

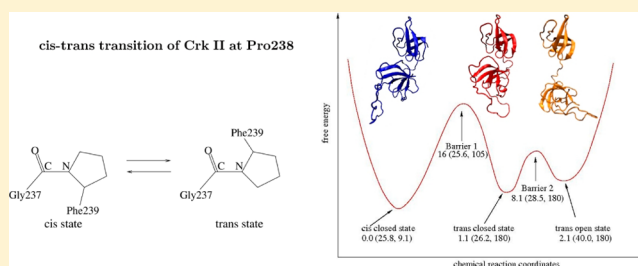
## Molecular Dynamics of the Proline Switch and Its Role in Crk Signaling

Junchao Xia and Ronald M. Levy<sup>\*,†</sup>

Department of Chemistry and Chemical Biology, Rutgers, the State University of New Jersey, 610 Taylor Road, Piscataway, New Jersey 08854, United States

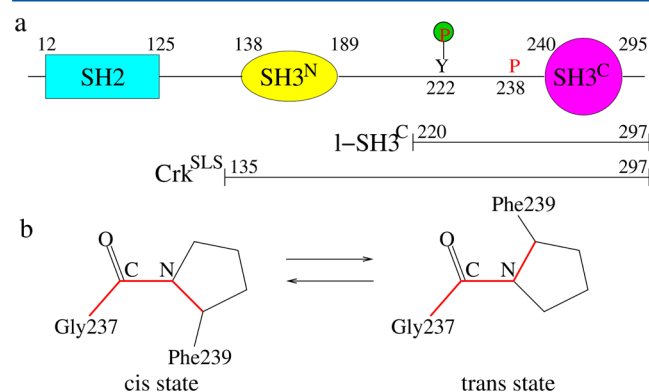
## Supporting Information

**ABSTRACT:** The Crk adaptor proteins play a central role as a molecular timer for the formation of protein complexes including various growth and differentiation factors. The loss of regulation of Crk results in many kinds of cancers. A self-regulatory mechanism for Crk was recently proposed, which involves domain–domain rearrangement. It is initiated by a cis–trans isomerization of a specific proline residue (Pro238 in chicken Crk II) and can be accelerated by Cyclophilin A. To understand how the proline switch controls the autoinhibition at the molecular level, we performed large-scale molecular dynamics and metadynamics simulations in the context of short peptides and multidomain constructs of chicken Crk II. We found that the equilibrium and kinetic properties of the macrostates are regulated not only by the local environments of specified prolines but also by the global organization of multiple domains. We observe the two macrostates (cis closed/autoinhibited and trans open/uninhibited) consistent with NMR experiments and predict barriers. We also propose an intermediate state, the trans closed state, which interestingly was reported to be a prevalent state in human Crk II. The existence of this macrostate suggests that the rate of switching off the autoinhibition by Cyp A may be limited by the relaxation rate of this intermediate state.



## INTRODUCTION

Proline cis–trans isomerization describes two distinct states ( $0^\circ$  for cis and  $\pm 180^\circ$  for trans, respectively; see Figure 1) of the backbone dihedral angle  $\omega$  (defined as  $C\alpha-C-N-C\alpha$ )



**Figure 1.** Domain organization and proline switch of chicken Crk II protein. (a) Schematic diagram of the domain arrangement of Crk. Pro238 can behave as a regulation switch through the cis–trans isomerization. Tyr222 can be phosphorylated by the enzyme Abl. I-SH3<sup>C</sup> and Crk<sup>SLS</sup> represent the one-domain (residue 220 to 297) and two-domain (residue 135 to 297) systems respectively studied in this paper. (b) cis–trans isomerization about the prolyl Gly237–Pro238 bond (a case of Xxx–Pro peptide). The corresponding dihedral angle  $\omega$  is defined by four atoms ( $C\alpha 237-C237-N238-C\alpha 238$  as shown in the red line).

presented in the X-Pro peptides. Proline isomerization is one important way to achieve large conformational changes and reach various macrostates of multidomain proteins without modifying the covalent structures.<sup>1–4</sup> Conformational changes resulting from proline switching are crucial to control protein activity in many biological processes including cell signaling,<sup>5–8</sup> neurodegeneration,<sup>9</sup> channel gating,<sup>10</sup> gene regulation,<sup>11</sup> and others.<sup>12–16</sup>

Many experimental studies<sup>17–21</sup> and theoretical investigations<sup>22–33</sup> on short-peptides containing only one proline residue show that the trans state population predominates (>95%) and the free energy barriers to rotation about the torsion angle  $\omega$  are relatively large ( $\sim 20$  kcal/mol), although the populations and the barriers can be adjusted due to different substitutions and side chain effects.<sup>18,34–37</sup> The well-known preference for the trans conformation has also been found through statistical analysis of proline-containing proteins from the Protein Data Bank,<sup>30,38,39</sup> and has been attributed to the steric effects<sup>18</sup> of ring atoms, although other contributions including electronic effects<sup>20</sup> may exist. Recent results also revealed that the equilibrium and exchange rates between macrostates varies considerably from short polyproline peptides to large proteins due to changes in the local and global environments.<sup>5,6,40–46</sup> For example, a recent ion mobility-mass

Received: February 6, 2014

Revised: March 27, 2014

Published: April 4, 2014

spectrometry (IM-MS) study found that the nonapeptide bradykinin (containing three proline residues at positions of 2, 3, and 7) contains up to 10 metastable states depending on the solution composition and the multiple structures are associated with different combinations of cis and trans states from the three proline residues.<sup>43,44</sup> Förster resonance energy transfer (FRET) experiments on polyprolines with 1–10 residues have also revealed structural heterogeneity with subpopulations that do not interconvert on time scales from nano- to milliseconds.<sup>45,46</sup> For large proteins, although statistical surveys of X-ray structures of nonredundant chains from the Protein Data Bank found that around 95% are in the trans configuration,<sup>30,38,39</sup> recent studies show that the population of the cis state can be dramatically increased for some systems such as staphylococcal nuclease,<sup>47</sup> 5-HT<sub>3</sub> receptor,<sup>10,22</sup> and Crk adaptor proteins.<sup>5,6</sup>

The Crk family of adaptor proteins is believed to act as a molecular bridge to form protein complexes by recruiting downstream effectors to upstream phosphorylated tyrosine motifs.<sup>48,49</sup> Crk proteins are expressed in most tissues and mediate timely formation of protein complexes including various growth and differentiation factors.<sup>50,51</sup> Crk proteins usually are overexpressed in many human cancers.<sup>52–55</sup> Crk II is one of the five types of Crk adaptor proteins and consists of three domains (Figure 1):<sup>55,56</sup> a single Src homology 2 (SH2) domain, a N-terminal Src homology 3 (SH3<sup>N</sup>) domain, and a C-terminal Src Homology 3 (SH3<sup>C</sup>) domain. Between the SH3<sup>N</sup> and SH3<sup>C</sup> domains there is an approximately 50-residue long linker containing a specific tyrosine residue (Tyr222 in chicken Crk II) that can be phosphorylated by Abl kinase.<sup>57</sup> The SH2 domain is used to achieve the binding of phosphorylated tyrosine motifs with a consensus sequence of pTyr-x-x-Pro.<sup>58,59</sup> The SH3<sup>N</sup> domain binds proline-rich motifs of the polyproline II (PPII) subtype with the consensus Pro-x-x-Pro-x-(Lys, Arg) (e.g., Abl kinase).<sup>60</sup> The SH3<sup>C</sup>, however, does not bind to these canonical PPII motifs due to the lack of aromatic residues at the binding surface.<sup>61</sup> However, recent NMR experiments (more details below) show that the SH3<sup>C</sup> domain plays a critical role in a new autoinhibition mechanism.<sup>5,6</sup>

The regulation of Crk II is achieved through an autoinhibition mechanism by reorganizing the arrangement of the three Crk II domains.<sup>55,56</sup> Early studies suggested that the regulation of Crk is accomplished by phosphorylation,<sup>62</sup> namely, that Abl kinase recruited by the SH3<sup>N</sup> domain can phosphorylate Tyr222 in the linker and promote the formation of an intramolecular interaction between the SH2 domain and Tyr222, which further decreases the availability of the SH2 domain for other tyrosine-phosphorylated binding partners. In contrast, recent findings showed that the unphosphorylated human CrkII adopts a compact structure where protein binding to the SH3<sup>N</sup> domain is partially occluded by the linker region and autoinhibition is achieved.<sup>63</sup> Recently a new autoregulatory mechanism has been proposed based on a study of the one-domain and two-domain subsystems of chicken Crk II, recognizing the importance of the SH3<sup>C</sup> domain.<sup>5,6</sup> The cis–trans transition of proline (Pro 238) results in two distinct macrostates, one is an autoinhibited (closed, 90%) cis state in which the canonical binding site of SH3<sup>N</sup> favors hydrophobic contacts with the SH3<sup>C</sup> domain, and the other is an uninhibited (open, 10%) trans state in which no interaction exists between the two SH3 domains, and the SH3<sup>N</sup> is accessible for binding to other molecules (e.g., Abl kinase). Namely, the SH3<sup>C</sup> domain is

able to block the binding of other ligands to the SH3<sup>N</sup> domain by reorganizing itself and forming hydrophobic contact interactions with the N domain.<sup>5,6</sup> NMR experiments from different groups also revealed that this autoregulatory mechanism might not be conserved across species.<sup>5,6,63–65</sup>

