

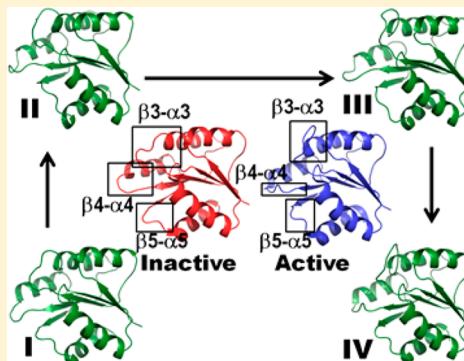
Conformational Transition of Response Regulator RR468 in a Two-Component System Signal Transduction Process

Rahul Banerjee,[†] Honggao Yan,^{*,‡} and Robert I. Cukier^{*,†}

[†]Department of Chemistry and [‡]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States

Supporting Information

ABSTRACT: Signal transduction can be accomplished via a two-component system (TCS) consisting of a histidine kinase (HK) and a response regulator (RR). In this work, we simulate the response regulator RR468 from *Thermotoga maritima*, in which phosphorylation and dephosphorylation of a conserved aspartate residue acts as a switch via a large conformational change concentrated in three proximal loops. A detailed view of the conformational transition is obscured by the lack of stability of the intermediate states, which are difficult to detect using common structural biology techniques. Molecular dynamics (MD) trajectories of the inactive and active conformations were run, and show that the inactive (or active) trajectories do not exhibit sampling of the active (or inactive) conformations on this time scale. Targeted MD (TMD) was used to generate trajectories that span the inactive and active conformations and provide a view of how a localized event like phosphorylation can lead to conformational changes elsewhere in the protein, especially in the three proximal loops. The TMD trajectories are clustered to identify stages along the transition path. Residue interaction networks are identified that point to key residues having to rearrange in the process of transition. These are identified using both hydrogen bond analysis and residue interaction strength measurements. Potentials of mean force are generated for key residue rearrangements to ascertain their free energy barriers. We introduce methods that attempt to extrapolate from one conformation to the other and find that the most fluctuating proximal loop can transit part way from one to the other, suggesting that this conformational information is embedded in the sequence.



■ INTRODUCTION

Two-component system (TCS) signal transduction is the predominant mechanism for bacteria to sense environmental conditions such as temperature, nutrients, and osmopressure.^{1,2} The critical role of TCS signal transduction in bacteria for their virulence and pathogenesis and the absence of this mode of signal transduction in higher organisms make TCSs an important target for future therapeutic interventions to treat bacterial infections.³ Integral to TCS pathways is a phosphotransfer reaction between a membrane bound histidine kinase (HK) and a cognate response regulator (RR) protein. A prototypical RR contains two domains, a receiver domain (REC) that hosts a phosphoacceptor aspartate residue and an effector or output domain responsible for downstream action.^{4,5} In RR468, the system of our interest, the effector domain is absent and the receiver domain itself functions as an effector domain.⁶ RR468 in *Thermotoga maritima* is involved in relaying a phosphoryl group from its cognate pair histidine kinase, HK853, in response to extracellular stimuli.^{6,7} The activated TCS is characterized by phosphorylated RR468 where the phosphoryl group is covalently bound to a conserved aspartate (D53) residue.

In spite of extensive literature focused on biochemical and mutational studies of TCS signaling, there was almost no structural basis to understand the mechanism until 2009 when

Casino and co-workers published a series of high resolution crystal structures of HK853, RR468, and a HK853-RR468 complex.⁶ The conformations of RR468 for free and BeF₃⁻-bound forms highlighted the structural difference between active (phosphorylated) and inactive (dephosphorylated) conformations of this protein. Beryllium trifluoride, covalently attached to an aspartate residue, mimics the phosphorylated D53 at the site of phosphorylation in phosphorylated RR468 (P~RR468).^{6,8} Because the conformations of the two end states (active and inactive) are known for RR468, an excellent model system is now available to computationally address some key questions related to the conformational transition in RR468 in the presence or absence of phosphorylation. The underlying hypothesis being that a simulation of the active form, without the phosphoryl group or metal ion, should enable the system to adopt the conformation of the inactive form. We also wanted to understand the energetic contributions associated with the event of phosphorylation that produce a conformational change by lowering energetic barriers associated with various local rearrangements.

Received: December 16, 2013

Revised: April 11, 2014

Published: April 14, 2014

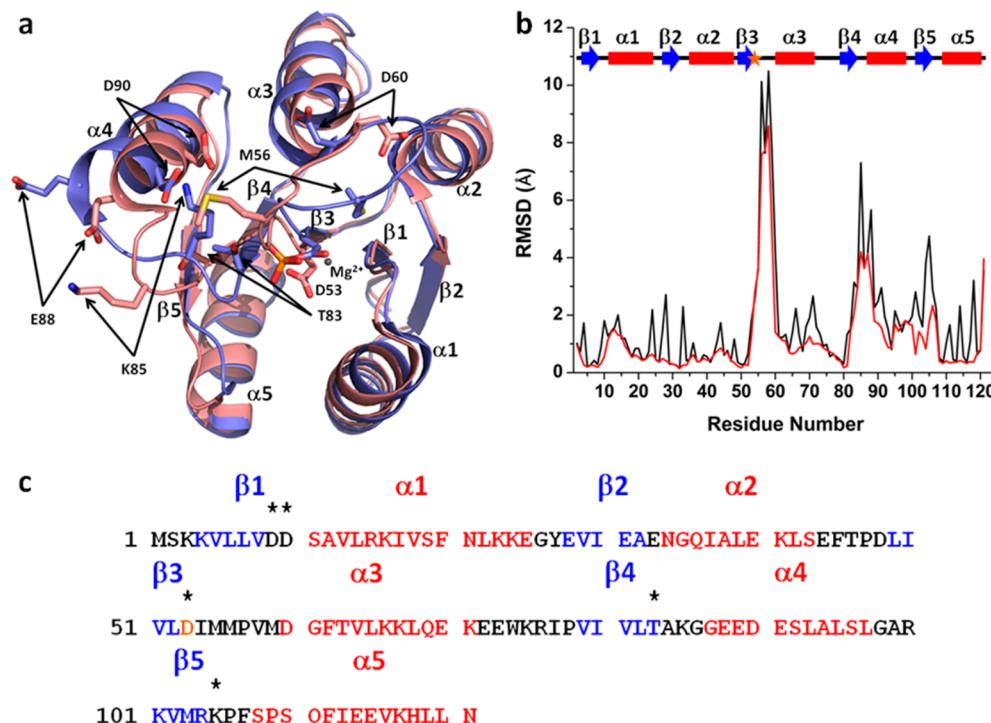


Figure 1. Comparison between the inactive (PDB id 3DGF in pink) and active conformations of RR468 (PDB id 3GL9 in blue) along with the residues that differ most significantly in conformation, orientation, and interaction in the active and the inactive form of RR468. (b) RMSD between active and inactive form crystal structures of all heavy atoms (black) and backbone heavy atoms (red), and (c) the sequence of RR468. The α -helices and the β -strands are shown in red and blue, respectively. The site of phosphorylation is D53, which is the last residue of β 3 (shown in orange in (b) and (c)). In (c), the residue positions marked with '*' are highly conserved residues in related response regulator proteins.

The active and inactive conformations of RR468 differ most significantly in the region adjacent to the site of phosphorylation. We will refer to it as the proximal region (Figure 1a). There are five loops toward the proximal side. Three proximal loops surround the site of phosphorylation closely: β 3- α 3, β 4- α 4, and β 5- α 5 loops. The difference between active and inactive conformations is most conspicuous in the β 3- α 3 loop (Figure 1 b). This is the longest loop in the proximal end (residues 54 to 59) and is closest to the site of phosphorylation, D53 (Figure 1b,c). In the HK853-RR468 complex the β 3- α 3 loop of RR468 has contacts with two regions of HK853: the α 4- β 4 loop and the ATP-lid region of the CA domain.⁶ Due to these contacts, D53 of RR468 is lined up with H260 of HK853 in the complex, favoring the phosphoryl transfer reaction. Residue M56, in the middle of the β 3- α 3 loop, is known to be crucial for complex formation in this and several other HK-RR cognate pairs.⁶ M56 is buried inside a hydrophobic pocket in RR468 consisting of V8, L52, and V64 in the active conformation whereas it is solvent exposed in the inactive conformation. D60 is the first residue in the α 3 helix immediately following the β 3- α 3 loop. It interacts with the N-termini of the α 3 and α 2 helices in the active and inactive conformations, respectively (Figure 1).

The RR468 β 4- α 4 loop interacts with DHp-CA linker residues in the HK853-RR468 complex crystal structure. Highly conserved T83 is the last residue the β 4 strand. This residue is known to be important for neutralizing or shielding the negative charge on the phosphoryl group and assisting phosphoryl transfer to a water molecule.⁸ Hence, in the active conformation, the T83 side chain flips to an inward orientation to interact with the phosphoryl group on P~D53, where it is outward toward the exterior of the protein in the inactive conformation.

The last, crucial RR468 loop β 5- α 5 clamps the α 1-helix from the DHp domain of HK853 with the α 1 helix of RR468 in the HK853-RR468 complex. K105 is the first residue in this loop. In the active conformation, K105 interacts with the highly conserved D9 (β 1) and the phosphoryl group on D53 (or phosphoryl group mimic BeF_3^-). The side chain of K105 has two alternative conformations in the crystal structure of the inactive form: one outward and another inward relative to the site of phosphorylation. In the inward conformation, the K105 side chain fills the void space of the phosphoryl group, hence close to the D53 side chain, and the salt bridge between K105 and D9 is missing in this conformation.

In this work, structural determinants of the active and inactive states and the transition between them are proposed using, respectively, equilibrium molecular dynamics (MD) and Targeted MD (TMD) simulations. TMD has proved to be a useful tool to identify high-energy intermediate conformational states along a transition path that are difficult to identify using common structural biology techniques.⁹ MD simulations together with TMD simulations were shown to be very useful to investigate the molecular basis for the conformational transition in various response regulator proteins.^{8,10-14} In this work, how a local event like the phosphorylation or dephosphorylation of a particular amino acid cascades down to other local conformational rearrangements, producing a global conformational change in RR468, is investigated by analyzing the MD simulations of the inactive and active forms, and the TMD simulations connecting them, using protein structure network (PSN)¹⁵ and hydrogen bond analyses. NMR experiments suggest that the conformational transition in RR468 takes place on μ s to ms time scale (unpublished data). This necessitates the introduction of TMD to span the

Table 1. Summary of MD and TMD Simulations along with the Acronyms Used in This Paper^a

acronym	starting conformation	target conformation	method	length (ns)	number of simulations
IAC	inactive conformation	-	MD	100	3
P~RR468	active conformation with P~D53 and Mg ²⁺	-	MD	100	1
ACT	active conformation without PO ₃ ²⁻ and Mg ²⁺	-	MD	100	3
I2A	inactive conformation	active conformation	TMD	20	3
A2I	active conformation without PO ₃ ²⁻ and Mg ²⁺	inactive conformation	TMD	20	3

^aThe results and conclusions were derived from one particular trajectory for each MD and TMD simulation unless exclusively mentioned in the text.

transition. From the TMD trajectories, there are key residues, MS6, D60, and T83, as noted above, which undergo substantial conformational changes. For these residues, we use biased MD simulations, starting from various points along the transition path, to estimate their free energy barriers to rearrangement from the potentials of mean force (PMFs) along relevant reaction coordinates.

With mobile loops in the active and inactive forms, it could be that the inactive conformation would sample active conformations, and vice versa, but that overlap would not be accessible within the MD time scale. To see if information about the active conformation is encoded in the inactive sequence, and vice versa, we use principle component analysis (PCA)^{16–18} to find the major modes of motion of the inactive and active MD trajectories. Then, these modes are used to probe whether the mobile loop fluctuations of one form point in a direction to sample the conformations of the other form.

In Section 2, we present the MD and TMD simulation details, along with the hydrogen bond and PSN analysis methods, and the construction of the free energy (potentials of mean force) for selected residue conformational transitions. Two methods for using PCA modes to investigate the encoding behavior of inactive and active forms are presented. Section 3 gives the results of the hydrogen bond and PSN analyses of the MD and TMD trajectories, and the PCA-based analysis. The PMFs of the critical residues for the transition between protein forms are shown. Our results are discussed in Section 4 to indicate which residues are mainly responsible for the conformational transition. Section 5 presents our conclusions.

