was found to H-bond with Lys89, Gln85, and Gln130.
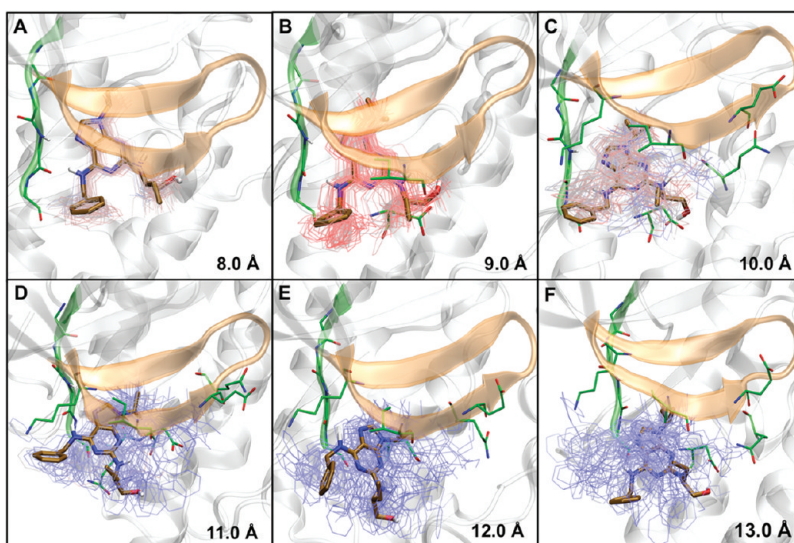
Figure 8. Representative snapshot of the exit route of the ligand from the binding pocket. The ligands represented in stick correspond to the central pose along the MDS coordinate in Figure 7. The thin lines represent the heavy atoms of the ligand, colored according the LMR measured and displayed in Figure 7.

Here too, the residues that interacted with roscovitine depended on the unbinding trajectories being investigated.

Conversely, we detected higher LMR values at the initial stage whenever the interactions with Cys83 were retained for a longer time. In this case, the ligand was found not to rotate but to favor a closer interaction with the hinge region. Here, the hydrogen donor of the benzylamine group interacted with Asp84, while charged nitrogen of Lys20 could form a cation−π interaction with the benzyl substituent of (R)-roscovitine. Additionally, the phenyl ring of the Phe82 was found to occasionally form T-stacking interactions with the benzyl substituent. On the other side, the H-bond interaction, between the hydroxyethyl group and the backbone oxygen of Gln130, was lost and replaced by new temporary polar interactions with the side chain of Asp86 (see Figure 8C) or backbone oxygen of Ile10 (depending on the trajectory). By coloring the ligand according to the LMRs values and aligning the configurations corresponding to a given distance, we produced an intuitive description of the energetics involved in the various interactions between the ligand and the target (Supporting Information, Figure S6 for an enhanced version of the picture).

In a later stage of the unbinding process, while the ligand approached the solvent, the amount of work decreased (blue dots at high pulling distance in Figure 7 and blue lines for conformations in Figure 8E and F) due to unspecific solvent interactions.

Then, we analyzed in depth the outlier configurations, which were labeled with red circles in Figure 7. These corresponded to values of extremely low or high LMR with respect to the central stream of the trajectories. According to our hypothesis, these outliers in the LMR space should correspond to peculiar conformations, showing specific (de)stabilizing contributions to the unbinding process that would otherwise be overlooked. As expected, most of the structurally relevant outliers were observed at small distance values, when the ligand started to move out of the protein entrance.

At a distance of 8.00 Å, the point labeled with A (see Figures 7 and 9) was characterized by a significantly well-directed H-bond pattern with Cys83. This is in agreement with the fact that removing such a well-directed H-bond network has a large energetic cost. At a subsequent stage, the formation of an H-bond with Cys83 and the coincidental interaction of hydroxyethylamine with Asp86 were detected (point B in Figure 7 and Figure 9). This seemed plausible since the formation of this interaction induced a rotation of the ligand that could more tightly interact with the pocket, thus preventing the unbinding from taking place. Conversely, a relevant decrease of LMR seemed to be associated with a conformational rearrangement in the hinge region of the protein (see point C in Figures 7 and 9). This could be ascribed to steering induced protein deformation, which facilitated the ligand unbinding. This highlights an additional feature of our postprocessing approach, namely the possibility of detecting the mechanical response of both the protein and the ligand, thus highlighting the induced rearrangement that occurs upon ligand unbinding.