Due to the high barrier (~20 kcal/mol) and the long time scale (>seconds) for the proline cis–trans switching, it is difficult to observe the exchange dynamics directly from experiments; furthermore, the transition paths by which Crk switches between the autoinhibited and uninhibited states are unknown. Some structural and dynamic information about the individual macrostates can be obtained from chemical shifts and fast NMR relaxation experiments.<sup>5,6,63–65</sup>

Computational methods are a complementary way to find the missing components of the autoregulatory mechanisms, but standard molecular dynamics (MD) simulations using high performance resources can only reach time scales on the order of microseconds.<sup>66</sup> For simulations on the order of milliseconds and beyond, we need to resort to more advanced sampling methods. Several MD simulations by accelerated methods<sup>23,24,26,27,29,31,41,42,67,68</sup> have been used to study the proline cis–trans isomerization in short peptides. Metadynamics<sup>69,70</sup> is a recently developed technique for accelerating simulations using a time-dependent biasing potential acting on certain prechosen chemical reaction coordinates. Metadynamics simulations have also been performed for short peptide fragments in 5-HT<sub>3</sub> receptor<sup>22</sup> and in the HIV-1 capsid protein (CA) and its complex with Cyclophilin A (Cyp A),<sup>71</sup> although few simulations of large proteins have been reported.

Due to the multidomain structures of Crk proteins, extending previous results from short peptides to large systems is challenging since no domain–domain interactions were considered in those studies. To understand the role of the proline switch on large Crk signaling proteins at the molecular level, we performed extensive standard MD and metadynamics simulations for short proline-containing peptides and the large protein systems related to Crk II proteins. Our simulations suggest the existence of a new intermediate state. Namely, the Proline 238 has switched “on”, but Crk remains in the autoinhibited “off” conformation due the existence of stabilizing hydrophobic contacts between the Src N and C terminal domains. We speculate on the possible effects of this new intermediate on the kinetics of autoinhibition.

## ■ MATERIALS AND METHODS

**Molecular Dynamics Simulations.** All MD and metadynamics simulations of short peptides and Crk proteins were performed using the molecular simulation package GROMACS,<sup>72,73</sup> an improved version of AMBER 99SB<sup>74</sup> force field by the Show group (AMBER99SB-ILDN in GROMACS), and the SPCE<sup>75</sup> explicit solvent model, at the same experimental temperature 298.15 K. The initial conformations for the simulations come from the NMR structures related to Crk proteins in the PDB and capped by the acetyl (ACE) and N-methyl (NME) groups: the fragment Gly237-Pro238-Phe239 of PDB 2L3P for ACE-GPF-NME, ACE-GP-NME, and ACE-P-NME, PDB 2L3P for the cis linker-SH3<sup>C</sup>, PDB 2L3Q for the trans linker-SH3<sup>C</sup>, and PDB 2L3S for the SH3<sup>N</sup>-linker-SH3<sup>C</sup>. These systems (in their NMR conformations) were solvated in a truncated octahedron of SPCE water molecules with the shortest distance of solute to box surfaces of 1.5 nm (for three short peptides), 2.5 nm (for the one-domain and two-domain Crk proteins), respectively. All production simulations were

under periodic boundary conditions in the constant volume, temperature, and number of particles (NVT) ensemble. Before the production runs, sequential equilibration processes in NVT (1.0 ns), NPT (1.0 ns), and three other NVT (1.0 ns) ensemble were performed to adjust the systems into the desired temperatures and volumes. During the first two equilibrium processes, the position restraints with the force constants of 1000 KJ/(mol·nm<sup>2</sup>) are applied to the solute. The position restraints were released gradually for the last 3 NVT equilibrium processes with the force constants from 100, 10, to 1 KJ/(mol·nm<sup>2</sup>). The integration time step was 0.002 ps. In all simulations, LINCS algorithm was used to constrain bonds involving hydrogen. The nonbonded cutoff for evaluating electrostatic and van der Waals forces was set to 1.0 nm. To deal with long-range electrostatic interactions, the PME algorithm was used with the default settings, including a real space grid of 0.12 nm.

**Metadynamics Simulations.** Metadynamics<sup>69,70</sup> accelerates the system out of a local free energy minimum and explores other free energy minima by continuously adding a history dependent biasing potential energy that forces the dynamics to explore conformations that were not previously visited. Compared with the equations of motion in standard MD simulations, there are additional forces acting on atoms derived from the history dependent biasing potential energy which is accumulated as Gaussian functions,

$$V_B(s(x), t) = w \sum_{t'=\tau_G, 2\tau_G, 3\tau_G, \dots} \exp\left(-\sum_{j=1}^M \frac{(s_j(x) - s_j(t'))^2}{2\delta_j^2}\right) \quad (1)$$

where  $s_j(x)$  is the  $j$ th of  $M$  chemical reaction coordinates (collective variables) constructed from the atomistic coordinates  $x$ . The constant  $w$  and  $\delta_j$  define the height and width of the Gaussian function added at constant time intervals of  $\tau_G$ . Well-tempered metadynamics<sup>76</sup> is one of many improved versions of metadynamics, which corresponds to adjusting the constant height of the Gaussian function as a time-dependent variable. Equation 1 becomes as below:

$$V_B(s(x), t) = w_0 e^{-V_b(s,t')/k_B \Delta T} \sum_{t'=\tau_G, 2\tau_G, 3\tau_G, \dots} \exp\left(-\sum_{j=1}^M \frac{(S_j(x) - S_j(t'))^2}{2\delta_j^2}\right) \quad (2)$$

where  $w_0$  is the initial height of Gaussian functions and  $\Delta T k_B$  defines a characteristic energy. Over a long time, the exact free energy can be estimated from the converged biasing potential as

$$F^w(S) = -\frac{\Delta T + T}{\Delta T} \lim_{t \rightarrow \infty} V_B(s(x), t) \\ = -\frac{B}{B - 1} \lim_{t \rightarrow \infty} V_B(s(x), t) \quad (3)$$

where  $B = (T + \Delta T)/T$  defines a scaling factor of fictitious higher temperature  $T + \Delta T$  to the normal temperature  $T$ .

The PLUMED<sup>77</sup> (version 1.3) package for the implementation of metadynamics was patched to GROMACS 4.5.5. Our well-tempered metadynamics simulations were started after finishing the five corresponding standard MD simulations, using the parameters of initial Gaussian height  $w_0 = 0.4$ , scaling factor  $B = 10$ , and Gaussian deposit time  $\tau_G = 1$  ps. The

Gaussian width  $\delta$  is 0.02 nm for distance space and 0.1 rad for dihedral space. These values were chosen so as to reach a balanced trade-off between a fast exploration of the conformations and good accuracy in the reconstructed free energy, and are close to the parameters chosen for other protein systems.<sup>71,78</sup>

**Umbrella Sampling Corrections.** The accuracy of the estimated free energy landscapes can be improved further by additional umbrella sampling using the free energy surface constructed from the metadynamics simulation as a biasing potential.<sup>79</sup> Corrections to the free energy were obtained from the probability distributions  $P^c(s)$  from the biased molecular dynamics trajectories as

$$F^c(s) = -k_B T \ln P^c(s) \quad (4)$$

where  $k_B$  is Boltzmann's constant and  $T$  is temperature. The final free energy landscapes become

$$F(s) = F^w(s) + F^c(s) \quad (5)$$

In this report we used metadynamics simulations to obtain low-resolution approximations of potentials of mean force (PMFs) and applied them as the time-independent biasing potentials during the umbrella sampling simulations for the fine-resolution results of PMFs. The umbrella sampling simulations allow some local relaxations along degrees of freedom other than that of the predefined reaction coordinates for PMFs and permit equilibration of the canonical ensembles. The simulation times for umbrella sampling were chosen in such a way that the populations integrated from all major basins do not have significant changes.

**Definition of Chemical Reaction Coordinates.** The center of mass calculations only involved the  $C\alpha$  atoms from the corresponding groups including the linker, SH3<sup>N</sup>, and SH3<sup>C</sup> domain. The peptide  $\omega$  dihedral angle was calculated from the four atoms  $C\alpha-C-N-C\alpha$  in the two neighboring residues.

The number of hydrophobic contacts between two groups was computed as

$$s_1 = \sum_{i,j} \frac{1 - (r_{ij}/r_0)^n}{1 - (r_{ij}/r_0)^m} \text{ for } r_{ij} > 0, \text{ and } s_1 = 1 \text{ for } r_{ij} \leq 0 \quad (6)$$

The summations of  $i$  and  $j$  go through all nonpolar side-chain carbon atoms belonging to two different groups.  $r_{ij} = |r_j - r_i - d_0|$  involves the distance between the  $i$  and  $j$ th atom.  $r_0$  and  $d_0$  are two preselected distances (0.4 and 0.2 nm in this paper) defining the switch function when combined with the other two constants ( $n$  and  $m$ ) whose values are 6 and 12, respectively, in our calculations.