2. METHODS

2.1. MD Simulation of RR468 without Phosphoryl Group and Mg²⁺. The inactive conformation of RR468 (PDB id 3DGF) was simulated (designated as IAC) using Amber10¹⁹ in the presence of explicit water molecules (see Table 1 for a list of all simulations with their designators). The ff99SB force field was used for the simulation. Water molecules in the crystal structure were removed, hydrogen atoms were added, and the system was neutralized by addition of an appropriate number of Na⁺ ions. The system was solvated with a 12 Å layer of TIP3P waters in a rectangular water box using leap. The system was minimized first using a combination of steepest descent and conjugated gradient methods for 8000 steps. A weak positional restraint (2 kcals/mol) was applied on all protein atoms during this minimization step. The whole system was minimized again using a combination of steepest descent and conjugated gradient methods for 20 000 steps without any positional restraint on any atoms. A time step of 2 fs was used for all subsequent heating, equilibration, and production runs with the SHAKE option on all bonds containing H atoms. Langevin dynamics was used for temperature control in the heating, equilibration, and production steps. The minimized system was heated from 0 to 300 K in 1 ns. Weak positional restraints (2

kcals/mol) on all protein atoms were applied during the heating cycle. Constant pressure equilibration was done at 300 K for 2 ns without any positional restraints. Finally, a trajectory of 100 ns was generated during a production run.

Root mean square deviation (RMSD) from the crystal structure was obtained along the time course of simulation and root-mean-square fluctuation (RMSF) for backbone heavy atoms for each residue was obtained using the Ambertools 10 ptraj module.²⁰ All individual frames were aligned on the first frame along the C α atoms, excluding the three mobile loop regions β 3- α 3, β 4- α 4, and β 5- α 5.

The active conformation of RR468 was simulated for 100 ns without phosphate (or BeF³⁻) and Mg²⁺ ion to investigate whether or not a conformational transition from active to inactive conformation takes place during this time scale. The explicit solvent simulation was done in the same way as outlined above.

2.2. MD Simulation of Active form RR468 (P~RR468).

The crystal structure of RR468 (3GL9) in its active conformation has a phosphoryl group mimic, BeF₃⁻, covalently attached to the aspartate (D53) residue. An Mg²⁺ ion is present in the active site. We replaced BeF₃⁻ with phosphoryl group (PO₃²⁻). An ab initio quantum chemical calculation was performed on a phosphorylated aspartate residue using density functional theory, with the B3LYP functional and 6-31G* basis set in Gaussian 2003.²¹ A system composed of phosphorylated aspartate capped with Acetyl (Ace) and N-methyl (Nme) groups at the N and C-terminals, respectively, was used for this calculation (Figure S1, Supporting Information). The Mg²⁺ ion, three water molecules, and an acetate ion were considered as a part of the system, resulting in a net charge -1 for this calculation. One of these water molecules mimics the backbone carbonyl group of M55, and the acetate ion mimics D10, which interacts with the Mg²⁺ ion, as in the crystal structure. The Gaussian input structure was built based on all heavy atom positions in the crystal structure (3GL9). MD charges on each atom in the phosphorylated aspartate residue were obtained by processing the Gaussian output using Antechamber.²² The other force field parameters for phosphorylated aspartate were obtained using the parmchk program in the Amber10 suite. Explicit solvent simulations were done using the same protocol as described above for the simulation of the inactive conformation.

2.3. Targeted Molecular Dynamics Simulations. To simulate a transition pathway between the inactive to active (I2A) structures, targeted molecular dynamics (TMD) was performed, starting from the equilibrated structure in the simulation of the inactive conformation. The target reference structure was the crystal structure conformation of the active form of the protein (PDB id 3GL9). In this method, a harmonic restraint in the mass-weighted RMSD between the initial and target structures is introduced that can, over a suitably slow transformation, provide a reasonable set of

intermediate conformations between the initial and target structures. All C α atoms were best fit on the target reference structure and the RMSD was calculated for all heavy atoms in the protein. A force constant of 8 kcal/mol was used for the TMD. A similar transition pathway along the active to inactive (A2I) conformation was simulated starting from the equilibrated structure in the simulation of the active conformation. The target structure was the crystal structure (PDB id 3DGF) conformation of the inactive form of the protein in this case. The simulation length was 20 ns for each TMD trajectories. Two additional sets of 20 ns independent transition pathways were simulated for I2A and A2I to reach a consensus among the transition path. Slightly different conformations, after equilibration in the IAC and ACT simulations, were used as starting structures in these TMD simulations.

2.4. Principal Component Analysis (PCA). Methods based on principal component analysis (PCA) were used to investigate conformational transitions of the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops in RR486 in IAC and ACT simulations. PCA diagonalizes the covariance matrix,

$$\mathbf{C} = \frac{1}{T} \int_0^T \delta\mathbf{R}(t) \delta\mathbf{R}^T(t) dt = \left\langle \delta\mathbf{R}(t) \delta\mathbf{R}^T(t) \right\rangle \quad (1)$$

of the atom fluctuations from their trajectory average values, $\delta\mathbf{R}(t) = \mathbf{R}(t) - \langle \mathbf{R}(t) \rangle$.

Here, $\mathbf{R}(t) = (x_1(t), y_1(t), \dots, z_N(t))^T$ denotes configurations along the trajectory and $\langle \dots \rangle$ denotes a time average over trajectory snapshots. The trajectory fluctuations are expressed as

$$\delta\mathbf{R}(t) = \sum_{\nu=1}^{3N} [\delta\mathbf{R}(t) \bullet \mathbf{m}_\nu^T] \mathbf{m}_\nu \equiv \sum_{\nu=1}^{3N} \delta p_\nu(t) \mathbf{m}_\nu \quad (2)$$

with the \mathbf{m}_ν ($\nu = 1, 2, \dots, 3N$) the (orthonormal) eigenvectors and λ_ν^2 the corresponding eigenvalues of the covariance matrix \mathbf{C} . The principal component $\delta p_\nu(t)$ is the projection of the trajectory onto the ν th eigenvector. In the rotated Cartesian coordinate basis defined by the \mathbf{m}_ν the largest eigenvalue λ_1^2 captures the largest fraction of the mean square fluctuation (MSF), the second largest the next largest fraction of the MSF, and so forth. In applying PCA, before diagonalization of the covariance matrix, the overall translational and rotational motion of the protein has to be removed. We do so in the present context by superimposing all trajectory snapshots onto the C α atoms of the crystal structure, except for those in $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops. That helps to isolate the fluctuations of the least stable parts of RR486.

PCA carried out on protein MD trajectories often show that their fluctuations correspond to many small amplitude, one-basin oscillatory modes along with a small number of large fluctuation and, typically, nonharmonic modes. If this separation held for the inactive protein and the amplitudes of all the modes were artificially increased, then it would be a remote possibility that active conformations would result or vice versa. Note that for PCA (along with normal mode analysis) only directions of motion are found, their amplitudes are not determined. On the other hand, if it is true that active conformations are encoded in the inactive fluctuations, but they are not seen in the MD trajectory because the trajectory is not run for a sufficiently long time, then it may be that some set of the large "productive" modes from the inactive PCA will point toward active conformations. If these properties are manifest in the trajectory then a small subspace contained in the covariance

matrix \mathbf{C} characterizes the important inactive motions and that subspace is the relevant one for active-directed motion. Based on these considerations, we present two methods that are useful for investigating if the fluctuations around one structural basin can point to the other structure's conformational basin.

2.5. DEVIATION Method. To explore the possibility that the IAC trajectory that is started from the inactive form crystal structure (here denoted as initial) encodes information about the active conformation (here denoted as final), we use a previously developed method²³ that projects the final structure onto the PCA modes of the initial structure. Thus, express the difference $\Delta\mathbf{X}$ between final and initial conformations in the initial state MD trajectory PCA basis:

$$\Delta\mathbf{X} = \mathbf{R}^{\text{final}} - \mathbf{R}^{\text{initial}} = \sum_{\nu}^{3N} (\Delta\mathbf{X} \bullet \mathbf{m}_\nu^T) \mathbf{m}_\nu \equiv \sum_{\nu}^{3N} a_\nu \mathbf{m}_\nu \quad (3)$$

In this fashion, the active conformation is expressed in the inactive PCA basis. If the inactive fluctuation directions that are expressed by the \mathbf{m}_ν basis do encode the active direction, then the a_ν expansion coefficients in eq 3 should track the λ_ν coefficients of the initial trajectory PCA. Equivalently, if this encoding hypothesis is true, then the function $\Delta\mathbf{X} - \Delta\mathbf{X}^{M_\nu} \equiv \sum_{\nu=M_\nu}^{3N} a_\nu \mathbf{m}_\nu$ should also decay rapidly with ν (again, on the same scale as the λ_ν). A convenient (normalized) measure to use is

$$\begin{aligned} \text{DEV} &= (\Delta\mathbf{X} - \Delta\mathbf{X}^{M_\nu})^T \bullet (\Delta\mathbf{X} - \Delta\mathbf{X}^{M_\nu}) / (\Delta\mathbf{X})^T \bullet (\Delta\mathbf{X}) \\ &= (\Delta\mathbf{X} - \Delta\mathbf{X}^{M_\nu})^T \bullet (\Delta\mathbf{X} - \Delta\mathbf{X}^{M_\nu}) / \text{MSD} \end{aligned} \quad (4)$$

with MSD the mean square deviation.

2.6. EXTEND Method. Another approach to see if, e.g., the inactive trajectory fluctuations point in the active form direction can be formulated. Consider the set of all inactive trajectory snapshots written in the basis of M_ν PCA modes. Again, a small set of modes are used to eliminate the small unproductive fluctuations from the trajectory. For each snapshot, introduce a multiplier, α ($-\infty$ to $+\infty$) and construct the following mean square metric:

$$\begin{aligned} f^{M_\nu}(\alpha(t)) &= (\mathbf{R}^{\text{in}} + \alpha \delta\mathbf{R}^{M_\nu}(t) - \mathbf{R}^{\text{ac}}) \bullet (\mathbf{R}^{\text{in}} + \alpha \delta\mathbf{R}^{M_\nu}(t) \\ &\quad - \mathbf{R}^{\text{ac}}) \\ &= (\alpha \delta\mathbf{R}^{M_\nu}(t) - \Delta\mathbf{R}) \bullet (\alpha \delta\mathbf{R}^{M_\nu}(t) - \Delta\mathbf{R}) \end{aligned} \quad (5)$$

Scan α on some grid to measure the overlap with the active conformation, for each snapshot, and find the snapshot that minimizes $f^{M_\nu}(\alpha(t))$. Then, the snapshot that can be extended to be as close as possible to the active conformation has been generated. The EXTEND method is more objective than the DEV method because the active form is only used to monitor the extent to which the inactive form extension overlaps with the active form. There is certainly no reason for closer approach to the active conformation with the addition of more modes in $\delta\mathbf{R}^{M_\nu}(t)$ as in the DEV method. To actually accomplish this minimization, the derivative $\partial f^{M_\nu}(\alpha)/\partial \alpha = 0$ leads to

$$f^{M_\nu}(\alpha_{\min}(t)) = \Delta\mathbf{R}^2 \sin^2 \theta(t) \quad (6)$$

with

$$\cos(\theta(t)) = \frac{\Delta\mathbf{R} \bullet \delta\mathbf{R}^{M_\nu}(t)}{\|\Delta\mathbf{R}\| \|\delta\mathbf{R}^{M_\nu}(t)\|} \quad (7)$$

the cosine of the angle between the indicated vectors and

$$\alpha_{\min}(t) = \cos(\theta(t)) \frac{\|\Delta \mathbf{R}\|}{\|\delta \mathbf{R}^{M_v}(t)\|} \quad (8)$$

It is straightforward to show that the second derivative, $\partial^2 f^{M_v}(\alpha)/\partial \alpha^2|_{\alpha_{\min}} > 0$; thus, a minimum results. Then, searching all the times (snapshots) for the $t = t_{\min}$ for which $f^{M_v}(\alpha_{\min}(t))$ is minimized produces the smallest deviation from the active form. This value of α is used to generate the extended protein conformation from the inactive trajectory. It is guaranteed to be as close as possible in the mean square sense to the active conformation based on a one-parameter measure.