To prove this, we repeated an identical simulation by applying an RMSD harmonic restraint on those residues involved in such rearrangements on the protein backbone. In this way, the decrease of LMR was not observed anymore, and the trajectory displayed a profile very similar to the one observed in the majority of cases (Supporting Information, Figure S7).

Similarly, the point denoted with D was characterized by a remarkably low LMR with respect to the majority of the trajectories. This seemed to be connected to the favorable pattern of release of the ligand, which lost the interaction with Cys83 and favored the interaction with Asp86 toward the glycine-rich loop.

Proceeding further, at a distance of 11 Å, two contrasting binding patterns emerged (see points E and F in Figures 7 and 9). These were characterized not only by a distinct pattern in the ligand−protein interactions but also by a sizable and opposite rearrangement of the hinge region of the protein.

Finally, while moving toward the solvent, several residual possible interactions were still detectable and sizable (see point G in Figures 7 and 9).

From a drug design perspective, it is potentially important to identify the outliers using local mechanical response analysis.
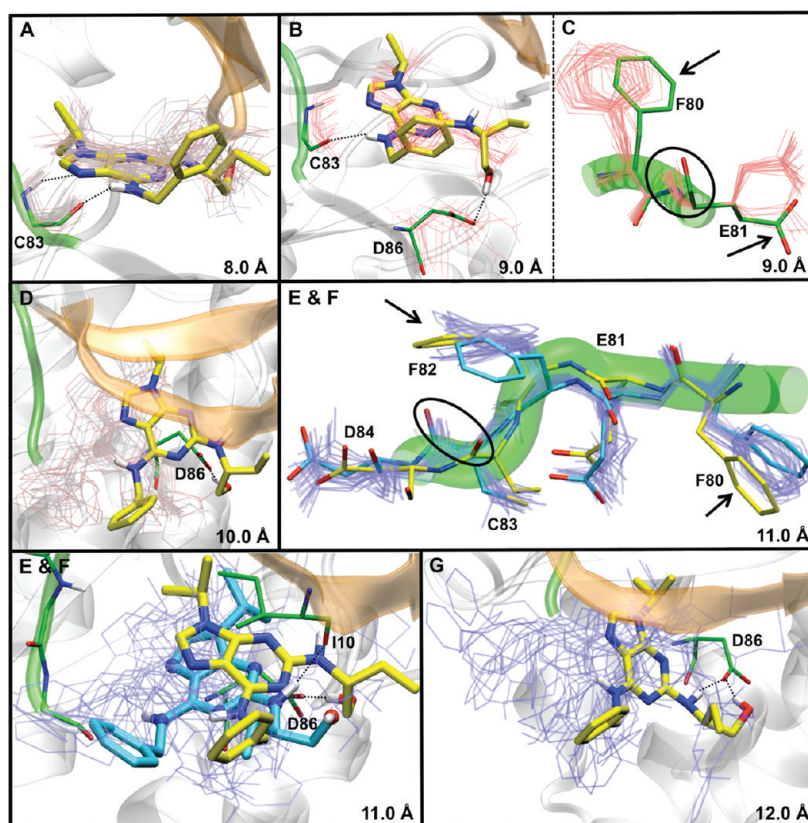
Figure 9. Structural features of the configurations corresponding to the outliers in LMR. The structures highlighted with letters correspond to the circles in Figure 7.

This is because it allows for: (i) the identification of the most relevant structural features responsible for ligand recognition and binding; and (ii) the identification of the key interactions that can be exploited to address chemical modifications aimed at improving the ligand affinity for both its biological counterpart and residence time.[11] In practice, this could be achieved by selectively suppressing the low LMR pathways and enhancing the high LMR pathways through specific modifications on the scaffold of the drug.