Similarly, the number of intramolecular hydrogen bonds between a group of donors and a group of acceptors is defined as

$$s_2 = \sum_{i,j} \frac{1 - (r_{ij}/r_0)^n}{1 - (r_{ij}/r_0)^m} \quad (7)$$

where  $i$  and  $j$  are O and H atoms belonging to two different groups, respectively.  $d_{ij} = |r_j - r_i|$  is the distance between the  $i$  and  $j$ th atom.  $r_0$  (= 0.27 nm) is the preselected distance, and  $n$  (= 6) and  $m$  (= 12) are two constants that determine the steepness of the switching function. The values, defining the chemical coordinates, are similar to previous work from other groups.<sup>71,80</sup>



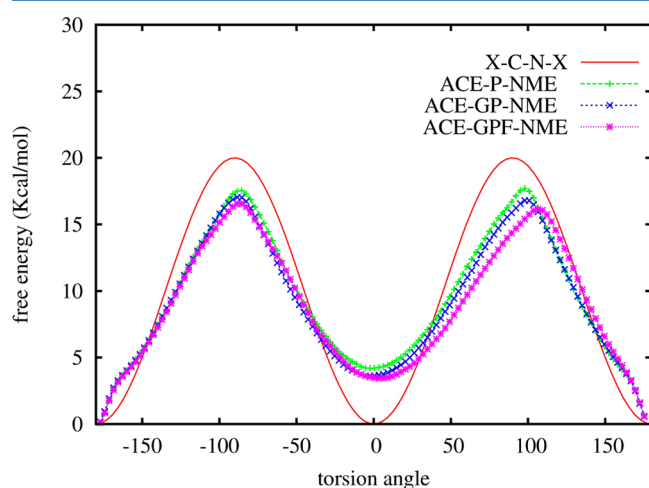
**Reweighting 2D PMFs for 1D PMFs.** In principle, the free energy surfaces (or potentials of mean force) for some variables other than the preselected chemical reaction coordinates for the biasing potentials can be reconstructed from the well-tempered metadynamics.<sup>81</sup> However, this requires many sampling snapshots between two continuous updates of the biasing potentials. In this paper, we obtained the 1D PMF (in  $s$ ) from the 2D PMFs (in  $x, y$ ) in an alternative way by using the biased MD trajectory for the umbrella sampling correction (in the time series of  $(s_i, x_i, y_i)$ ) in a similar way as WHAM.<sup>71,82</sup> The unbiased probability distribution can be calculated as

$$P^0(s_k) = \sum_i \frac{e^{\beta V_i(x_i, y_i)}}{\langle e^{\beta V_i(x_i, y_i)} \rangle^1} \quad (8)$$

where  $\beta = 1/k_B T$ ,  $i$  goes through all snapshots with the  $s$  values in the range of bin  $k$ , and  $V_B(x_i, y_i)$  is the 2D biasing potential at  $(x_i, y_i)$ . The superscript 0 denote the unbiased (real) distribution instead of 1 for the biased one. Finally the reweighted free energy can be obtained using eq 4 from the unbiased probability distribution.

## RESULTS AND DISCUSSION

**Trans State Is Predominant for Short Proline Peptides.** Figure 2 displays the free energy results as a



**Figure 2.** Free energy of short peptides as a function of  $\omega$  dihedral angle of the proline obtained from the metadynamics simulations. The initial structures of these three peptide fragments with one to three amino acid residues are from the first model of NMR derived chicken Crk II (PDB: 2L2P) (ACE-P238-NME, ACE-G237P238-NME, and ACE-G237P238F239-NME). The red line shows the function of torsion potential energy in AMBER force field,<sup>74</sup>  $V(\phi) = V_2[1 + \cos(2\phi - 180^\circ)]/2$ ,  $V_2 = 20$ .

function of the dihedral angle (one-dimensional PMFs in  $\omega$ ) involved in the proline cis–trans transition for three peptide fragments with 1 to 3 amino acid residues from chicken Crk II (PDB: 2L2P) (ACE-P238-NME, ACE-G237P238-NME, and ACE-G237P238F239-NME.) Adding the neighboring two residues does not have a pronounced effect on the free energy landscape of  $\omega$ : the trans state ( $\omega = \pm 180^\circ$ ) is the global free energy minimum and predominates (population >99%); the cis state has free energy 3–5 kcal/mol higher than that of the trans state; the transition barriers from the trans to cis state are between 15 to 20 kcal/mol (see Table 1 from the exact values).

**Table 1. Free Energy Differences and Barriers for the Short Proline-Containing Peptides (in Units of kcal/mol) from Metadynamics Simulations**

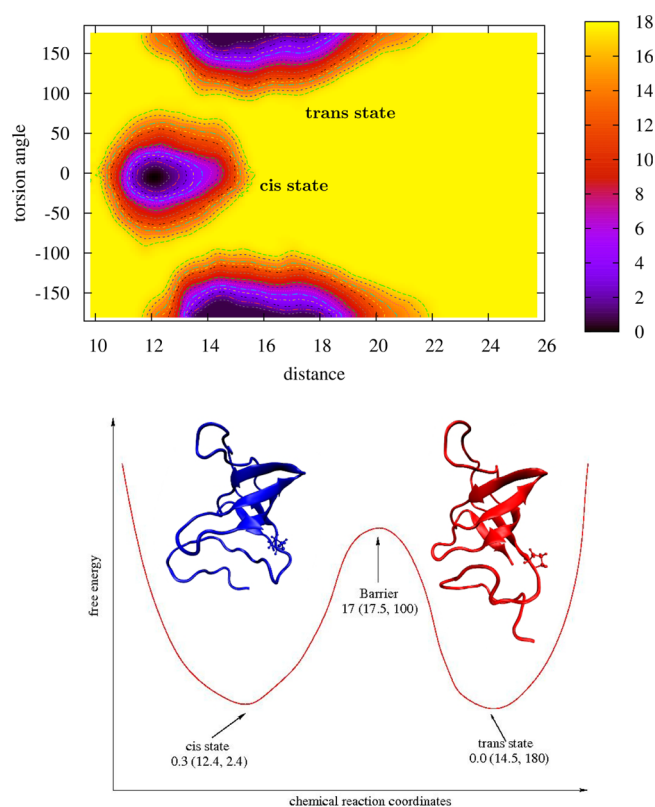
| peptides    | $F_{\text{cis}} - F_{\text{trans}}$ | $\Delta F_{\text{trans} \rightarrow \text{cis}}$ | $\Delta F_{\text{cis} \rightarrow \text{trans}}$ |
|-------------|-------------------------------------|--|--|
| ACE-P-NME   | 4.2                                 | 17.5   | 13.3   |
| ACE-GP-NME  | 3.6                                 | 17.0   | 13.4   |
| ACE-GPF-NME | 3.5                                 | 16.6   | 13.1   |

As a model peptide for proline cis–trans isomerization, ACE-P-NME has been extensively studied experimentally<sup>17–21</sup> and using computational methods, including classical molecular simulations<sup>22–24,26,27,29</sup> and quantum mechanics.<sup>25,28,30,83,84</sup> The general results of ACE-P-NME from our metadynamics simulation are in reasonable agreement with the experimental data<sup>17–21</sup> (with energy differences of  $\sim 1.5$  kcal/mol and barriers of  $\sim 20$  kcal/mol), although the free energy difference between the trans and cis state is overestimated and the barrier is underestimated. It should be noticed that it appears possible to improve these values by recently optimized force field parameters.<sup>26</sup>

**Two Macrostates of Crk Linker-SH3<sup>C</sup> with Distinct Sets of Hydrophobic Contacts.** Previous NMR results<sup>5,6</sup> showed that the cis–trans isomerization of Pro238 in Crk results in two distinct equally populated conformational states of the linker-SH3<sup>C</sup> fragment of the chicken Crk II protein. Figure 1S in the Supporting Information shows the energy landscapes of Crk I-SH3<sup>C</sup> in two dimensions: (1) the center of mass distance between the second half of the linker (residues 231 to 238) and the SH3<sup>C</sup> domain (residues 239 to 297) and (2) the  $\omega$  dihedral angle of P238 defined as in Figure 1, constructed from the 200 ns standard MD simulations started from the cis state (first model of PDB: 2L3P) and the trans state (first model of PDB: 2L3Q) respectively. The cis and trans macrostates are stable during the simulations, and no transitions occur during the 200 ns period. The comparisons of residue contact maps (Figure 2S) calculated from the NMR models with that from MD snapshots reveal that most linker–domain interactions from NMR are observable in the MD simulations. This confirms that both stable structures from the MD simulations are consistent with the NMR results. Further analysis of the structures from simulations will be presented below.