In a similar manner, the PCA analysis was carried out using the 100 ns MD trajectory of the active conformation (ACT) of the APO-protein and projected on the crystal structure of the inactive conformation.

2.7. PMF Construction Using WHAM. When using a window method to obtain a potential of mean force (PMF) for some reaction coordinate q , the trajectory data for the restrained (biased) windows need to be combined. The weighted histogram analysis method (WHAM)^{24,25} combines such data from the N different bias window potentials by expressing the true, unbiased estimated probability density $\rho^{(u)}(q)$ along q as a linear combination of the window biased probability densities

$$\rho_w^{(b)}(q) \rho^{(u)}(q) = \sum_{w=1}^N c_w \rho_w^{(b)}(q) \quad (9)$$

The coefficients, c_w in this linear combination, found by minimizing the statistical error of the density estimation along the reaction coordinate, lead to the iterative determination of the window free energies and are used to construct

$$\text{PMF}(q) = -k_B T \ln \rho^{(u)}(q) \quad (10)$$

using a program developed by Grossfield.²⁶ A few snapshots along the transition path in the I2A TMD trajectory were used directly as end points for biased window simulations for local transitions in MS6, D60, and T83.

2.8. Conformational Clustering of the Transition Path Using Wordom. Changes in interactions among different residues in RR468 along the TMD trajectory were studied using the protein structure network (PSN) analysis in Wordom.^{15,27} The TMD trajectory was segregated into 4 conformational clusters based on Distance RMS (DRMS) of all C_α atoms in the $\beta 3-\alpha 3$, $\beta 4-\alpha 4$, and $\beta 5-\alpha 5$ loops with respect to that in the active conformation of RR468. The DRMS between two structures (a and b) is defined as the square root of the averaged squared distance deviations (d_{ij}) between pairs of atoms i and j , i.e.,

$$\text{DRMS} = \sqrt{\frac{1}{N} \sum_{ij}^n (d_{ij}^a - d_{ij}^b)^2} \quad (11)$$

The DRMS for n atoms, thus, takes into account the fluctuations of $N = n(n - 1)/2$ distance pairs. Clustering was done using the Leader algorithm in Wordom. A DRMS cutoff 1.65 Å was used for clustering.

2.9. Change of Pairwise Interaction along Transition Path Using Wordom. A PSN analysis was done on each individual conformational cluster obtained from the I2A TMD trajectory.²⁷ The PSN output was compared for two

subsequent clusters to determine interactions formed/broken during transition. The interaction strength (I_{ij}) between residue pair i and j , is defined as

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} 100 \quad (12)$$

Here I_{ij} is the interaction percentage of residue i and j , n_{ij} is the number of side-chain atom pairs within a given distance cutoff (4.5 Å as a default), and N_i and N_j are, respectively, the normalization factors for residues i and j , that take into account the differences in size of the different residues, as larger residues are likely to make more contact pairs, n_{ij} .²⁷ This interaction strength provides a quantitative estimate of the overall interaction between two residues based on their proximity. In our experience it is a better estimate of the hydrophobic interaction, because n_{ij} will be higher in the case of hydrophobic interactions compared to polar interactions, the latter being strongly directional and involving fewer atoms.

Pairwise interactions with interaction percentage 2.0% (I_{\min}) in two consecutive clusters were compared. If a given pair of residues interacts with strength ($\geq 2.0\%$) in a significant number of snapshots (frequency $\geq 50\%$) in one cluster has significantly lower frequency (frequency $\leq 20\%$) in the next cluster, it was considered that the interaction had broken for the transition between one cluster and the next cluster. Similarly, residue pairs with interaction frequency less than 20% in one cluster but more than 50% in the next cluster were considered to represent the formation of an interaction.

2.10. Hydrogen Bond Analysis. Minimized conformations of the inactive and active crystal structures of RR468 were compared for their hydrogen bonding patterns using the “ptraj” module in AmberTools 10. Also, hydrogen bond analyses for the four independent trajectories, IAC, ACT, and I2A and A2I were done using ptraj. In the ptraj output for hydrogen bond analysis, all hydrogen bonds with percentage occurrences more than 1% were listed. Distance cutoff 3.5 Å (between two heavy atoms involved in the hydrogen bond) and angle cutoff 120° (heavy atom-hydrogen-heavy atom) was used to detect the hydrogen bonds. Multiple ptraj outputs were analyzed together using a script to extract meaningful information. All the hydrogen bonds in the crystal structures protein can be classified in 3 categories: hydrogen bonds that are present only in the inactive conformation, only in the active conformation, and hydrogen bonds that are there in both conformations. Hydrogen bonding patterns in two consecutive clusters along the I2A transition path (as obtained using Wordom) were analyzed to obtain changes in hydrogen bonding patterns along the transition path. The hydrogen bond analysis was done for each TMD cluster and the formation and dissociation of characteristic hydrogen bonds (of the inactive and active conformations) were compared in two consecutive clusters. The hydrogen bond analysis in each TMD cluster was obtained using ptraj and the change in hydrogen bonding pattern in two consecutive TMD cluster was obtained. The same distance and angular criterion for hydrogen bonding was used as before. A percentage occurrence of less than 15% was considered as absent and over 40% was considered as present.

3. RESULTS

MD trajectories of Apo-RR468 were generated starting from the inactive conformation (PDB id 3DGF) and active conformation (PDB id 3GL9) without BeF_3^- attached to

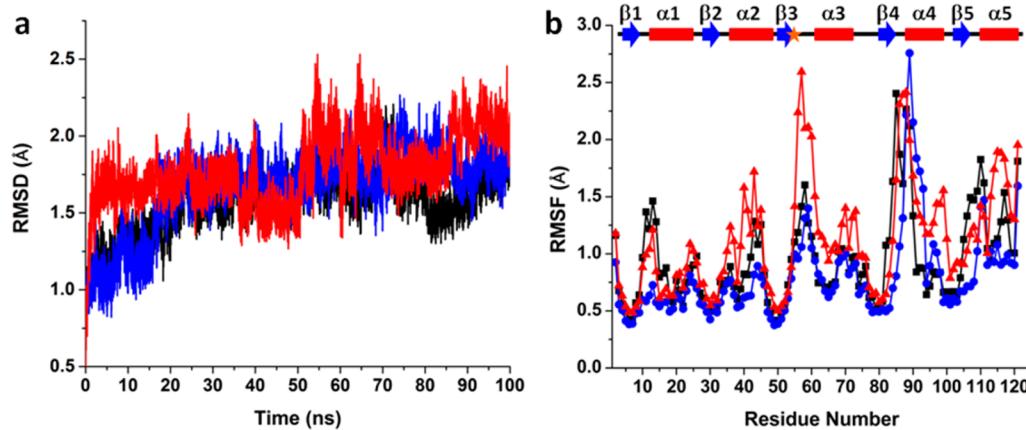


Figure 2. Backbone RMSD versus time (a) and backbone RMSF (b) plots for the 100 ns simulations starting from the inactive conformation (red), active conformation with PO_3^{2-} and Mg^{2+} (black), and active conformation without PO_3^{2-} and Mg^{2+} (blue). It is evident from the RMSF plot that the secondary structural elements in RR468 remain stable during all MD simulations.

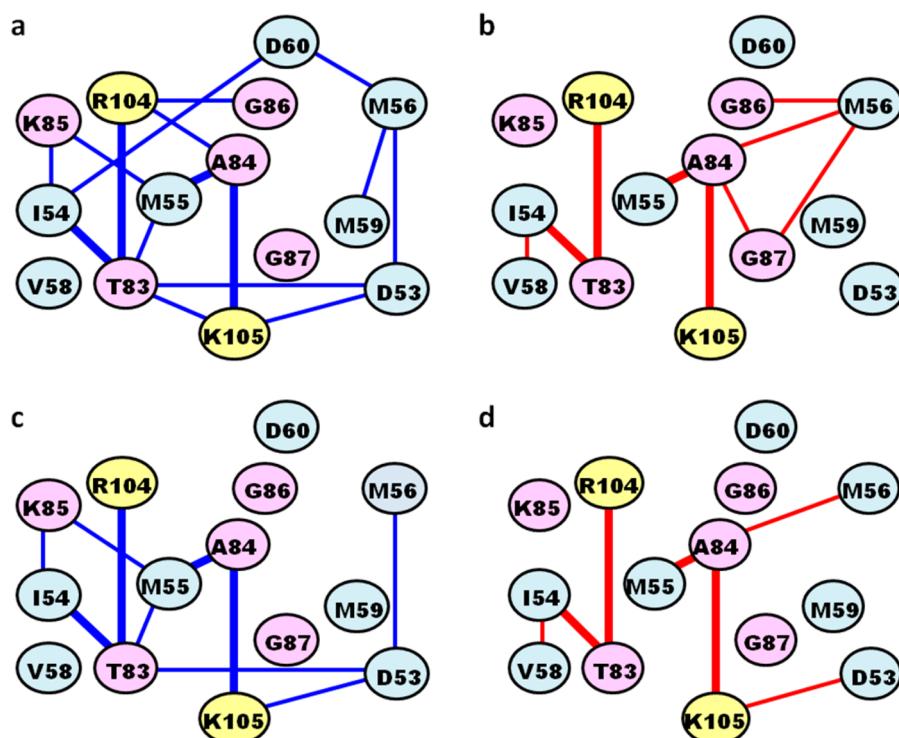


Figure 3. Intra- and interloop interactions of the residues from the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops. One residue adjacent to the loops in each N- and C-terminus of these three loops was also considered as part of the loop for their contacts in the crystal structure of the (a) active form, (b) inactive form, and corresponding simulation trajectories for (c) ACT and (d) IAC simulations. All interactions in the active and inactive form crystal structures are shown with blue and red connectors, respectively. Intra- and interloop interactions in the ACT and IAC simulations are shown with same color code in (c) and (d), respectively. Interactions shown in (c) and (d) have interaction strength more than 5.0% in more than 30% of the total number of snapshots during MD trajectories. Common interactions that are present in both crystal structures or during simulations are shown with bold connectors in all the interaction maps. Residues from $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops are shown as cyan, pink, and yellow circles, respectively.

D53 and without a Mg^{2+} ion. A phosphorylated RR468 complex was obtained by modifying BeF_3^- to PO_3^{2-} . All three systems, the inactive, apo-RR468 denoted as IAC, the active without BeF_3^- and Mg^{2+} denoted as ACT, and active with BeF_3^- and Mg^{2+} denoted as P~RR468 (see Table 1 for a summary of the simulations) were simulated for 100 ns in the presence of explicit water molecules. Since root-mean-square deviation (RMSD) versus time traces reached a plateau during the 5 ns equilibration cycle (not shown) prior to the 100 ns

production simulation, the whole 100 ns segments of the three simulation trajectories were considered equilibrated and used for further analysis (Figure 2a). Backbone root-mean-square fluctuation (RMSF) during the production runs shows that the inactive conformation of RR468 fluctuates more than the active conformation with or without the phosphoryl group and the Mg^{2+} ion. Apart from N- and C-terminal residues, fluctuations are greater in the loop regions, especially in three proximal

loops where the inactive and active conformations differ significantly (Figure 2b).

3.1. Pairwise Interactions among Residues in the Crystal Structures Are Maintained in the MD Trajectories. The protein structure networks in the crystal structures of RR468 and in the MD trajectories ACT and IAC were studied using Wordom.²⁷ Protein structure network (PSN) analysis on a given segment of trajectory provides a pairwise interaction map among all residue pairs. For a given pair of residues it provides in a number of snapshots the interaction strength, I_{ij} , as defined in the Methods section, is greater than a cutoff, I_{\min} . Intra- and interloop interactions among residues in the three proximal mobile loops are summarized in Figure 3(a,b). All the common interactions among the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops, as observed in the crystal structures of RR468, are stable during MD simulation (Figure 3c,d). Interaction between D53 and T83 in the active conformation and the ACT trajectory is due to orientation of the T83 side chain toward the site of phosphorylation (D53) in the active form. In the active conformation, M56 is buried in a hydrophobic pocket. In this conformation, M56 is in contact with D53. This interaction is stable in the ACT simulation, but in the IAC simulation M56 makes weak contact with other loop residues. All the interactions involving the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops along with one residue each following and preceding these loops are summarized in Table S6, Supporting Information. For the sake of simplicity, interactions not there in the crystal structures but formed during simulations were ignored in this analysis.