A crucial aspect of this work is that MDS representations of LMR allow us to group together trajectories that display an analogous mechanical response along the unbinding pathway. This is of great help in focusing the structural analysis on a limited number of representative configurations, which can be crucial when large statistics of trajectories are produced. Additionally, the analysis of outliers in LMR can further strengthen the hypothesis of specific unbinding mechanisms, thus saving a large amount of work in terms of human time. Although the robustness of the results can be improved by increasing the statistics, we have here shown that even a limited statistics of unbinding events can lead to important structural insights.

Moreover, while for this case the use of one single MSD dimension was sufficient, additional dimensions may provide a further source of flexibility helping to track different exit routes displaying a different mechanical signature.

■ CONCLUSIONS

In this work, we developed a postprocessing tool for out-of-equilibrium steered MD simulations to identify distinctive

structural features of particular relevance for a specific biological event. Steered MD is nowadays becoming very attractive. Many MD programs now include native routines to perform it.[36,62] It can also be made available through external plugins.[38] One of the major limitations to a wide application of steered MD for proper free energy difference calculations is related to the applicability of the Jarzynski equation, which requires huge computational resources for realistic cases. This is prohibitive in the drug design field, where a major trade-off between speed and accuracy is usually required. We have here shown that very useful structural information can be retrieved from a relatively small statistics of pulling trajectories obtained in a non-quasistatic regime. In particular, in the alanine dipeptide case study, the structural features observed and the topological differences in paths have a clear correspondence in the mechanical response of the system to the external pulling force. The conformational transition of the alanine dipeptide is relevant because of the following points: (i) alanine dipeptide can exist in two well-defined energy minima; (ii) three different paths connect the two minima; (iii) the transitions among minima are usually affected by a very low dissipative work; and (iv) the representation of the conformational transition in the local mechanical space can be straightly converted into a transition in the Cartesian space. Remarkably, while points (i) and (ii) also apply directly to a ligand—protein unbinding, (iii) and (iv) do not. In fact, the ligand unbinding from a protein via steered MD experiences a non-negligible dissipative work, and local mechanical space cannot be directly transformed into a Cartesian space. This is because the protein atoms involved in the unbinding change along the reactive path. Despite this, local

mechanical response can still be very informative from a structural standpoint, as we have here demonstrated by investigating biological systems of increasing complexity. Our second case study was the unbinding of (R)-roscovitine from CDK5. Here, the steering coordinate was intentionally chosen to be of general applicability, with the consequent increase in the dissipative work. Nevertheless, even in this regime, many structural details could be obtained. Additionally, from a drug design standpoint, such an approach could also be used in discriminating the understanding of the energetic relationships among alternative docking poses. By adding the nontrivial step of the mechanical response pattern analysis, it could be possible to discriminate those poses that, bringing completely different pathways, are structurally and energetically more or less stable than others.

In conclusion, we have shown that the analysis of the local mechanical response of the system to a forced unbinding can be very informative with respect to the unbinding process itself. Moreover, from a drug design standpoint, it captures the relevant structural events that can be directly exploited to design novel ligands with a potentially increased affinity for and residence time at the biological counterpart. In addition, a large amount of work in terms of human time is saved. This is because researchers need to analyze only the most promising regions where structural dissimilarities are evidenced by MDS. Although the quality of the results can be improved by increasing the statistics, we have shown here that even limited statistics can lead to important insights, which complement a more computationally demanding accurate free energy calculation. In our opinion, this is remarkably important because, as more computer power becomes available and more extended MD studies become possible,[63] new analysis techniques[29-31] as well as new data selection strategies must be devised to optimize storage usage and human effort by directing the action to a smaller but more relevant set of data drawn from a consistent statistics. Additionally, such approach is not limited to ligand—protein binding problems but can be directly applied in simulations of single-molecule force spectroscopy experiments.[64]

■ ASSOCIATED CONTENT

ⓈSupporting Information. Selection of spring constant and pulling velocity for alanine dipeptide; representative sketch of the residues of CDK5 involved in the binding process; Cartesian MDS representation of drug unbinding trajectories; selection of the pulling length in undocking of (R)-rosocovitine/CDK5; effect of steering on the structure of the CDK5; enhanced picture for LMR representation, and steering experiment on (R)-roscovitine/CDK5 system with constrained residues. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: davide.branduardi@gmail.com; andrea.cavalli@iit.it.