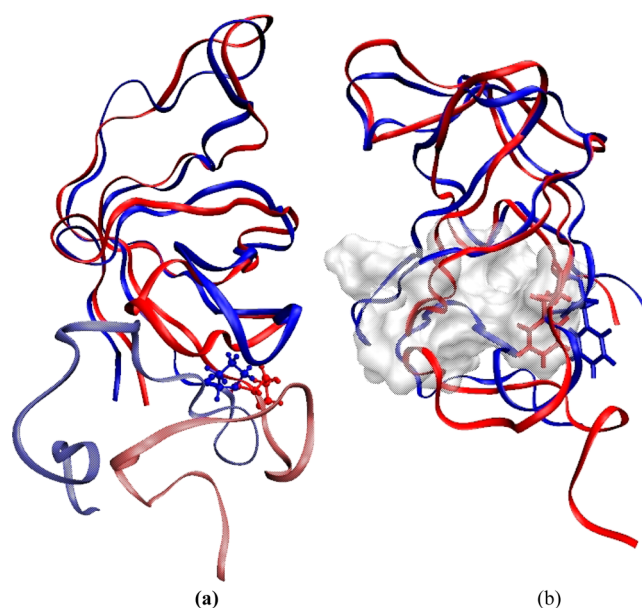
Due to the intrinsically slow interconversion rate between the two conformational states, no transitions were observed in the standard MD simulation. In contrast, in the metadynamics simulation of Crk I-SH3<sup>C</sup> many transitions are observed. The PMF along the same two dimensions as in Figure 1S is displayed in Figure 3. The metadynamics simulation explored more high energy regions in addition to the low-energy cis and trans basins sampled in the two standard MD simulations. The low energy conformers for each macrostate are shown in Figure 3S. The residue contact maps calculated from these low energy structures are exhibited in Figure 4S; they closely resemble the maps from the NMR models and also from the standard MD simulations in Figure 2S.

More detailed comparison of Crk I-SH3<sup>C</sup> structures from the NMR with that from the metadynamics simulations is presented in Figure 5S. The linker–domain (type A) and linker–linker interactions (type B) in the structures representing the two free energy minima (macrostates) are well conserved from the NMR models.<sup>5</sup> Further analysis of the residues within the interaction interfaces reveals that the cis–



**Figure 3.** (a) PMF of linker-SH3<sup>C</sup> in 2D (the center of mass distance between the second half of the linker (residue 230 to 238) and the SH3<sup>C</sup> domain (residue 239 to 297) and the  $\omega$  dihedral angle of P238) from the 60 ns metadynamics simulation started from the cis state (first model of PDB: 2L3P) and corrected by the 100 ns umbrella sampling simulation with the fixed biasing potential from the metadynamics. Two minima, (12.4, 2.4°) and (14.5, 180°), represent the two macrostates, the cis and the trans state, respectively. (b) Schematic representation of the free energy path for the cis to trans transition of linker-SH3<sup>C</sup> extracted from panel a. The numbers show the free energy values in kcal/mol and the positions in 2D PMF of panel a.

trans isomerization results in two stable macrostates with two different sets of linker–domain interactions. In the cis state, the major hydrophobic interaction region includes two residues from the second half linker (Leu231 and Leu234) and five from the SH3<sup>C</sup> domain (Ala241, Val267, Trp276, Phe289, and Val292). In contrast, the major hydrophobic contacts in the trans state involve four residues from the linker (ala223, Ile227, Pro230 and Pro 232) and five from the SH3<sup>C</sup> domain (Phe239, Lys269, Ile270, Trp276, and Leu 294). Moreover, the linker–linker interactions are also different. Pro230 and Pro232 from the second half linker form hydrophobic contacts with the domain in the trans state instead of with Pro225 and Ile227 from the first half linker in the cis state. In addition to these hydrophobic interactions, a hydrogen bond can be found between Asn236 (from the linker) and Gln297 (from the SH3<sup>C</sup> domain) in the trans state. From Figure 4 we can see that this rearrangement of the hydrophobic interactions is associated with a substantial change in the environment of Phe239. Namely, it is exposed to the solvent in the cis state instead of buried in the hydrophobic core in the trans state. As described below the conformation of Phe239 in the cis state promotes the association between the SH3<sup>C</sup> and SH3<sup>N</sup> domain in the case of the two-domain system.



**Figure 4.** Structure comparison of two macrostates of Crk 1-SH3<sup>C</sup> from the metadynamics simulation as shown in Figure 3. (blue: cis state, Red: trans state) (a) The cis–trans isomerization of Pro238 (shown as the blue sticks and balls in the cis state and red in the trans state) results in macrostates with distinct linker–domain interactions. The linkers (residue Gly220 to Pro238) are displayed in light color. (b) The residue Phe239 of trans state (shown as red bonds) is buried into the hydrophobic interaction surface (shown as transparent surface) between the linker (Ile227, Pro230 and Pro232) and the SH3<sup>C</sup> domain (Phe239, Lys269, Ile270, Asn271, Trp276, and Leu294). Instead, it is exposed to solvent in the cis state (shown as blue bonds).

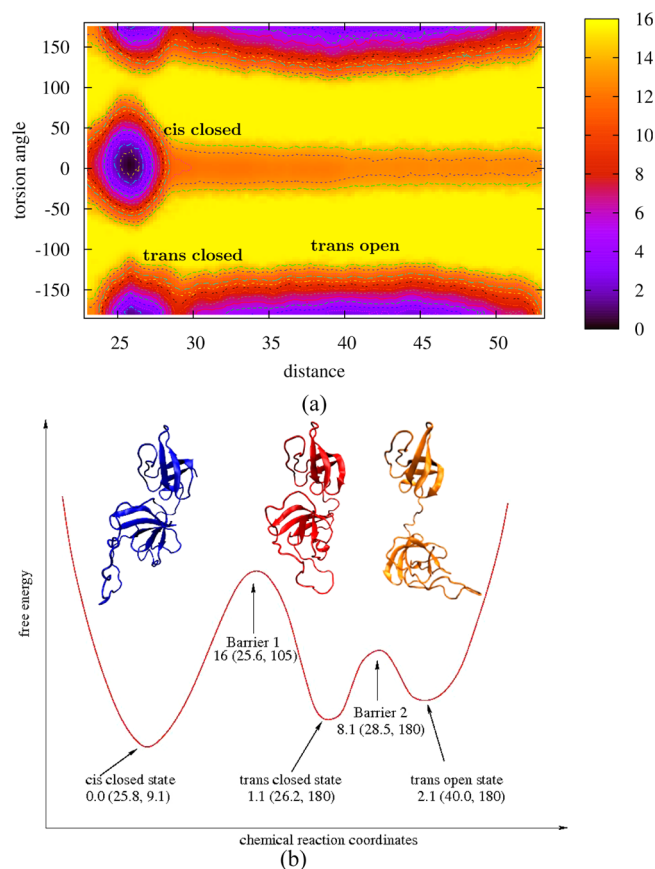
Integrating the PMF in Figure 3 results in the following populations: 38% cis and 62% trans state, which is in reasonable agreement with the NMR results (50%/ 50%). The transition barriers are calculated to be around 17.0 kcal/mol, which is close to the values of short peptides above and consistent with the slow transition rate of cis–trans isomerization.

**Three Macrostates of the Crk Two-Domain SH3<sup>N</sup>-Linker-SH3<sup>C</sup> System.** Previous chemical shift and relaxation analyses<sup>5,6</sup> found that Crk SH3<sup>N</sup>-linker-SH3<sup>C</sup> (Crk<sup>SLS</sup>) exists in two conformations in solution due to the proline isomerization switching of Pro238: a major one (with population of ~90%) in which the switch is in the autoinhibited cis state and the two SH3 domains interact each other (the closed conformation), and a minor one (~10%) in which the switch is in the trans state and the two domains do not have direct interactions. In comparison with the short peptides and the one-domain linker-SH3<sup>C</sup> construct, the population of the cis state becomes predominant in the larger construct, which contains both domains. Due to the interactions between the SH3<sup>C</sup> and SH3<sup>N</sup> domains, the binding of PPII ligands to the SH3<sup>N</sup> domain is blocked, which results in a biologically autoinhibited state. Our simulations provide new insights into the structural transformation between the autoinhibited and uninhibited macrostates.

Figure 6S displays the PMF of Crk SH3<sup>N</sup>-1-SH3<sup>C</sup> along the reaction coordinates including the center of mass distance between the SH3<sup>N</sup> and SH3<sup>C</sup> domain and the  $\omega$  dihedral angle of P238, constructed from the 200 ns standard MD simulation started from the closed cis state (first model of PDB: 2L3S). The cis macrostate is stable during the simulations and no

transition can be observed during the 200 ns period. The residue contact patterns on the map in Figure 7S (especially for the domain–domain interactions) calculated from the MD conformations are very similar to that from NMR models, implying that the extensive domain–domain interactions from NMR are reproduced in the MD simulation.

The 2D PMF from the metadynamics simulation (Figure 5) and the corresponding population analysis by integrating the



**Figure 5.** (a) PMFs of Crk SH3<sup>N</sup>-I-SH3<sup>C</sup> in 2D (the center of mass distance between the SH3<sup>N</sup> and SH3<sup>C</sup> domains and the  $\omega$  dihedral angle of P238) from the 50 ns metadynamics simulation started from the cis state (first model of PDB: 2L3S) and corrected by the 250 ns umbrella sampling simulation with the fixed biasing potential from the metadynamics. Three minima can be found at (25.8, 9.1°), (26.2, 180°), and (40,180°), representing the three macrostates: the cis closed, the trans closed, and the trans open state, respectively. (b) Schematic representation of the free energy path for the cis closed to trans closed and to trans open transition extracted from panel a. The numbers show the free energy values in kcal/mol and the positions in 2D PMF of panel a.