3.2. Hydrogen Bond Comparison of the Crystal Structures and MD Trajectories. We first consider hydrogen bonding interactions in the crystal structures and simulation trajectories, ACT and IAC, starting from the respective crystal structure conformations in order to obtain end point patterns that will be connected later in the TMD simulations. All hydrogen bonds in RR468 can be classified into three categories: hydrogen bonds that are present only in the inactive conformation, only in the active conformation, and hydrogen bonds that are there in both. Obviously not all of these hydrogen bonds are stable in the presence of explicit water molecules during the simulation and there are new hydrogen bonds formed during the equilibration process. The results of the hydrogen bond analysis based on the minimized crystal structures of the both forms of the RR468 are summarized in Table S1, Supporting Information.

There are 14 hydrogen bonds that are identified in the inactive conformation but not there in the active conformation crystal structures, but only three of these hydrogen bonds are stable (percentage occurrence more than 40%) during the IAC trajectory. These three hydrogen bonds are also somewhat stable in the ACT simulation as well. It can be concluded that these characteristic hydrogen bonds present in the inactive but not in the active conformation as identified from the crystal structure are not very stable during MD simulation.

There are 26 hydrogen bonds that are there in the active conformation but not in the inactive conformation. Two out of these 26 hydrogen bonds have comparable stability in the IAC trajectory as well (backbone–backbone hydrogen bond, T83-K105, and side chain–side chain hydrogen bond, R15-E31). Eight of these hydrogen bonds are stable only in the ACT trajectory but not in the IAC trajectory. The salt bridge between D9 and K105 brings the $\beta_1\text{-}\alpha_1$ and $\beta_5\text{-}\alpha_5$ loops closer. The N-termini of all 5 α -helices are proximal to the site of

phosphorylation. The N-termini of the α_1 , α_2 , and α_3 helices are more stable in the ACT simulation. This is evident from the stability of the backbone hydrogen bonds for A12-K16 (α_1), I37-K41 (α_2), G61-K65, and T63-K67 (α_3) in the ACT simulation to that compared to the IAC simulation.

There are 35 hydrogen bonds that are there in both the active and inactive conformations. Most of these common hydrogen bonds involve backbone interactions only and have comparable stability in both these trajectories.

3.3. PCA, DEVIATION, and EXTEND Show that the $\beta_3\text{-}\alpha_3$ Loop Can Partially Transit between Inactive and Active Forms in IAC and ACT Simulations. During simulations starting from the inactive (IAC) and active (ACT) states, the trajectories fluctuate around their respective starting conformations in the sense that the flexible regions, the three crucial loops ($\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$), do not, on the 100 ns MD time scale, sample each other's conformations. It could be that, with mobile loops in the active and inactive forms, an overlap of loop conformations would occur over times beyond MD practicality. To monitor the relative motions of the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops, representative distances between residues were chosen in such a way that the differences in their distances were maximized between the active and inactive conformations in the crystal structures (Table 2). The

Table 2. Relative Orientation of the $\beta_3\text{-}\alpha_3$ Loop with Respect to the $\beta_4\text{-}\alpha_4$ and $\beta_5\text{-}\alpha_5$ Loops in RR468 Was Monitored by Tracking the Distances between the C α Atoms^a

residue	distance between C α atoms (Å)		
	inactive	active	difference
M56-N34	13.1	5.8	7.3
V58-G87	9.1	23.4	14.3
V58-P106	17.8	23.0	5.2

^aFollowing residues were tested: M56 and V58 from the $\beta_3\text{-}\alpha_3$ loop, G87 from the $\beta_4\text{-}\alpha_4$, and P106 from the $\beta_5\text{-}\alpha_5$ loop. N34 marks a very stable region in the protein, as clear from the RMSF plot in Figure 2. Residues in the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops are shown in bold.

M56-N34, V58-G87, and V58-P106 distances maintain their characteristic values during the simulations starting from the active conformation with or without phosphoryl group and from the inactive conformation (Figure 4). The possibility of any rearrangement spanning the other's conformation in the $\beta_3\text{-}\alpha_3$ loop was ruled out using this analysis. Since the difference in distance among any such residue pairs from the $\beta_4\text{-}\alpha_4$ and $\beta_5\text{-}\alpha_5$ loops is small in the active and inactive conformation, the possibility of any conformational transition among backbone atoms in three mobile proximal loops were studied based exclusively on the position and relative orientation of $\beta_3\text{-}\alpha_3$ with respect to the core of the protein and with respect to the other two mobile proximal loops: $\beta_4\text{-}\alpha_4$ and $\beta_5\text{-}\alpha_5$. Note that the stable (nonloop) residues are very closely aligned in the two structures (Figure 1a,b). Using these criteria, we do find that there is no loop conformational overlap in the 100 ns IAC and ACT MD trajectories.

To see if this lack of connection is due to the physically short time scale that is practical with MD we use the DEVIATION and EXTEND methods (sections 2.5 and 2.6) that are based on a principal component analysis of the trajectory data. PCA decomposes the atom fluctuations of a trajectory into modes that successively incorporate decreasing amounts of the total

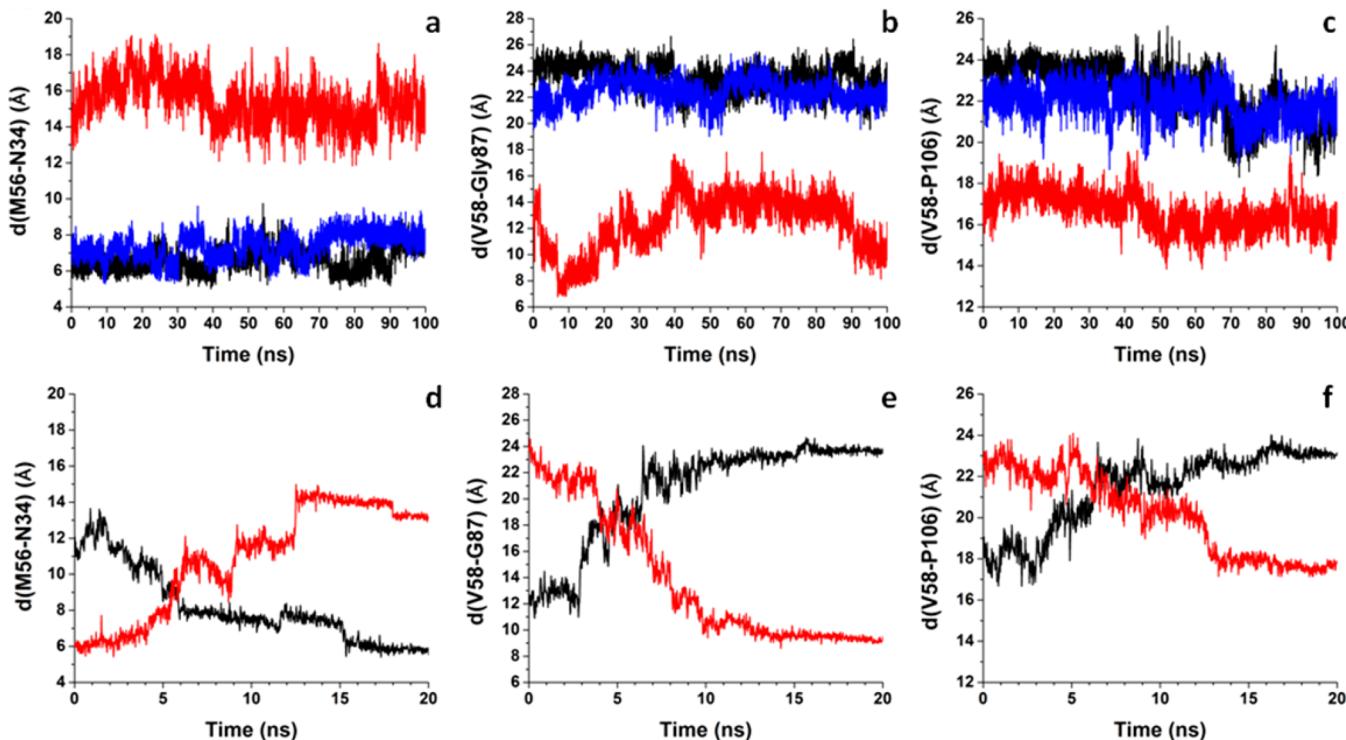


Figure 4. Top panel shows distances between C α atoms of (a) M56 and Asn34, (b) V58 and G87, (c) V58 and P106 residues in the IAC, ACT, and P~RR468 trajectories (red, black, and blue, respectively). Bottom panel shows the same distances during TMD pathways I2A and A2I (in black and red, respectively).

mean square fluctuation over the trajectory duration. If there is a separation in the sizes of the mode eigenvalues λ_ν^2 with a small set of large eigenvalues and a large set of small eigenvalues, then those first few modes of large amplitude may account for most of the “productive” motion, while the many small-amplitude modes may correspond to “noise”. PCA was carried out using the backbone heavy atoms after best fitting the trajectory snapshots onto the crystal structure on all CA atoms of the residues, excluding the loops, as summarized in Table 3. The

Table 3. PCA Mode Eigenvalues Derived from the Inactive and Active Trajectories

trajectory	mode	λ_ν (Å)	λ_ν^2 (Å 2)	cumulative fraction
IAC	1	0.480 79	0.231 16	0.231 16
	10	0.143 96	0.020 73	0.693 99
	20	0.085 88	0.007 38	0.799 90
	30	0.062 90	0.003 96	0.852 90
	47	0.045 28	0.002 05	0.900 30
ACT	1	0.459 16	0.210 83	0.210 83
	10	0.130 94	0.017 15	0.602 57
	20	0.090 96	0.008 27	0.720 03
	30	0.071 15	0.005 06	0.783 83
	74	0.037 58	0.001 41	0.900 85

first 10 and 20 eigenvalues out of 1440 eigenvalues obtained from the inactive trajectory account for 69.4% and 80.0% of the total fluctuation in the system, respectively. The fluctuations are relatively less in the active simulation, where the first 10 and 20 principle components obtained from active trajectory account for 60.3% and 72.0% of the total fluctuation in the system, respectively. Thus, in both simulations, a small number of modes do capture the bulk of the loop fluctuations.

As discussed in the Methods section, if it were true that, e.g., active conformations are encoded in the inactive trajectory fluctuations, but they are not seen in the MD trajectory because the trajectory is not run for a sufficiently long time, then it may be that some set of the large “productive” modes from the inactive trajectory PCA will point toward active conformations, and vice versa. The DEV method is one way to address this issue. It projects a final state structure onto the PCA modes of a trajectory based on an initial state structure. If the initial state trajectory does “point” in the direction of the final state structure, then DEV defined in eq 4 should track the decay of the λ_ν^2 eigenvalues from the initial state PCA. The scaling of DEV with λ_ν^2 is shown in Figure 5. As noted above, there is good separation between a small number of the large PCA eigenvalues from the many small eigenvalues. The DEV measure modestly tracks the decay of the λ_ν^2 indicating some degree of (loops) movement of the inactive toward the active form. Also displayed in Figure 5 is the backbone conformation (gray) of RR468 reconstructed from the first 50 PCA modes of the inactive trajectory, the equilibrated inactive conformation (red), and active conformation crystal structure (violet). Focusing on loop $\beta 3-\alpha 3$, the predicted position is intermediate between the inactive and the active conformation, which is consistent with the modest agreement between DEV and the eigenvalue decays. Also displayed are snapshots at 2 (yellow), 4 (green), and 6 (blue) ns of the 20 ns TMD trajectory that was used to span the inactive to active conformations. The loop movement in DEV is between the 4 and 6 ns of the TMD trajectory, indicating that a reasonable direction of loop movement toward the active form is encoded in the inactive trajectory.