■ ACKNOWLEDGMENT

■ REFERENCES

(1) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J. Chem. Phys.* **1988**, *89*, 3742–3746.

(2) Gilson, M.; Given, J.; Bush, B.; McCammon, J. The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* **1997**, *72*, 1047–1069.

(3) Hamelberg, D.; McCammon, J. A. Standard free energy of releasing a localized water molecule from the binding pockets of proteins: double-decoupling method. *J. Am. Chem. Soc.* **2004**, *126*, 7683–7689.

(4) Deng, Y.; Roux, B. Calculation of standard binding free energies: aromatic molecules in the T4 Lysozyme L99A mutant. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.

(5) Wang, J.; Deng, Y.; Roux, B. Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* **2006**, *91*, 2798–2814.

(6) Boyce, S. E.; Mobley, D. L.; Rocklin, G. J.; Graves, A. P.; Dill, K. A.; Shoichet, B. K. Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. *J. Mol. Biol.* **2009**, *394*, 747–763.

(7) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **2007**, *371*, 1118–1134.

(8) Mobley, D. L.; Chodera, J. D.; Dill, K. A. On the use of orientational restraints and symmetry corrections in alchemical free energy calculations. *J. Chem. Phys.* **2006**, *125*, 084902–084916.

(9) Mobley, D. L.; Chodera, J. D.; Dill, K. A. Confine-and-release method: obtaining correct binding free energies in the presence of protein conformational change. *J. Chem. Theory Comput.* **2007**, *3*, 1231–1235.

(10) 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.

(11) Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discovery* **2006**, *5*, 730–739.

(12) Swinney, D. C. Biochemical mechanisms of drug action: what does it take for success? *Nat. Rev. Drug Discovery* **2004**, *3*, 801–808.

(13) Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10184–10189.

(14) Bui, J. M.; Henchman, R. H.; McCammon, J. A. The dynamics of ligand barrier crossing inside the acetylcholinesterase gorge. *Biophys. J.* **2003**, *85*, 2267–2272.

(15) Roux, B. The calculation of the potential of mean force using computer-simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282.

(16) Laio, A.; Parrinello, M. Escaping Free Energy Minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *20*, 12562–12566.

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

(18) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120*, 5946–5961.

(19) Masetti, M.; Cavalli, A.; Recanatini, M.; Gervasio, F. L. Exploring complex protein—ligand recognition mechanisms with coarse metadynamics. *J. Phys. Chem. B* **2009**, *113*, 4807–4816.

(20) Bayas, M. V.; Schulten, K.; Leckband, D. Forced detachment of the CD2-CD58 complex. *Biophys. J.* **2003**, *84*, 2223–2233.

(21) Gao, M.; Wilmanns, M.; Schulten, K. Steered molecular dynamics studies of titin I1 domain unfolding. *Biophys. J.* **2002**, *83*, 3435–3445.

(22) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **1997**, *72*, 1568–1581.

(23) Krammer, A.; Lu, H.; Isralewitz, B.; Schulten, K.; Vogel, V. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1351–1356.

(24) 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.

(25) Chen, L. Y. Exploring the free-energy landscapes of biological systems with steered molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6176–6183.

(26) Ozer, G.; Valeev, E. F.; Quirk, S.; Hernandez, R. Adaptive steered molecular dynamics of the long-distance unfolding of neuropeptide Y. *J. Chem. Theory Comput.* **2010**, *6*, 3026–3038.

(27) Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690.

(28) Xiong, H.; Crespo, A.; Marti, M.; Estrin, D.; Roitberg, A. E. Free energy calculations with non-equilibrium methods: applications of the Jarzynski relationship. *Theor. Chem. Acc.* **2006**, *116*, 338–346.

(29) Rajan, A.; Freddolino, P. L.; Schulten, K. Going beyond clustering in MD trajectory analysis: an application to villin headpiece folding. *PLoS ONE* **2010**, *5*, e9890.

(30) Seeber, M.; Cecchini, M.; Rao, F.; Settanni, G.; Caflisch, A. Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics* **2007**, *23*, 2625–2627.