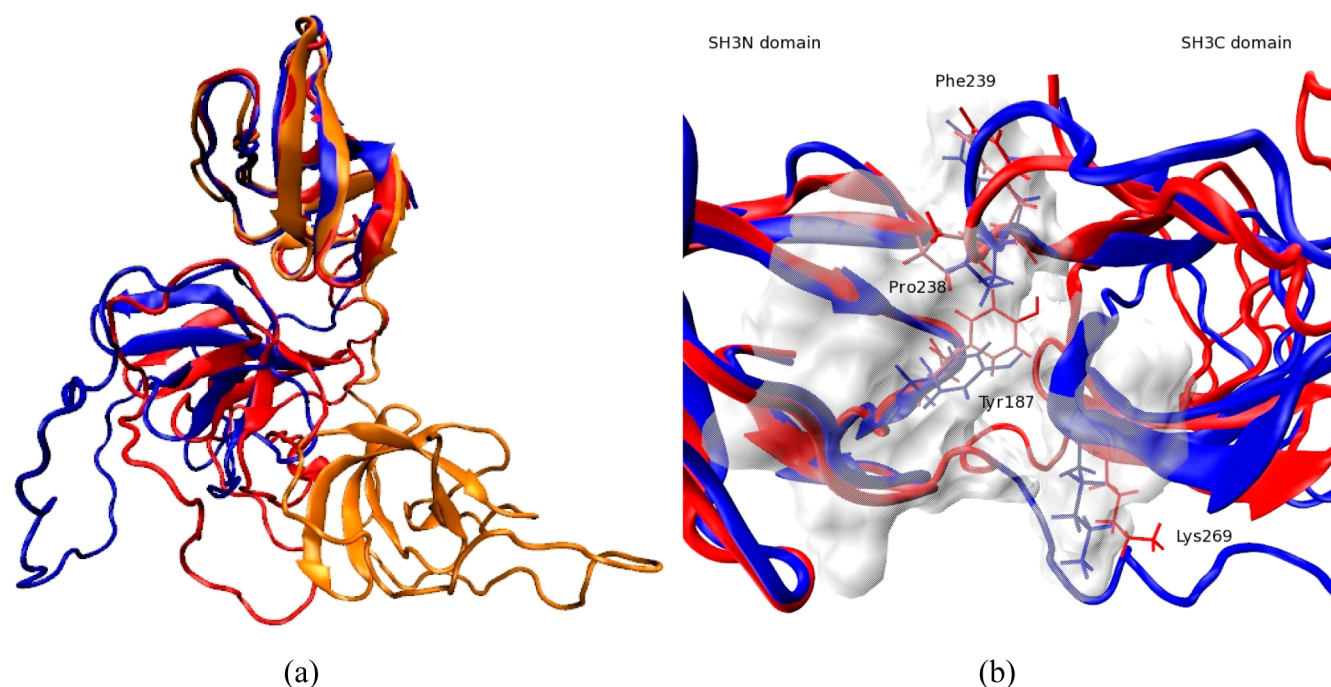
entire basins corresponding to the individual macrostates instead reveal the existence of three macrostates (see Figure 8S for 20 low-energy conformations for each state). The population of the cis state (cis closed/autoinhibited) is 89%, very close to the population derived from NMR data (~90%).<sup>5,6</sup> The trans state has two subgroup distributions: besides one macrostate (4.5%, trans open/uninhibited, already found in NMR) with a wide-distribution of domain–domain distance, we found an intermediate metastable trans state (6.5%, trans closed) with the distribution of domain–domain distance similar to that of the cis closed state (see Figure 5). Although previous NMR experiments on chicken Crk II<sup>5,6</sup> did

not report the existence of a trans closed state, corresponding experiments on human Crk II<sup>63</sup> in fact reported a trans closed state while the cis closed state was not observed. Our results suggest that Crk II can exist in both trans and cis closed states and that the most prevalent state may depend on the species.

**The Two Closed States of Crk<sup>SLS</sup> Have a Similar Domain–Domain Interface but the Flexible Linker Configuration Is Different.** The residue contact maps (Figure 9S) of the cis closed and trans closed states calculated from the simulation snapshots share very similar patterns of domain–domain interactions with that of NMR models of the cis closed state (PDB: 2L3S), indicating a similar domain–domain interface. The detailed structure comparison of the cis closed state between the NMR and metadynamics simulation is shown in Figure 10S. Both of them have the same intermolecular interactions utilizing the hydrophobic contacts which include the canonical binding site of SH3<sup>N</sup> (aromatic residues Phe142, Phe144, Trp170, and Tyr187 along with Pro184 and Pro186) and three other hydrophobic residues from SH3<sup>C</sup> (Pro238, Phe239, and Ile270). In addition, Gln169 in SH3<sup>N</sup> and Lys266 in SH3<sup>C</sup> form a hydrogen bond, and Asp143 in SH3<sup>N</sup> and Lys269 in SH3<sup>C</sup> make a salt bridge in the cis closed state.

Figure 6a shows the structure comparison of three macrostates of Crk<sup>SLS</sup> from the metadynamics simulation. Both closed (cis/trans) states have very similar domain–domain orientations and interactions although the flexible linker has different arrangements. In contrast, the domain–domain interactions are lost in the trans open state, which results in a wide distribution of domain–domain distances (Figure 5a) due to the effects of the flexible linker. Some important differences also exist between the two closed states, which become clearer when the domain–domain interface is examined (Figure 6b) and reweighting is performed to create 1D free energy curves as a function of the number of hydrophobic interactions between the SH3<sup>N</sup> and SH3<sup>C</sup> domains (Figure 7a), the number of hydrogen bonds between Gln169 and Lys266 (Figure 7b), and the distance between Asp143 and Lys269 forming a salt bridge (Figure 7c). From Figure 6b, we can see that the cis to trans transition of Pro238 disturbs the hydrophobic interface induces an outward motion of Tyr 187 from the interface, and reduces the number of hydrophobic contacts between the two domains (Figure 7a). At the same time, the transition also breaks other interactions between two domains including the hydrogen bond between Gln 169 and Lys266 (smaller number in Figure 7b), and the salt bridge between Asp143 and Lys269 (larger distance in Figure 7c). From Figure 6b, it can be seen that Phe239 of the cis closed state is no longer buried into the SH3<sup>C</sup> domain but exposed to SH3<sup>N</sup> as is the cis state of the linker-SH3<sup>C</sup> construct. Such exposure of Phe239 helps to form the hydrophobic contact interface between the two domains as found by the NMR structure analysis.<sup>6</sup> From Figure 7b,c, we can see that the cis closed state is not the global free energy minimum when projected along this reaction coordinate; but this is not consistent with the 2D PMF in Figure 5 and NMR results.<sup>6</sup> This implies the hydrogen bond between Gln169 and Lys266 and the salt bridge between Asp 143 and Lys269 are not the major contributors to stabilize the cis closed state. In contrast, from Figure 7a, it can be seen that the cis closed state is the global minimum, in agreement with the 2D PMF in Figure 5 and NMR results,<sup>6</sup> indicating that the number of hydrophobic contacts between the SH3<sup>C</sup> and SH3<sup>N</sup> domains is