In the DEV method, as more modes are included, the reconstructed coordinates will eventually match the target

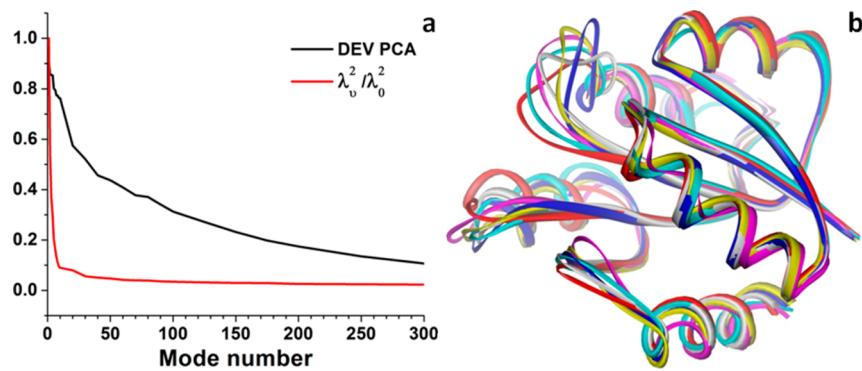


Figure 5. (a) DEV that monitors the approach toward the inactive conformation as a function of the number of modes used in the expansion of the difference of equilibrated inactive conformation and active conformation crystal structures in the PCA basis, along with the decay of the inactive fluctuation eigenvalues, λ_v^2 . The decay of DEV with the number of modes is slower than that of the eigenvalues, showing that the inactive fluctuations modestly encode active conformations. (b) Backbone conformation (gray) of RR468 reconstructed from first 50 PCA modes in the inactive trajectory compared with the equilibrated inactive conformation (red) and active conformation crystal structure (blue). Conformational snapshots at 2, 4, and 6 ns of the TMD trajectory are shown in cyan, magenta, and yellow, respectively. The reconstructed $\beta_3\text{-}\alpha_3$ loop structure (top-left loop) shows that an intermediate position between inactive and active conformations can be achieved.

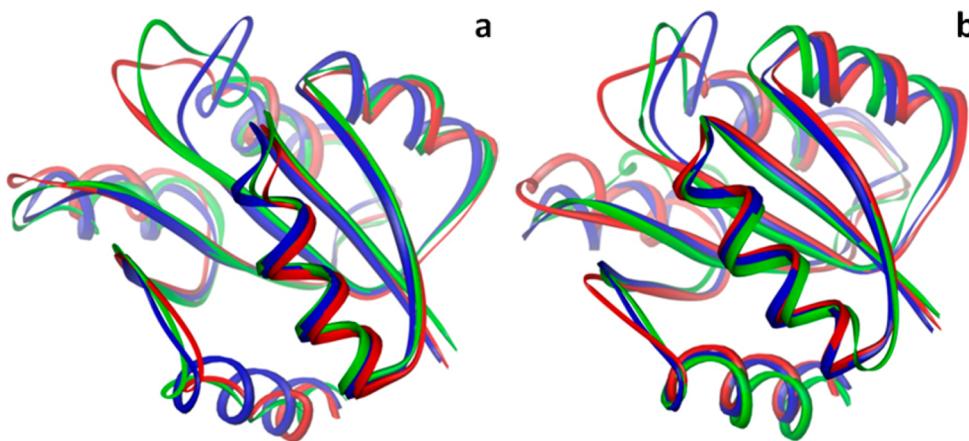


Figure 6. EXTEND method. (a) Equilibrated structure of the inactive (red) conformation and crystal structure of the active conformation (blue) of RR468 were aligned with the atomic coordinates reconstructed from the first 10 PCA modes derived from the inactive trajectory. (b) Equilibrated structure of the active (blue) conformation and crystal structure of the inactive conformation (red) of RR468 were aligned with the atom coordinates reconstructed from the first 20 PCA modes derived from the active trajectory. Loop $\beta_3\text{-}\alpha_3$ in the inactive trajectory points toward the active conformation and vice versa.

conformation (active conformation in this case). Thus, a more objective method is desired. The EXTEND method only uses the active conformation to measure how well the inactive trajectory succeeds in matching its configuration when minimizing the RMS distance between them. Atomic coordinates of the backbone atoms reconstructed from the first few PCA-modes in the inactive trajectory were compared with the equilibrated conformation of the inactive form and the crystal structure of the active form. Similarly, atomic coordinates of the backbone atoms reconstructed from the first few PCA-modes in the active trajectory were compared with the equilibrated conformation of the active form and the crystal structure of the inactive form. In either case, the large amplitude fluctuations in a trajectory started with one end state indeed points in the direction that samples the other substrate as well, as shown in Figure 6. Again, as in the DEV method, the $\beta_3\text{-}\alpha_3$ loop distinctly points in a direction that, for the active trajectory, spans the inactive-to-active conformations, and vice versa.

3.4. Targeted Molecular Dynamics (TMD) Can Span Inactive to Active, and Vice Versa, Conformations.

Starting from the inactive or active conformation of RR468, a conformational transition to the other form was not observed within the 100 ns equilibrium simulation (IAC or ACT). Thus, we applied TMD to find a possible transition path from the inactive to active conformation, or the reverse direction.

The convergence of beginning to final conformations in TMD was confirmed by monitoring three representative distances noted in Table 2, and from the positions and interactions of three residues M56, D60, and T83, that undergo significant rearrangement during transition (Figures 3 and 7). We repeated the TMD simulations two more times for I2A or A2I with slightly different starting points (from the equilibrated structures during MD simulations) to reach a consensus. The active and inactive conformations from the crystal structure were considered as target structures for the I2A and A2I trajectories, respectively. In all I2A trajectories, the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loop conformations finally converge to that of the active form between 10 and 15 ns (Figure 3d,e,f, respectively). Rearrangements in the M56, D60, and T83 side chains occur somewhere near to 15 ns in all three I2A transition paths (Figure 7a,b,c, respectively). In the A2I trajectories

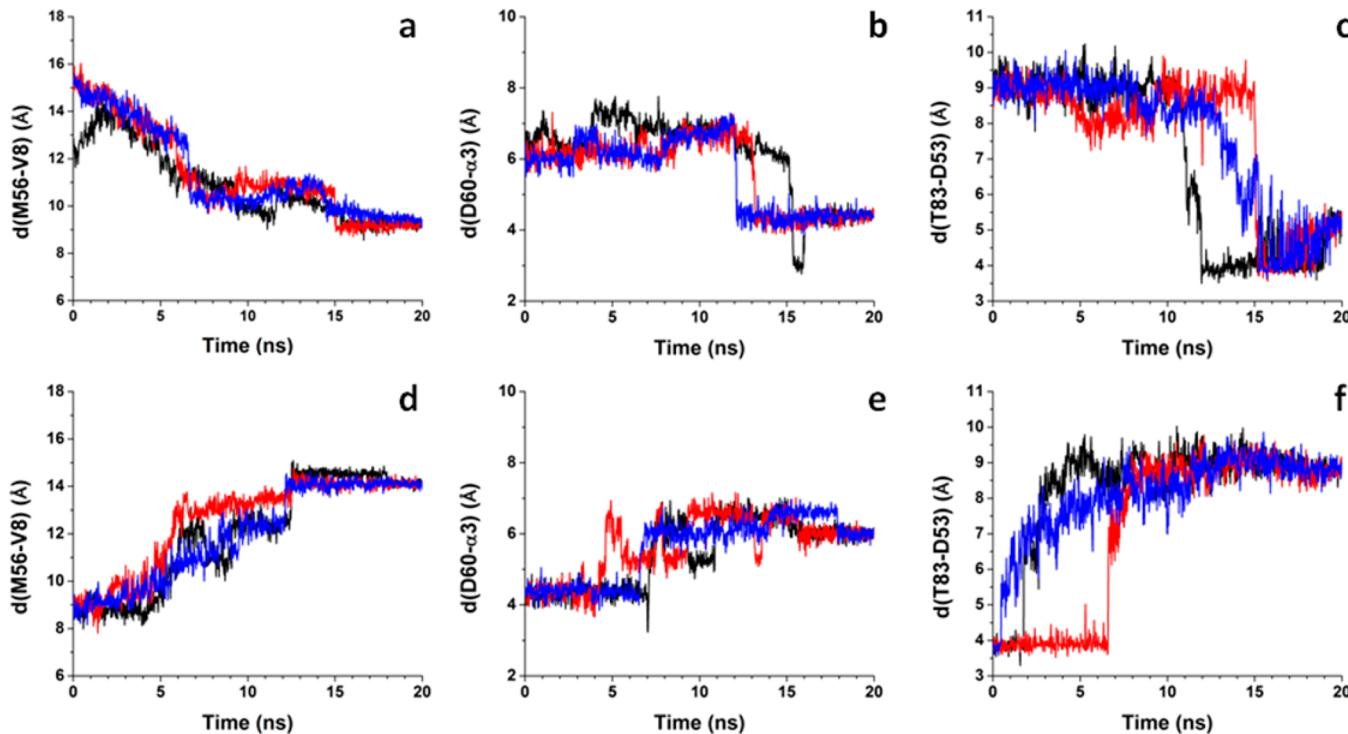


Figure 7. Time series of parameters that best describe the conformation and position of M56, D60, and Th83 in the I2A (a, b, c) and A2I TMD trajectories (d, e, f): distance between C α atoms of M56 and V8 in a hydrophobic core (a, b); average distance between C γ atom of D60 and backbone N atoms of three N-terminal residues of the α 3 helix (G61, F62, T63) (c, d), and distance between the O γ atom of T83 and the C γ atom of D53 (e, f). Results from 3 independent TMD trajectories in either direction (I2A or A2I) are marked in 3 colors.

(Figure 7d,e,f, respectively) transitions are more-or-less a reflection of the I2A path. During conformational transitions in either direction, the M56 rearrangement follows multiple stable intermediate states, but the D60 and T83 rearrangements are more drastic in nature (Figure 7).

3.5. Potentials of Mean Force (PMF) for Transitions of Key Residues in the TMD Trajectory. As shown in Figure 7, residues M56, D60, and T83 undergo significant rearrangements during the TMD trajectories, which were not observed during the MD simulations (Figure S2, Supporting Information). These transitions were monitored by various atom distances. However, to construct potentials of mean force (PMF), more meaningful reaction coordinates that best describe these transitions needed to be introduced (Table 4). For M56 we use the distance between the C α atoms in M56 and V8. For this gradual M56 transition, end points for the PMF are 10.5 and 9.25 Å. To span the reaction coordinate, four windows were needed to obtain good window overlap (Figure S3, Supporting Information). For D60 and T80, with their abrupt nature of rearrangement in the TMD trajectory, we had

to perform brief steered molecular dynamics (SMD) spanning the reaction coordinates, from a characteristic value in the inactive conformation to that in the active conformation (details in Table S9, Supporting Information). Starting conformations for the biased window simulations were obtained from these SMD trajectories. Because of the mainly side chain based rearrangements of these residues, we found that the T83 transition was best described by the indicated side chain atoms and the D60 transition best described by the indicated dihedral angle in Table 4. To obtain good window overlap with these reaction coordinates, eight and ten windows were needed for the construction of the PMF for the D60 and T83 rearrangements, respectively (Figure S3, Supporting Information).

The PMF plots are shown in Figure 8. From the offset of the PMF value in the inactive conformation versus that in the active conformation, it is evident that the active conformation is a relatively high energy conformation. The energy barrier for the M56 rearrangement is relatively low compared to that for D60 and T83. Conformational rearrangement in M56 follows several intermediate conformations, whereas for D60 and T83 the rearrangements are very abrupt in nature (Figure 8 and Figure S3, Supporting Information).