(31) Wriggers, W.; Stafford, K. A.; Shan, Y.; Piana, S.; Maragakis, P.; Lindorff-Larsen, K.; Miller, P. J.; Gullingsrud, J.; Rendleman, C. A.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. Automated event detection and activity monitoring in long molecular dynamics simulations. *J. Chem. Theory Comput.* **2009**, *5*, 2595–2605.

(32) Borg, I.; Groenen, P. J. F. *Modern Multidimensional Scaling: theory and applications*; 2nd ed.; Springer-Verlag: New York, 2005.

(33) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.

(34) Crooks, G. E. Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *J. Stat. Phys.* **1998**, *90*, 1481–1487.

(35) Kearsley, S. K. On the orthogonal transformation used for structural comparison. *Acta Cryst. A* **1989**, *45*, 208–210.

(36) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(37) Paterlini, M. G.; Ferguson, D. M. Constant temperature simulations using the Langevin equation with velocity Verlet integration. *Chem. Phys.* **1998**, *236*, 243–252.

(38) 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.

(39) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(40) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(41) Mapelli, M.; Massimiliano, L.; Crovace, C.; Seeliger, M. A.; Tsai, L. H.; Meijer, L.; Musacchio, A. Mechanism of CDK5/p25 binding by CDK inhibitors. *J. Med. Chem.* **2005**, *48*, 671–679.

(42) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65*, 712–725.

(43) 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.

(44) 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.

(45) 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.

(46) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

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

(48) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(49) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. String method in collective variables: minimum free energy paths and isocommittor surfaces. *J. Chem. Phys.* **2006**, *125*, 24106.

(50) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **2007**, *126*, 054103.

(51) Bolhuis, P. G.; Dellago, C.; Chandler, D. Reaction coordinates of biomolecular isomerization. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877–5882.

(52) Apostolakis, J.; Ferrara, P.; Caflisch, A. Calculation of conformational transitions and barriers in solvated systems: application to the alanine dipeptide in water. *J. Chem. Phys.* **1999**, *110*, 2099–2108.

(53) Ferrara, P.; Apostolakis, J.; Caflisch, A. Targeted molecular dynamics simulations of protein unfolding. *J. Phys. Chem. B* **2000**, *104*, 4511–4518.

(54) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **1963**, *7*, 95–99.

(55) Bibb, J. A. Role of Cdk5 in neuronal signaling, plasticity, and drug abuse. *Neurosignals* **2003**, *12*, 191–199.

(56) Cheng, K.; Ip, N. Y. Cdk5: a new player at synapses. *Neurosignals* **2003**, *12*, 180–190.

(57) Dhavan, R.; Tsai, L. H. A decade of CDK5. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 749–759.

(58) Gupta, A.; Tsai, L. H. Cyclin-dependent kinase 5 and neuronal migration in the neocortex. *Neurosignals* **2003**, *12*, 173–179.

(59) Lau, L. F.; Ahlijanian, M. K. Role of cdk5 in the pathogenesis of Alzheimer's disease. *Neurosignals* **2003**, *12*, 209–214.

(60) Nguyen, M. D.; Julien, J. P. Cyclin-dependent kinase 5 in amyotrophic lateral sclerosis. *Neurosignals* **2003**, *12*, 215–220.

(61) Smith, P. D.; Crocker, S. J.; Jackson-Lewis, V.; Jordan-Sciutto, K. L.; Hayley, S.; Mount, M. P.; O'Hare, M. J.; Callaghan, S.; Slack, R. S.; Przedborski, S.; Anisman, H.; Park, D. S. Cyclin-dependent kinase 5 is a mediator of dopaminergic neuron loss in a mouse model of Parkinson's disease. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13650–13655.

(62) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(63) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. Atomic-level characterization of the structural dynamics of proteins. *Science* **2010**, *330*, 341–346.

(64) Irback, A.; Mitternacht, S.; Mohanty, S. Dissecting the mechanical unfolding of ubiquitin. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13427–13432.

(65) Stierand, K.; Maab, P. C.; Rarey, M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* **2006**, *22*, 1710–1716.