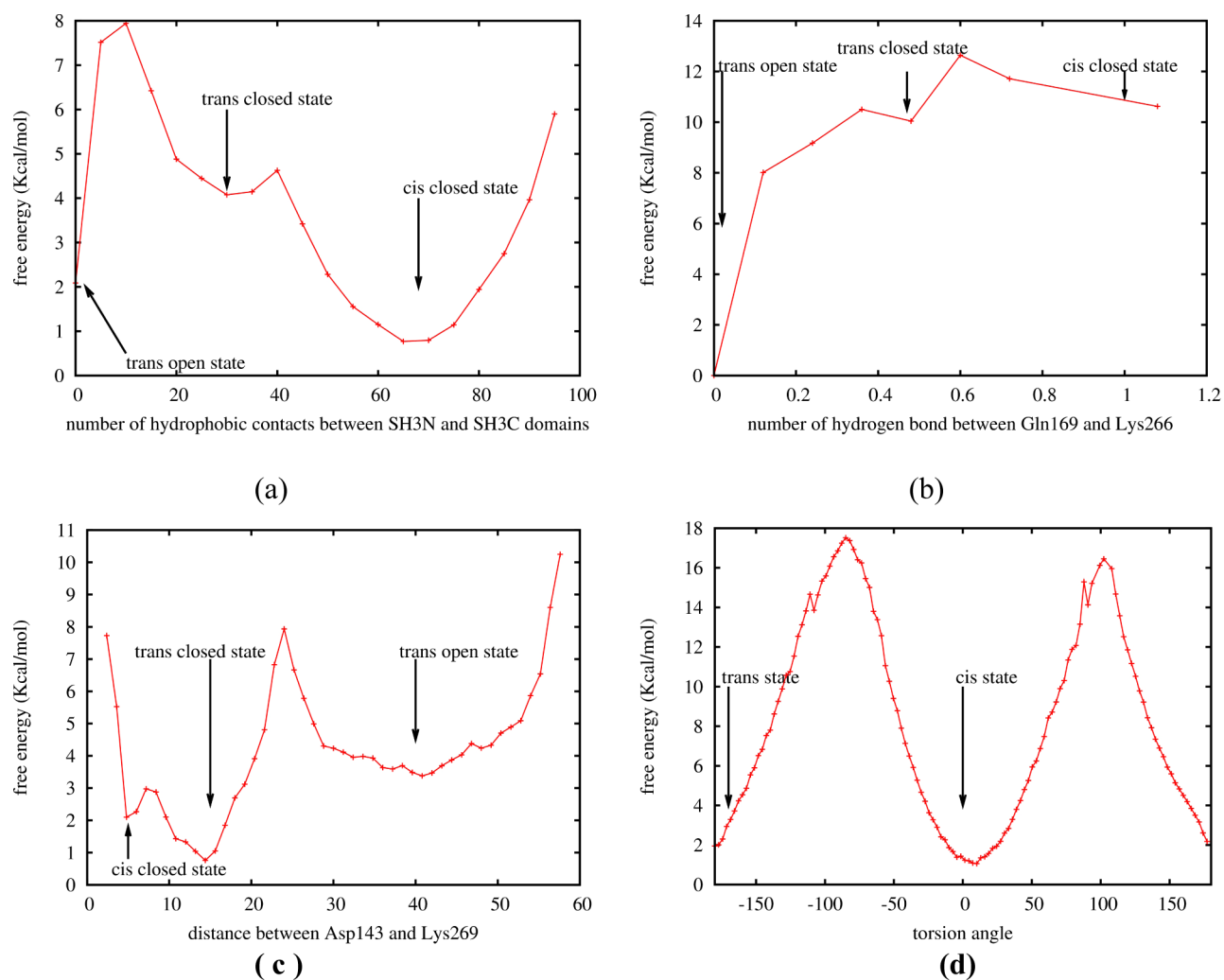
**Figure 6.** Comparison of three structure minima of Crk SH3<sup>N</sup>-1-SH3<sup>C</sup> from the medynamics simulation. (a) The 3D structures representing the three macrostates—the cis closed (blue), the trans closed (red), and the trans open (orange)—were aligned using the SH3<sup>N</sup> domain. (b) Domain–domain interaction surfaces of both the cis closed (blue) and the trans closed (red) state. The transparent surface area shows the hydrophobic interaction regions (for the cis closed state) includes the canonical binding site of SH3<sup>N</sup> (aromatic residues Phe142, Phe144, Trp170, and Tyr187 along with Pro184 and Pro186), three hydrophobic residues from SH3<sup>C</sup> (Pro238, Phe239, and Ile270), two residues (Gln 169 and Lys 266) forming a hydrogen bond, and the other two residues (Asp143 and Lys269) involving in a salt bridge. Tyr187, Pro238, Phe239, and Lys269 are displayed as bonds.

the major contributor to stabilize the cis-closed state. Although the 2D dimensional PMF shown in Figure 5 represents a projection of the many dimensions macrostate transition onto preselected reaction coordinates, the corresponding macrostate populations and transition barriers are in good agreement with that derived from experiments.<sup>6</sup> This suggests that the 2D PMF analysis captures the major characteristics of the cis–trans transition of Crk II.

**Possible Mechanism for the Catalysis of Crk Signaling by Cyp A.** The cis–trans transition of proline can be catalyzed by peptidyl-prolyl isomerases (PPIase)<sup>85,86</sup> such as Pin 1, Cyp A and FKBP. The experimental results<sup>6</sup> show that Cyp A is able to accelerate the transition from the closed to open state for the two domain systems of Crk II Crk<sup>SL5</sup> by several thousand-fold. The mechanisms of catalysis at the molecular level are still not clear although several classical<sup>71,87–89</sup> and quantum simulations<sup>90–93</sup> have been performed for Cyp A and Prolyl isomerases. Previous theoretical studies of Cyp A<sup>71,87–89,91,92,94,95</sup> only focused on short peptide–enzyme complexes and searched for the transition states and paths with low energy barriers from the local structure rearrangements. It is difficult to extend these results to large protein systems that have extensive domain–domain interactions. For the case of Crk II, the domain–domain interactions have to be broken for the transition from the closed to the open state, as shown in Figures 6 and 7.

As shown in Figure 5, the free energy barrier for the transition between the cis closed and trans closed state is  $\sim 16$  kcal/mol and  $\sim 15$  for the reverse transition. The major component of the barrier is the intrinsic proline peptide torsion energy. In contrast, the free energy barrier between the trans

closed to trans open state is much lower ( $\sim 7$  kcal/mol), which comes from the favorable domain–domain interactions mostly due to the dissociation of hydrophobic contacts as analyzed above. There are two possible control points that determine the “timing mechanism” of Crk II proteins: (1) proline isomerization and (2) breaking hydrophobic contacts. This suggests that the mechanism of enzyme catalysis for the Crk proteins occurs in two steps: first a PPIase such as Cyp A acts locally on the Pro238 of the cis closed state, accelerating the cis to trans transition in a way similar to that of short peptides; the transition disturbs the domain–domain interactions and results in an intermediate state (trans closed state). The subsequent relaxation from the trans closed to the trans open state (with a much smaller barrier) breaks the hydrophobic contacts further and completes the transition. The first step is required to overcome the high barrier ( $\sim 16$  kcal/mol) from the torsion energy of  $\omega$  of Pro238 as is the case for short peptides. This step is catalyzed by Cyp A by forming the peptide–enzyme transition state, which lowers the height of the barriers (by 6–13 kcal/mol).<sup>71,87–89,91,92,94,95</sup> In contrast, the second step is required to overcome a barrier ( $\sim 7$  kcal/mol) due to the domain–domain interactions. The need to break the domain–domain interaction (in the second step) determines an intrinsic rate to turn off the autoinhibition. No matter how efficiently a PPIase accelerates the transition from the cis closed to the trans closed state, the relaxation of the domain–domain interactions defines an upper limit to the reaction rate. That is, no enzyme can accelerate switching off the autoinhibition with a rate faster than the self-relaxation rate in the second step. This limit becomes very important when a PPIase can decrease the barrier of proline isomerization below 7 kcal/mol. Furthermore, we



**Figure 7.** 1D PMFs of SH3<sup>N</sup>-1-SH3<sup>C</sup> reweighted from 2D PMF as a function of (a) the number of hydrophobic interactions between the SH3<sup>N</sup> and SH3<sup>C</sup> domains, (b) the number of hydrogen bonds between Gln 169 and Lys266, (c) the distance between Asp143 and Lys269 for a salt bridge, and (d) the dihedral angle  $\omega$  of Pro 238.

suggest that mutations that destabilize the domain–domain association can further accelerate the timing mechanism.

## CONCLUSION

In this work, we provided the results from molecular dynamics and metadynamics simulations for several Crk II protein constructs. The ability to surmount large free energy barriers and accelerate the large conformational changes using metadynamics sampling allow us to investigate the autoregulatory mechanism of Crk, which results from the cis–trans isomerization of the proline switch at Pro 238. The cis–trans transition of Pro 238 induces the linker–domain and domain–domain reorganization and results in three distinct macrostates. The large increase in the cis macrostate population can be explained by linker–linker, linker–domain, and domain–domain interactions, which stabilize the closed cis state primarily by hydrophobic contacts, but also by the formation of hydrogen bonds, and salt bridges. In contrast, the high barrier to the cis–trans transitions is mainly associated with the torsion energy of the intrinsic peptide  $\omega$  dihedral angle of Pro238. For the Crk SH3<sup>N</sup>-linker-SH3<sup>C</sup> construct, our simulations predict an intermediate state, the trans closed state, in which the domain–domain interface is disturbed, the

number of hydrophobic contacts is reduced, and the interdomain hydrogen bond and salt bridge are broken, compared with the stable cis closed (autoinhibited) state. Based on the existence of this intermediate state, we propose that the accelerating process of switching off the autoinhibition by Crp A has two distinct steps: (1) The enzyme binds to the Pro 238 region of the cis closed state catalyzes the transition of the proline switch, and this results in an intermediate state (the trans closed); (2) subsequently the intermediate state relaxes to the trans open states with a much smaller barrier and completes the activation process.