3.6. Hydrogen Bond Analysis of the TMD Trajectory Identifies Residues Critical to the Conformational Transition. In the TMD trajectories, several hydrogen bonds, which we will refer to as non-native, transiently exist during the transition path, and are absent in either of the end state simulations. The role of a particular non-native hydrogen bond was confirmed using the multiple TMD trajectories. To identify critical non-native hydrogen bonds that may assist conformational transitions in RR468, we selected hydrogen

Table 4. Reaction Coordinates Used for Calculation of the PMF for Local Transitions

local changes	reaction coordinate	acronym
M56 transition from exposed (inactive) to buried (active) state	distance between C α atoms in M56 and V8	d(M56-V8)
shift in D60 interaction with α 2 in inactive to that with α 3 in active conformation	torsion angle C γ -C β -C α -C of D60	tD60(C γ -C β -C α -C)
T83 flip from outward to inward	distance between T83 O γ and D53 C γ	d(T83-D53)

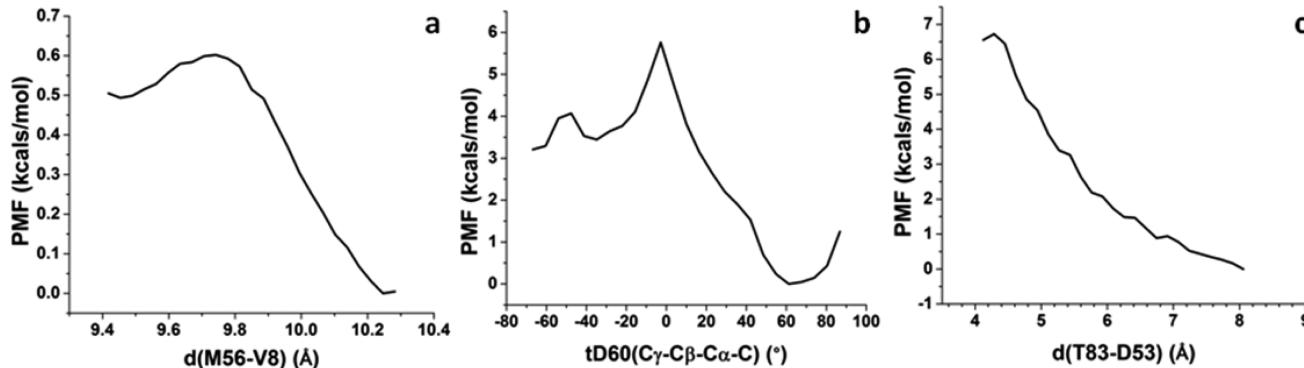


Figure 8. Potentials of mean force along reaction coordinates that best describe local conformational rearrangements in (a) M56, (b) D60, and (c) T83 for the inactive to active transition. In all three plots, the reaction coordinates proceed from left to right during the inactive to active transition.

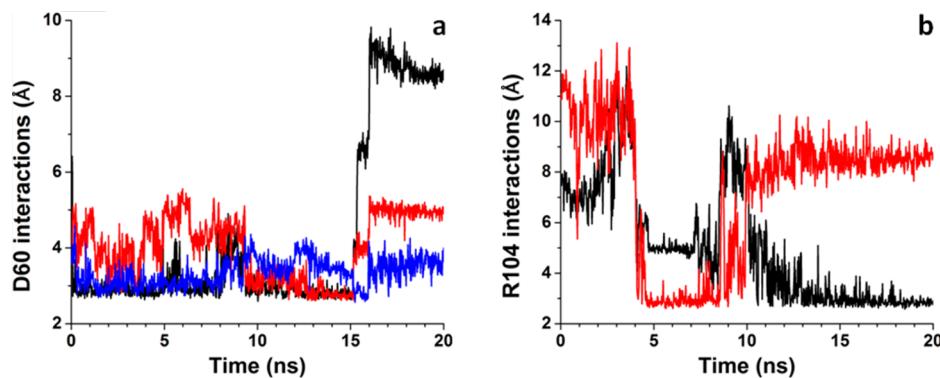


Figure 9. (a) Transient hydrogen bond formed by D60 O δ with backbone N of G35 (black), G61 (red), and D60 (blue) during the inactive to active TMD trajectory. (b) Hydrogen bonding of R104 with several residues in the β 4- α 4 loop is a significant feature of the I2A TMD trajectory.

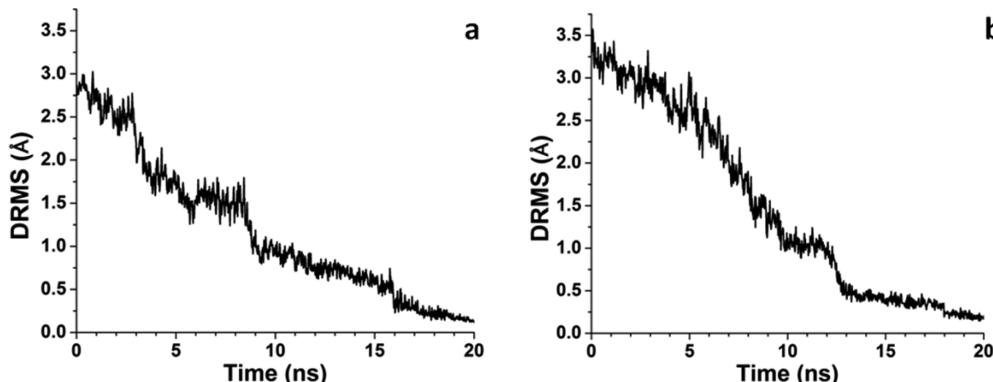


Figure 10. (a) DRMS of the all C α atoms in the β 3- α 3, β 4- α 4, and β 5- α 5 loops with respect to the active conformation of RR468 in the I2A TMD trajectory. The partitioning into four clusters is I (0–2.88 ns), II (2.90–4.02 ns), III (4.04–9.22 ns), and IV (9.24–20.00 ns). (b) DRMS of the all C α atoms in the β 3- α 3, β 4- α 4, and β 5- α 5 loops with respect to the inactive conformation of RR468 in the A2I TMD trajectory. The partitioning into four clusters is I (0–6.58 ns), II (6.60–8.12 ns), III (8.14–12.50 ns), and IV (12.52–20.00 ns).

bonds with lower occurrences in both the ACT and IAC simulations (less than 20%) but relatively higher occurrence in the TMD trajectories (e.g., I2A). We have chosen a small cutoff for occurrences (10%) in TMD simulations for this purpose, since conformational sampling is limited in the biased simulation. In the I2A TMD trajectory, 26 such hydrogen bonds were found that are present in at least 15 consecutive snapshots (lifetime 300 ps or more). Out of 26 hydrogen bonds identified, 19 involve side chain interactions (Table S7, Supporting Information).

Two notable non-native hydrogen bonds involve residues D60 and R104. Hydrogen bonds between the D60 side-chain

and main chain N of G35 can be crucial for the D60 rearrangement. This hydrogen bond either never appears or has very rare appearances during the IAC and ACT simulations (Table S7, Supporting Information). The occurrence of the hydrogen bond between D60 side chain and G35 backbone N atom is 54.0% in the I2A trajectory, but this hydrogen bond never forms during the IAC or ACT simulations (Figure 9a). The same hydrogen bond is present in 29.3% and 24.4% of the total number of snapshots in the other two I2A trajectories. D60 forms similar non-native hydrogen bonds with adjacent residues (F62 and T63) prior to its local conformational rearrangement in the other I2A trajectories as well. Significant

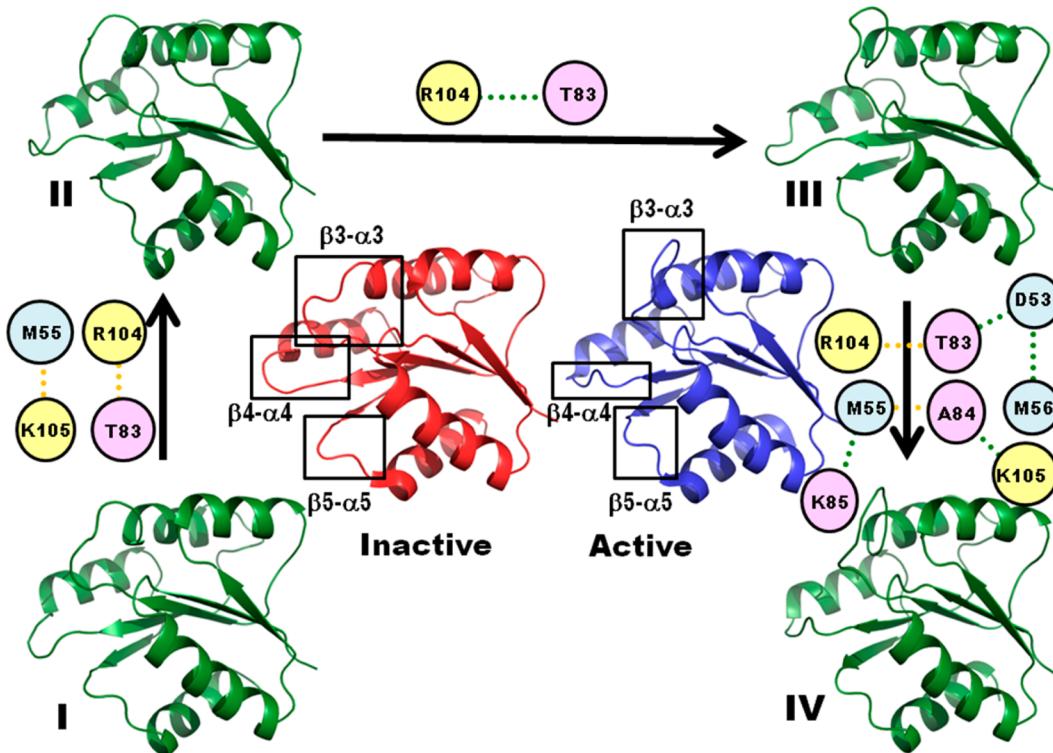


Figure 11. Representative structure from each cluster during the I2A TMD trajectory is shown in green cartoon. Inter- and intraloop pairwise interactions among $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops that are broken and formed during each transition are shown with orange and green broken lines, respectively. Conformations of the three proximal mobile loops are shown for the inactive (red) and active (blue) forms of RR468.

interactions of D60 with T63 were also observed in A2I TMD trajectory (occurrence 23.10%) (Table S8, Supporting Information). Transient hydrogen bonds formed by R104 with the $\beta_4\text{-}\alpha_4$ loop residues and their adjacent residues are likely to couple the $\beta_5\text{-}\alpha_5$ and $\beta_4\text{-}\alpha_4$ loop movements during the transition. This feature was observed in two out of three independent I2A trajectories. A salt bridge between R104 and E91 found in the A2I TMD trajectory is also crucial for the conformational transition in the reverse direction (occurrence 48.60% in A2I) (Table S7, Supporting Information). R104 remains solvated most of the time during the ACT and IAC equilibrium simulations.

3.7. Conformational Clustering of the TMD Trajectories Permits Identification of Intermediates. The above hydrogen bond analysis presents a global view over the entire TMD trajectories. A similar analysis based on conformational clustering along the TMD trajectory was carried out to identify changes in pairwise interactions among different residues. Conformational clustering of the TMD trajectories was based on the DRMS (Distance RMS) of all the $C\alpha$ atoms in the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loop regions to identify crucial intermediate states during the conformational transition (Figure 10). The 20 ns I2A TMD trajectory was segmented into 4 distinct clusters (0–2.88 ns, 2.90–4.02 ns, 4.04–9.22 ns, and 9.24–20.00 ns) based on the DRMS with respect to the active conformation. The $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loop conformations during the first 2.88 ns of the TMD trajectory remain close to those of the inactive conformation; while these loop conformations converge to those of the active conformation during the entire last half of the TMD trajectory (cluster IV).

The TMD trajectory A2I was segmented into four distinct clusters: I (0–6.58 ns), II (6.60–8.12 ns), III (8.14–12.50 ns), and IV (12.52–20.00 ns). Conformations of the $\beta_3\text{-}\alpha_3$, $\beta_4\text{-}\alpha_4$, and $\beta_5\text{-}\alpha_5$ loops remain close to that of the active form during first 6–7 ns of the A2I TMD trajectory.

3.8. PSN to Identify Key Pairwise Interactions along the I2A and A2I TMD Trajectories. As noted above, PSN analysis on a given segment of trajectory provides a pairwise interaction map among all residue pairs. Residue pairs having significant changes in the frequency of interaction at a given interaction strength in two consecutive clusters are likely to point to the interactions formed or broken during this stage of the transition. Changes in interaction patterns from one to the next cluster during the inactive to active or reverse direction, of the conformational change are summarized in the Table S2 and S3 (Supporting Information), respectively. Some notable changes in the pairwise interactions along with the hydrogen bond analysis among the TMD clusters are given below.