The full Crk II proteins have three domains, SH2, SH3<sup>N</sup>, and SH3<sup>C</sup>. NMR analysis<sup>6</sup> of Chicken Crk SH3<sup>N</sup>-linker-SH3<sup>C</sup> two-domain construct has shown that the SH2 domain does not have significant effects on the other two domains. For Crk II proteins from other species<sup>63,64</sup> and Crk like proteins,<sup>65</sup> this assumption may not be valid. Recent NMR experiments show that Crk like and phosphorylated Crk like proteins have markedly different domain–domain structures and are regulated in a distinct manner, which may involve the SH2 domain.<sup>65</sup> NMR results on mouse Crk II<sup>64</sup> and human Crk II<sup>63</sup> also showed that the autoinhibition mechanism can be accomplished without cis–trans isomerization by combining



the intradomain and interdomain, or linker–domain rearrangements. NMR experiments on human Crk II<sup>63</sup> also revealed the existence of a predominant trans closed state as we have observed for chicken Crk II. Our simulations show that both closed states can coexist in chicken Crk II, and we suggest that the predominant state may depend on the species. Besides these structural reorganizations, phosphorylation of specific tyrosines is an alternative mechanism of autoregulation.<sup>62</sup> Simulations of phosphorylated systems will also shed new light on the mechanisms of autoinhibition.<sup>96</sup>

In this report we were not able to provide the exact time scales for the transitions of Crk II, although no similar results are available in experiments. The metadynamics simulations accelerated the convergence of equilibrium properties, but direct information about the real dynamics is lost. However, the barrier of ~16 kcal/mol for the cis closed to trans closed transition is reasonably close to the experimentally estimated limit for the time scale (>seconds).<sup>6</sup> The substantially reduced barrier for the trans closed to trans open state conformational change is also consistent with the several thousand-fold acceleration of the cis–trans transition by the Cyp A enzyme.<sup>6</sup> To obtain further information about the kinetics and pathways, we plan to build Markov state models starting from the metadynamics simulations in order to generate stochastic trajectories for sampling the transitions,<sup>97,98</sup> which can be used to investigate transition paths systematically for the activation of Crk II.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

Results for standard MD simulations, more low-energy conformations of all macrostates of Crk constructs, and structural comparisons between NMR and simulations are included. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: ronlevy@temple.edu; Phone: 215-204-1927.

### Present Address

†Center for Biophysics & Computational Biology, Department of Chemistry, Temple University, 1901 N. 13th Street, Philadelphia, PA 19122.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work has been supported by a grant from the National Institutes of Health (GM30580) awarded to R.M.L. MD simulations were carried out on the Gyges cluster at Rutgers supported by NIH shared instrumentation Grants Nos. S10-RR022375 and S10-RR027444, and XSEDE resources (TG-MCB100145) awarded to R.M.L. We thank Dr. Emilio Gallicchio for the support of computer resources and helpful discussion. We also thank Dr. Charalampo (Babis) Kalodimos for his central role in helping to motivate this work.

## ■ REFERENCES

(1) Dugave, C.; Demange, L. Cis–trans Isomerization of Organic Molecules and Biomolecules: Implications and Applications. *Chem. Rev.* **2003**, *103*, 2475–2532.

(2) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. Prolyl cis–trans Isomerization as a Molecular Timer. *Nat. Chem. Biol.* **2007**, *3*, 619–629.

(3) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. Proline cis–trans Isomerization and Protein Folding. *Biochemistry* **2002**, *41*, 14637–14644.

(4) Andreotti, A. H. Native State Proline Isomerization: An Intrinsic Molecular Switch. *Biochemistry* **2003**, *42*, 9515–9524.

(5) Sarkar, P.; Saleh, T.; Tzeng, S. R.; Birge, R. B.; Kalodimos, C. G. Structural Basis for Regulation of the Crk Signaling Protein by a Proline Switch. *Nat. Chem. Biol.* **2011**, *7*, 51–57.

(6) Sarkar, P.; Reichman, C.; Saleh, T.; Birge, R. B.; Kalodimos, C. G. Proline cis–trans Isomerization Controls Autoinhibition of a Signaling Protein. *Mol. Cell* **2007**, *25*, 413–426.

(7) Wulf, G.; Finn, G.; Suizu, F.; Lu, K. P. Phosphorylation-Specific Prolyl Isomerization: Is There an Underlying Theme? *Nat. Cell Biol.* **2005**, *7*, 435–441.

(8) Brazin, K. N.; Mallis, R. J.; Fulton, D. B.; Andreotti, A. H. Regulation of the Tyrosine Kinase Itk by the Peptidyl–Prolyl Isomerase Cyclophilin A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1899–1904.

(9) Pastorino, L.; Sun, A.; Lu, P. J.; Zhou, X. Z.; Balastik, M.; Finn, G.; Wulf, G.; Lim, J.; Li, S. H.; Li, X. J.; et al. The Prolyl Isomerase Pin1 Regulates Amyloid Precursor Protein Processing and Amyloid-beta Production. *Nature* **2006**, *404*, 528–534.

(10) Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. Cis–trans Isomerization at a Proline Opens the Pore of a Neurotransmitter-Gated Ion Channel. *Nature* **2005**, *438*, 248–252.

(11) Nelson, C. J.; Santos-Rosa, H.; Kouzarides, T. Proline Isomerization of Histone H3 Regulates Lysine Methylation and Gene Expression. *Cell* **2006**, *126*, 905–916.

(12) Fathers, K. E.; Bell, E. S.; Rajadurai, C. V.; Cory, S.; Zhao, H.; Mourskaia, A.; Zuo, D. M.; Madore, J.; Monast, A.; Mes-Masson, A. M.; et al. Crk Adaptor Proteins Act as Key Signaling Integrators for Breast Tumorigenesis. *Breast Cancer Res.* **2012**, *14*, R74.

(13) Miller, C. T.; Chen, G.; Gharib, T. G.; Wang, H.; Thomas, D. G.; Misek, D. E.; Giordano, T. J.; Yee, J.; Orringer, M. B.; Hanash, S. M.; et al. Increased C-Crk Proto-Oncogene Expression is Associated With an Aggressive Phenotype in Lung Adenocarcinomas. *Oncogene* **2003**, *22*, 7950–7957.

(14) Wang, J.; Che, Y. L.; Li, G.; Shen, T. M.; Wang, H.; Hua, L. H. Crk and CrkL Present with Different Expression and Significance in Epithelial Ovarian Carcinoma. *Mol. Carcinog.* **2011**, *50*, 506–515.

(15) Dai, Y. Q.; Qi, L.; Zhang, X. B.; Li, Y.; Chen, M. F.; Zu, X. B. CrkI And p130Cas Complex Regulates the Migration and Invasion of Prostate Cancer Cells. *Cell Biochem. Funct.* **2011**, *29*, 625–629.

(16) Yanagi, H.; Wang, L.; Nishihara, H.; Kimura, T.; Tanino, M.; Yanagi, T.; Fukuda, S.; Tanaka, S. CRKL Plays a Pivotal Role in Tumorigenesis of Head and Neck Squamous Cell Carcinoma through the Regulation of Cell Adhesion. *Biochem. Biophys. Res. Commun.* **2012**, *418*, 104–109.

(17) Schoetz, G.; Trapp, O.; Schurig, V. Determination of the cis–trans Isomerization Barrier of Several L-Peptidyl-L-Proline Dipeptides by Dynamic Capillary Electrophoresis and Computer Simulation. *Electrophoresis* **2011**, *22*, 2409–2415.

(18) Beausoleil, E.; Lubell, W. D. Steric Effects on the Amide Isomer Equilibrium of Prolyl Peptides. Synthesis and Conformational Analysis of N-Acetyl-5-tert-butylproline N'-Methylamides. *J. Am. Chem. Soc.* **1996**, *118*, 12902–12908.

(19) Kern, D.; Schutkowski, M.; Drakenberg, T. Rotational Barriers of Cis/Trans Isomerization of Proline Analogues and Their Catalysis by Cyclophilin. *J. Am. Chem. Soc.* **1997**, *119*, 8403–8408.

(20) Hinderaker, M. P.; Raines, R. T. An Electronic Effect on Protein Structure. *Protein Sci.* **2003**, *12*, 1188–1194.

(21) Trapp, O.; Trapp, G.; Schurig, V. Determination of the cis–trans Isomerization Barrier of Enalaprilat by Dynamic Capillary Electrophoresis and Computer Simulation. *Electrophoresis* **2004**, *25*, 318–323.