For the inactive to active (I2A) transition, interactions between the $\beta_5\text{-}\alpha_5$ loop and the other two mobile loops become weak during the initial stage of the I2A TMD simulation. This is evident from the breaking of interactions of R104 and K105 with $\beta_3\text{-}\alpha_3$ and $\beta_4\text{-}\alpha_4$ residues M55 and T83, respectively (Figure 11). Among other significant changes involving three proximal mobile loop residues, there is the loss of interaction between V58 and F62, signifying weakening of the N-terminus of the α_3 helix. Movement of the N-terminus of the α_4 helix is implicated in the loss of interaction between T83 and K91. The N-terminus of the α_4 helix is not stable during any simulation. A stable salt bridge between D9 and K105 is a feature in the ACT simulation. This salt bridge forms in the early stage of the I2A simulation. Changes in the interaction

pattern involving proximal mobile loop residues are summarized in Table S2, Supporting Information.

The interaction between T83 and R104 forms and breaks in two intermediate segments of the I2A simulation (Figure 11). The side-chain of M56 resides in the void space of the phosphoryl group and becomes close to the D10 residue. D10 is coordinated with the Mg^{2+} ion in the phosphorylated complex. D60 and P57 form a contact with N34 prior to side-chain rearrangement of D60. The T83 side chain is tethered in an outward orientation and interacts with V102 and R104 (Figure 11). This interaction finally breaks during the last stage of the TMD simulation and the T83 side chain flips toward D53, forming a hydrogen bond with the D53 side-chain. Hydrophobic collapse was observed in the side-chains of residues in the proximal end that involves L14, V18, L22, L39, M59, and F107 during the last phase of the conformational transition. The frequency of interaction involving these residues for the interaction strength cutoff, 2% increases for the TMD cluster IV compared to that in cluster III (Table S2, Supporting Information). M56 forms contacts with D53 in the buried conformation, as observed in the active form.

For the active to inactive (A2I) transition, many hydrophobic interactions involving V8, A32, I37, L52, V79, and M56 are lost during the early stage of the TMD trajectory. Interactions of M56 with V8 and D53 are broken and that with M59 is formed as a result of the expulsion of M56 side chain from the deep hydrophobic pocket consisting of V8, L52, and V64 (Table S3, Supporting Information). Toward the end of the A2I TMD simulation, the hydrophobic core of the protein is regenerated through expulsion of solvent molecules. Formation of hydrophobic contacts involving A32, I37, I54, V58, and F62 in the later stages of this simulation leaves the distortion among two end states to a minimum, except for the three crucial loop regions.

3.9. Hydrogen Bond Analysis to Identify Key Pairwise Interactions along the I2A and A2I TMD Trajectories. The definition of interaction in the PSN analysis is based on total number of atom pairs in contact between two residues and thus can be ambiguous about the exact nature of the interaction. Therefore, as complementary information to PSN, the hydrogen bonding pattern was also constructed that compares each of the consecutive TMD clusters. The hydrogen bond analysis performed on each of these TMD clusters was carried out as discussed in the Methods section 2.10. The criteria for formation and dissociation of different hydrogen bonds along two consecutive segments along TMD trajectories are as follows. If a hydrogen bond significantly present (occurrence >40.0%) in one segment is found to be rare (occurrence <15.0%) in the next segment, it was considered a critical hydrogen bond that has to dissociate for the formation of the second segment from the first. Similarly, a hydrogen bond that is rare (<15.0%) in one segment but significant (>40.0%) in the next segment was considered a critical hydrogen bond that has to form for the formation of the second segment from the first. A summary of all hydrogen bonds that are broken or formed during conformational transition in I2A or A2I trajectory is given in Tables S4 and S5, Supporting Information, respectively. Some notable changes in the hydrogen bonding interaction are as follows.

In the inactive to active transition direction, at the beginning of the transition the T83 side chain is stabilized in an outward conformation due to its hydrogen bonding interaction with M103. This hydrogen bond is observed in the IAC simulation

as well. This interaction breaks toward the end of the I2A simulation, and T83 flips inward to form a hydrogen bond with D53, as in the active conformation. Formation and dissociation of several hydrogen bonds among the R104 side chain and different residues of the $\beta 4-\alpha 4$ loop is a notable feature of the I2A simulation. Loss of backbone interaction between D60 and V64 at the early stage of the simulation also points toward the weakening of the N-terminus of the $\alpha 3$ helix.

In the active to inactive transition direction, breaking of the N-terminal of the $\alpha 3$ backbone interaction was also observed. A backbone hydrogen bond between G61 and L65 breaks during the early stage of this transition path. R104 also plays a crucial role in this process as it forms a hydrogen bond with T83 in the intermediate stages, but this hydrogen bond is absent in either end state conformations.

4. DISCUSSION

The crystal structure of the inactive and active conformations differ most in the three proximal loops, $\beta 3-\alpha 3$, $\beta 4-\alpha 4$, and $\beta 5-\alpha 5$ (Figure 1b). The overall backbone RMSD for the IAC, P~RR468, and ACT trajectories show similar deviations from their respective starting structures, while the respective residue RMSFs show that the three proximal loops closely surrounding the site of phosphorylation exhibit greater fluctuations compared to the rest of the protein (Figure 2 b). Backbone atoms in the inactive trajectory fluctuates more, especially around the $\beta 3-\alpha 3$ loop, whereas the active conformation, with or without the phosphoryl group and Mg^{2+} , is relatively stable during the 100 ns simulations. There are four distal loops on the other side of the protein away from the site of phosphorylation: loops $\alpha 1-\beta 2$ (residues 26–27), $\alpha 2-\beta 3$ (44–48), $\alpha 3-\beta 4$ (72–78), and $\alpha 4-\beta 5$ (98–100) (Figure 1 a). All these distal loops are very stable in all the trajectories as compared to the proximal loops (Figure 2b). The greater overall fluctuations in the IAC trajectory relative to the other trajectories may indicate a larger conformational space for the inactive conformation. The active conformation without the phosphoryl group on D53 and Mg^{2+} is also trapped in a similar conformation as in the P~RR468 trajectory. Charge–charge repulsions among a set of ionized residues spanning E88, E89, D90, and E91 at the N-terminal of the $\alpha 4$ helix are a somewhat unstable region in the protein (Figure 2b). This region is involved in dimerization of the receiver domain in another response regulator protein and hence may have other physiological significance.⁵

There are a greater number of intra- and interloop interactions of the proximal loops $\beta 3-\alpha 3$, $\beta 4-\alpha 4$, and $\beta 5-\alpha 5$ in the active conformation compared to the inactive conformation crystal structure (Figure 4a,b) and they are more stable during the respective simulation trajectories (Figure 4c,d). Interactions that are present in the crystal structures of both forms are stable in the dynamic scenario as well. The greater number of contacts of M55, M56, M59, and F107 with other loops and the core of the protein in the ACT simulation (Table S6, Supporting Information) indicate more interactions of the proximal loops in the active conformation. A strong hydrogen bond between the conserved T83 and phosphoryl group on D53 (or the D53 side chain) brings $\beta 3$ and $\beta 4$ closer, which facilitates more hydrophobic contacts between the $\beta 3-\alpha 3$ and $\beta 4-\alpha 4$ loops. The hydrogen bond between T83 and D53 was not observed at all during the IAC simulation.

In the presence of the phosphoryl group, the side chain of K105 is positioned deep within a pocket making contacts with

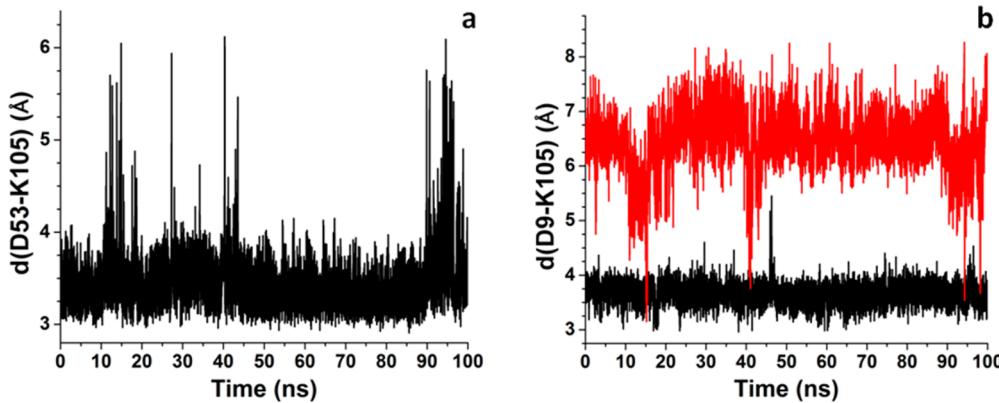


Figure 12. (a) Distance between K105 N ξ and D53 C γ in the IAC simulation. (b) The salt bridge between D9 and K105 in the ACT (black) is very stable, whereas this interaction is not present except transiently during the IAC simulation (red).

D9 and the phosphoryl group on D53. The adjacent D10 residue interacts with the Mg $^{2+}$ ion; thus, its conformation is directly affected by phosphorylation or dephosphorylation of the D53 residue. In the absence of the phosphoryl group, K105 can still maintain the salt bridge with D9. A salt bridge between K105 and D9 is one reason for trapping the active conformation in a similar conformation in the absence of the phosphoryl group. In the crystal structure of the inactive form, K105 has two alternative conformations. In one conformation, K105 can form a salt bridge with D53, but this salt bridge is not very stable, as the K105 side chain can sometimes move toward D9, and D53 becomes exposed to the bulk solvent (Figure 12a,b). The dephosphorylation reaction of the phosphorylated aspartate is water mediated, whereby a water molecule is believed to act as the nucleophile to attack the phosphorus atom.⁶ The salt bridge between the highly conserved D9 from $\beta 1\text{-}\alpha 1$ and K105 from $\beta 5\text{-}\alpha 5$ (Figure 12b) and several stable hydrophobic contacts involving proximal loop residues and the core of the protein limits the exposure of the site of phosphorylation to the bulk solvent, making the phosphoryl group on D53 very stable (half-life, $t_{1/2}$ is ~ 19 h. at 4 °C in P~RR468).⁶ However, P~RR468 is quickly dephosphorylated in the presence of HK, as several residues in the proximal loop regions interact with the HK domain as shown in the HK853-RR468 complex crystal structure. The K105 position is also inferred to be crucial for complex formation with HK853.⁶ It seems that the higher mobility of the proximal loops may actually prevent the process of complex formation.

The active conformation of the $\beta 3\text{-}\alpha 3$ loop is stabilized by hydrophobic interactions of the M56 side-chain with residues V8, L52, and V64 in a deep hydrophobic pocket (Table S6, Supporting Information). M56 is completely solvent exposed during the inactive simulation. Flexibility within the $\beta 3\text{-}\alpha 3$ loop is completely arrested when M56 is docked into the deep hydrophobic pocket in the active conformation. This finding is consistent with the fact that complex formation with HK853 is affected when M56 is mutated, as the $\beta 3\text{-}\alpha 3$ loop conformation in the inactive form of RR468 is not compatible to interact with HK853, as in the HK853-RR468 crystal structure (3DGE).⁶ However, the side-chain rearrangements in M56 do not have a significant barrier (~ 0.6 kcals/mol for exposed to buried). The change in the M56-V8 C α distance during the inactive to active transition is from 13 to 15 Å in the inactive conformation to ~ 9 Å in the active conformation (Figure 7). This side-chain rearrangement follows several intermediate states in either direction (Figures 7a and 8a). For the initial conformational

transition from 13 to 15 to 10.2 Å, the side-chain is just sampling solvated states and should have a flat PMF in this range. Therefore, we focused the PMF simulation for side-chain orientation from an exposed (10.2 Å) to buried (9.4 Å) distance that corresponds to the dramatic change in the conformation.

A specific conformation of the $\beta 3\text{-}\alpha 3$ and $\beta 4\text{-}\alpha 4$ loops is necessary for the HK853-RR468 interaction.⁶ This conformation in $\beta 3\text{-}\alpha 3$ is accompanied by D60 interacting with the N-terminus of the $\alpha 3$ helix in the active conformation, while in the inactive conformation, D60 interacts weakly with the N-terminus of the $\alpha 2$ helix. In the I2A TMD trajectory, the D60 side chain interaction with G61 (Figure 9a) coincides with formation of cluster IV in the I2A TMD trajectory (Figure 10). The energetic barrier of the side-chain rearrangement in D60 is relatively high (~ 6 kcals/mol). The interaction between RR468 and HK853 may help lower this barrier. Also, some transient hydrogen bonds of the D60 side-chain with the backbone N of G35, G61 and backbone N of itself, at the same time brings the N-terminus of the $\alpha 2$ and $\alpha 3$ helices closer in the I2A TMD trajectory, to facilitate the transition of the D60 side chain (Figure 9).

Our PCA based analysis techniques, DEV and EXTEND support the idea that in spite of the lack of transition from inactive to active state and vice versa in the equilibrium trajectories, the primary sequence and the conformation of one state do have information for a transition direction toward the other state. Quasi harmonic analysis on the receiver domain of nitrogen regulatory protein C done by Lei et al. is similar to the DEV method in our analysis. They found that, e.g., the inactive form modestly points toward the active conformation. Backbone conformations constructed from our PCA based methods point in the right direction of the conformational change, especially for the $\beta 3\text{-}\alpha 3$ loop (Figures 5 and 6). However, the precise conformation that enables side-chain rearrangement and complete conformational transition was not achieved in the equilibrium simulation, suggesting that the state of phosphorylation is also important for the conversion. . The direction of conformational change in the $\beta 4\text{-}\alpha 4$ and $\beta 5\text{-}\alpha 5$ loops do not clearly approach the other form due to two reasons. First, the difference between the end states is quite small for these loop regions. Second, PCA modes may be affected by the anomalous fluctuation in the $\alpha 4$, $\beta 5$, $\alpha 5$ regions in the C-terminal region of the protein (Figure 2b).

The TMD simulations were performed sufficiently slowly such that the restraint energy never becomes too large, to

minimize deviations from the equilibrium-like scenario. Of course, in finite length trajectories various paths will still be sampled. The relative orientations of the three mobile proximal loops become similar to those of the active conformation somewhere between 10 and 15 ns in the I2A TMD simulation, and they do not change much thereafter (Figure 3d,e,f). From the DRMS plot it also is evident that the backbone conformations in these loops remain the same during the last half of the I2A trajectory (Figure 10). The entire last half of the I2A TMD trajectory fits into the fourth TMD cluster. In all the I2A trajectories, side chain rearrangements in M56, D60, and T83 take place after all three loops adopt a specific active-like conformation, i.e., around 15 ns (Figure 7a,b,c). In the A2I trajectories, side chain rearrangements in those three residues take place at an early stage (before 7 ns) of the simulation (Figure 7d,e,f) where the conformation of three proximal loops is still close to that in the active form (Figure 3d,e,f). Side-chain rearrangements are very local conformational changes within the protein that depend on backbone conformations in the system. Apparently, a specific active-like conformation of the three proximal loops is a prerequisite for the side chain rearrangement in M56, D60, and T83.

Combining hydrogen bond and PSN analyses along the TMD trajectory provides complementary information about the time evolution of pairwise contacts between residues and the physical nature of their interactions. Formation of a contact between R104 and T83 at the third segment of the I2A TMD trajectory (Figure 11) is due to hydrogen bond formation between the R104 side-chain and the A84 backbone (Table S4, Supporting Information). Breaking of the T83-M103 hydrogen bond toward the end of I2A trajectory (Table S4, Supporting Information) is replaced by one between T83 and D53. Rearrangement of T83 from an outward, relative to the site of phosphorylation, inactive conformation to an inward, active conformation orientation (Figure 1) is energetically expensive, i.e., free energy barrier ~ 7 kcal/mol (Figure 8). This is partly because the T83 side-chain can be stabilized in the outward conformation due to an interaction with the M103 backbone.

Interactions between N34 and V58, D53 and M56, and M59 and T63 form toward the end of the I2A conformational transition. These interactions have to break at the early stage of the A2I simulation. The formation of contacts between Q36 and V58, L39 and M59 at the end of I2A and breaking of contacts between Q36 and M59 at the early stage of the A2I simulation also involve the same region in the protein (Tables S2 and S3 in Supporting Information). Changes in conformation at the early stage of the A2I trajectory proceed in an approximately reversible manner as compared to the I2A trajectory.

Our hydrogen bond and the Wordom PSN analyses rule out the possibility of any global rearrangement within the protein over the course of the conformational transition. Hydrogen bonds that have more than 50% occurrence in both ACT and IAC trajectories have a very strong presence in the I2A TMD trajectory as well. Thus, the stable regions in the IAC and ACT regions are mostly maintained over the TMD trajectory. The PSN analysis suggests that the conformational rearrangements occur within the proximal loops and their adjacent residues during the transition. Conformational rearrangements in the distal loops and their adjacent residues are minimal during the I2A or A2I TMD trajectories (Table S2 and S3, Supporting Information). Also, the distal loops were identified as one of the most stable regions in the protein from the RMSF calculation.

The role of R104 seems to be crucial for the conformational transition in the protein as the interactions of R104 form and break at different stages of the transition (Figure 9b). The interaction between the T83 and R104 side-chains might also facilitate this local rearrangement of the T83 residue (Figure 11).

5. CONCLUDING REMARKS

The conformational transition in RR468 is a much slower process (μ s-ms order) than can be captured in 100 ns time scale MD simulations. By complementing the inactive and active form based trajectories with TMD trajectories that span these end states, we have identified structural determinants that may be responsible for the conformational transition. The active conformation without the phosphoryl group is locked close to the conformation of the active form with the phosphoryl group, mainly due to interactions of the M56 and K105 residues. The TMD simulations connecting the active and inactive state conformations of RR468 provide a possible path for this conformational transition and thereby indicate crucial interactions involving residues D60 and R104 that make the conformational transition possible. The relative population of active and inactive states at a given temperature should be affected if these positions are mutated. This hypothesis could be tested using mutation experiments. We demonstrated here how a local event like the phosphorylation or dephosphorylation of a particular amino acid can act as a conformational switch for other local conformational rearrangements leading to a global conformational change in RR468. Biased MD simulations from various starting points along the transition path enabled us to estimate the free energy barrier for several key local conformational rearrangements. From the IAC and ACT trajectories, using PCA-based methods, we found that there is some information embedded in one trajectory about the direction of the β 3- α 3 loop conformational change required to access the other conformation, but other interactions may be required to complete the transition—a situation in between induced fit and conformational selection.

■ ASSOCIATED CONTENT

S Supporting Information

Figure S1 provides a description of the system used for DFT calculation. Figures S2 provides descriptors for the side chain orientations of three crucial residues that are different in the active and inactive conformation and remains close to the respective starting structure during simulation. Figure S3 describes background information for PMF calculation. Table S1 shows the fate of the native hydrogen bonds in the crystal structures during simulations. Tables S2, S3 summarize changes in pairwise interactions during TMD simulations in inactive to active and reverse direction. Tables S4, S5 summarize changes in hydrogen bonding interactions during TMD simulations in inactive to active and reverse directions. Table S6 describes the fate of pairwise interactions involving residues for three crucial residues. Table S7–S8 summarizes non native hydrogen bonds identified from the TMD trajectories. Table S9 summarizes details of the Steered MD simulations used for conformational sampling of the intermediate states of the D60 and T83 transitions. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: cukier@chemistry.msu.edu. Phone: 517-355-9715 x 263.
*E-mail: yanh@msu.edu. Phone: 517-353-5282.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Kaillathe Padmanabhan of the Macromolecular Computer Facility and Dr. Dirk Colbray of the High Performance Computer Center at Michigan State University for technical support.

REFERENCES

- (1) Bourret, R. B.; Silversmith, R. E. Two-Component Signal Transduction. *Curr. Opin. Microbiol.* **2010**, *13*, 113–115.
- (2) Stewart, R. C. Protein Histidine Kinases: Assembly of Active Sites and Their Regulation in Signaling Pathways. *Curr. Opin. Microbiol.* **2010**, *13*, 133–141.
- (3) Casino, P.; Rubio, V.; Marina, A. The Mechanism of Signal Transduction by Two-Component Systems. *Curr. Opin. Struct. Biol.* **2010**, *20*, 763.
- (4) Bourret, R. B. Receiver Domain Structure and Function in Response Regulator Proteins. *Curr. Opin. Microbiol.* **2010**, *13*, 142–149.
- (5) Gao, R.; Stock, A. M. Molecular Strategies for Phosphorylation-Mediated Regulation of Response Regulator Activity. *Curr. Opin. Microbiol.* **2010**, *13*, 160–167.
- (6) Casino, P.; Rubio, V.; Marina, A. Structural Insight into Partner Specificity and Phosphoryl Transfer in Two-Component Signal Transduction. *Cell* **2009**, *39*, 325–336.
- (7) Gao, R.; Stock, A. M. Biological Insights from Structures of Two-Component Proteins. *Annu. Rev. Microbiol.* **2009**, *63*, 133–154.
- (8) Knaggs, M. H.; Salsbury, J. F. R.; Edgell, M. H.; Fetrow, J. S. Insights into Correlated Motions and Long-Range Interactions in Chey Derived from Molecular Dynamics Simulations. *Biophys. J.* **2007**, *92*, 2062–2079.
- (9) Huang, H.; Ozkirimli, E.; Post, C. B. Comparison of Three Perturbation Molecular Dynamics Methods for Modeling Conformational Transitions. *J. Chem. Theory Comput.* **2009**, *5*, 1304–1314.
- (10) Roche, P.; Mouawad, L.; Perahia, D.; Samama, J.-P.; Kahn, D. Molecular Dynamics of the FixJ Receiver Domain: Movement of the B4–A4 Loop Correlates with the in and out Flip of Phe101. *Protein Sci.* **2002**, *11*, 2622–2630.
- (11) Lei, M.; Velos, J.; Gardino, A.; Kivenson, A.; Karplus, M.; Kern, D. Segmented Transition Pathway of the Signaling Protein Nitrogen Regulatory Protein C. *J. Mol. Biol.* **2009**, *392*, 823–836.
- (12) Ma, L.; Cui, Q. Activation Mechanism of a Signaling Protein at Atomic Resolution from Advanced Computations. *J. Am. Chem. Soc.* **2007**, *129*, 10261–10268.
- (13) Tatsis, V. A.; Tsoulosb, I. G.; Krinasa, C. S.; Alexopoulos, C.; Stavrakoudis, A. Insights into the Structure of the PmrD Protein with Molecular Dynamics Simulations. *Int. J. Biol. Macromol.* **2009**, *44*, 393–399.
- (14) Formaneck, M. S.; Ma, L.; Cui, Q. Reconciling the "Old" and "New" Views of Protein Allostery: A Molecular Simulation Study of Chemotaxis Y Protein (CheY). *Proteins: Struct. Funct. Bioinform.* **2006**, *63*, 846–867.
- (15) 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.
- (16) Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer Science: New York, 2004.
- (17) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412–425.
- (18) García, A. E. Large-Amplitude Nonlinear Motions in Proteins. *Phys. Rev. Lett.* **1992**, *68*, 2696–2699.
- (19) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; et al. *Amber 10*; University of California, San Francisco, 2008.
- (20) Cheatham, T. E., III. *Ambertools Users' Manual*, 2008.
- (21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Rob, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; et al. *Gaussian 03*; Gaussian, Inc.; Wallingford, CT, 2003.
- (22) Wang, J. *Ambertools Users' Manual*, 2008.
- (23) Cukier, R. I. Apo Adenylate Kinase Encodes Its Holo Form: A Principal Component and Varimax Analysis. *J. Phys. Chem. B* **2009**, *113*, 1662–1672.
- (24) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (25) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135*, 40–57.
- (26) Grossfield, A. *Wham: The Weighted Histogram Analysis Method*, v 1.4, 2003.
- (27) Seeber, M.; Felline, A.; Raimondi, F.; Muff, S.; Friedman, R.; Rao, F.; Caflisch, A.; Fanelli, F. Wordom: A User-Friendly Program for the Analysis of Molecular Structures, Trajectories, and Free Energy Surfaces. *J. Comput. Chem.* **2010**, *32*, 1183–1194.