- (22) Melis, C.; Bussi, G.; Lummis, S. C. R.; Molteni, C. Trans-cis Switching Mechanisms in Proline Analogues and Their Relevance for the Gating of the 5-HT<sub>3</sub> Receptor. *J. Phys. Chem. B* **2009**, *113*, 12148–12153.
- (23) Faller, C. E.; Reilly, K. A.; Hills, R. D.; Guvench, O. Peptide Backbone Sampling Convergence with the Adaptive Biasing Force Algorithm. *J. Phys. Chem. B* **2013**, *117*, 518–526.
- (24) Yonezawa, Y.; Shimoyama, H.; Nakamura, H. Multicanonical Molecular Dynamics Simulations Combined with Metadynamics for the Free Energy Landscape of a Biomolecular System with High Energy Barriers. *Chem. Phys. Lett.* **2011**, *501*, 598–602.
- (25) Yonezawa, Y.; Nakata, K.; Sakakura, K.; Takada, T.; Nakamura, H. Intra- and Intermolecular Interaction Inducing Pyramidalization on Both Sides of a Proline Dipeptide during Isomerization: An Ab Initio QM/MM Molecular Dynamics Simulation Study in Explicit Water. *J. Am. Chem. Soc.* **2009**, *131*, 4535–4540.
- (26) Doshi, U.; Hamelberg, D. Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics. *J. Phys. Chem. B* **2009**, *113*, 16590–16595.
- (27) Hamelberg, D.; Shen, T.; McCammon, J. A. Phosphorylation Effects on cis/trans Isomerization and the Backbone Conformation of Serine-Proline Motifs: Accelerated Molecular Dynamics Analysis. *J. Am. Chem. Soc.* **2005**, *127*, 1969–1974.
- (28) Kang, Y. K. Conformational Preferences of Non-Prolyl and Prolyl Residues. *J. Phys. Chem. B* **2006**, *110*, 21338–21348.
- (29) Fischer, S.; Dunbrack, R. L.; Karplus, M. Cis-trans Imide Isomerization of the Proline Dipeptide. *J. Am. Chem. Soc.* **1994**, *116*, 11931–11937.
- (30) Zhong, H. Z.; Carlson, H. A. Conformational Studies of Polyprolines. *J. Chem. Theory Comput.* **2006**, *2*, 342–353.
- (31) Wu, M. H. G.; Deem, M. W. Analytical Rebridging Monte Carlo: Application to cis/trans Isomerization in Proline-Containing, Cyclic Peptides. *J. Chem. Phys.* **1999**, *111*, 6625–6632.
- (32) Valiaev, A.; Lim, D. W.; Oas, T. G.; Chilkoti, A.; Zauscher, S. Force-Induced Prolyl cis-trans Isomerization in Elastin-Like Polypeptides. *J. Am. Chem. Soc.* **2007**, *129*, 6491–6497.
- (33) Valiaev, A.; Lim, D. W.; Oas, T. G.; Clark, R. L.; Chilkoti, A.; Zauscher, S. Force-Induced Proline cis-trans Isomerization in Elastin-Like Polypeptides. *Biophys. J.* **2005**, *88*, 168A.
- (34) Jamet, H.; Jourdan, M.; Dumy, P. NMR and theoretical calculations: A unified view of the cis/trans isomerization of 2-substituted thiazolidines containing peptides. *J. Phys. Chem. B* **2008**, *112*, 9975–9981.
- (35) Kern, D.; Schutkowski, M.; Drakenberg, T. Rotational Barriers of cis/trans Isomerization of Proline Analogues and Their Catalysis by Cyclophilin. *J. Am. Chem. Soc.* **1997**, *119*, 8403–8408.
- (36) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. Side-Chain Effects on Peptidyl-Prolyl cis/trans Isomerisation. *J. Mol. Biol.* **1998**, *279*, 449–460.
- (37) Reimer, U.; Fischer, G. Local Structural Changes Caused by Peptidyl-Prolyl cis/trans Isomerization in the Native State of Proteins. *Biophys. Chem.* **2002**, *96*, 203–212.
- (38) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. Non-Proline cis Peptide Bonds In Proteins. *J. Mol. Biol.* **1999**, *286*, 291–304.
- (39) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. Peptide Bonds Revisited. *Nat. Struct. Biol.* **1998**, *5*, 676–676.
- (40) Kang, Y. K.; Jhon, J. S.; Park, H. S. Conformational Preferences of Proline Oligopeptides. *J. Phys. Chem. B* **2006**, *110*, 17645–17655.
- (41) Moradi, M.; Babin, V.; Roland, C.; Sagui, C. A Classical Molecular Dynamics Investigation of the Free Energy and Structure of Short Polyproline Conformers. *J. Chem. Phys.* **2010**, *133*, 125104–125121.
- (42) Moradi, M.; Babin, V.; Roland, C.; Darden, T. A.; Sagui, C. Conformations and Free Energy Landscapes of Polyproline Peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 20746–20751.
- (43) Pierson, N. A.; Chen, L. X.; Russell, D. H.; Clemmer, D. E. Cis-trans Isomerizations of Proline Residues are Key to Bradykinin Conformations. *J. Am. Chem. Soc.* **2013**, *135*, 3186–3192.
- (44) Pierson, N. A.; Chen, L. X.; Valentine, S. J.; Russell, D. H.; Clemmer, D. E. Number of Solution States of Bradykinin from Ion Mobility and Mass Spectrometry Measurements. *J. Am. Chem. Soc.* **2011**, *133*, 13810–13813.
- (45) Sahoo, H.; Roccatano, D.; Hennig, A.; Nau, W. M. A 10-Angstrom Spectroscopic Ruler Applied to Short Polyprolines. *J. Am. Chem. Soc.* **2007**, *129*, 9762–9772.
- (46) Doose, S.; Neuweiler, H.; Barsch, H.; Sauer, M. Probing Polyproline Structure and Dynamics by Photoinduced Electron Transfer Provides Evidence for Deviations from a Regular Polyproline Type II Helix. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 17400–17405.
- (47) Hodel, A.; Rice, L. M.; Simonson, T.; Fox, R. O.; Brunger, A. T. Proline cis-trans Isomerization in Staphylococcal Nuclease - Multi-Substate Free-Energy Perturbation Calculations. *Protein Sci.* **1995**, *4*, 636–654.
- (48) Birge, R. B.; Kalodimos, C.; Inagaki, F.; Tanaka, S. Crk and CrkL Adaptor Proteins: Networks for Physiological and Pathological Signaling. *Cell Commun. Signal.* **2009**, *7*, 13.
- (49) Feller, S. M. Crk Family Adaptors - Signalling Complex Formation and Biological Roles. *Oncogene* **2001**, *20*, 6348–6371.
- (50) Isakov, N. A New Twist to Adaptor Proteins Contributes To Regulation of Lymphocyte Cell Signaling. *Trends Immunol.* **2008**, *29*, 388–396.
- (51) Pawson, T. Dynamic Control of Signaling by Modular Adaptor Proteins. *Curr. Opin. Cell Biol.* **2007**, *19*, 112–116.
- (52) Rodrigues, S. P.; Fathers, K. E.; Chan, G.; Zuo, D. M.; Halwani, F.; Meterissian, S.; Park, M. CrkL and CrkII Function as Key Signaling Integrators for Migration and Invasion of Cancer Cells. *Mol. Cancer Res.* **2005**, *3*, 183–194.
- (53) Takino, T.; Nakada, M.; Miyamori, H.; Yamashita, J.; Yamada, K. M.; Sato, H. CrkI Adapter Protein Modulates Cell Migration and Invasion in Glioblastoma. *Cancer Res.* **2003**, *63*, 2335–2337.
- (54) Cabodi, S.; Camacho-Leal, M. D.; Di Stefano, P.; Defilippi, P. Integrin Signalling Adaptors: Not Only Figurants in the Cancer Story. *Nat. Rev. Cancer* **2010**, *10*, 858–870.
- (55) Birge, R. B.; Kalodimos, C.; Inagaki, F.; Tanaka, S. Crk and CrkL Adaptor Proteins: Networks for Physiological and Pathological Signaling. *Cell Commun. Signal.* **2009**, *7*, 13.
- (56) Feller, S. M. Crk Family Adaptors - Signalling Complex Formation and Biological Roles. *Oncogene* **2001**, *20*, 6348–6371.
- (57) Feller, S. M.; Knudsen, B.; Hanafusa, H. C-Abl Kinase Regulates the Protein-Binding Activity of C-Crk. *EMBO J.* **1994**, *13*, 2341–2351.
- (58) Groesch, T. D.; Zhou, F.; Mattila, S.; Geahlen, R. L.; Post, C. B. Structural Basis for the Requirement of Two Phosphotyrosine Residues in Signaling Mediated by Syk Tyrosine Kinase. *J. Mol. Biol.* **2006**, *356*, 1222–1236.
- (59) Zhou, S.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J. SH2 Domains Recognize Specific Phosphopeptide Sequences. *Cell* **2004**, *S116*, 767–778.
- (60) Knudsen, B. S.; Feller, S. M.; Hanafusa, H. 4 Proline-Rich Sequences of the Guanine-Nucleotide Exchange Factor C3g Bind with Unique Specificity to the First Src Homology-3 Domain of Crk. *J. Biol. Chem.* **1994**, *269*, 32781–32787.
- (61) Muralidharan, V.; Dutta, K.; Cho, J.; Vila-Perello, M.; Raleigh, D. P.; Cowburn, D.; Muir, T. W. Solution Structure and Folding Characteristics of the C-Terminal SH3 Domain of C-Crk-II. *Biochemistry* **2006**, *45*, 8874–8884.
- (62) Rosen, M. K.; Yamazaki, T.; Gish, G. D.; Kay, C. M.; Pawson, T.; Kay, L. E. Direct Demonstration of an Intramolecular SH2-Phosphotyrosine Interaction in the Crk Protein. *Nature* **1995**, *374*, 477–479.
- (63) Kobashigawa, Y.; Sakai, M.; Naito, M.; Yokochi, M.; Kumeta, H.; Makino, Y.; Ogura, K.; Tanaka, S.; Inagaki, F. Structural for the Transforming Activity of Human Cancer-Related Signaling Adaptor Protein Crk. *Nat. Struct. Mol. Biol.* **2007**, *14*, 503–510.
- (64) Cho, J. H.; Muralidharan, V.; Vila-Perello, M.; Raleigh, D. P.; Muir, T. W.; Palmer, A. G. Tuning Protein Autoinhibition by Domain Destabilization. *Nat. Struct. Mol. Biol.* **2011**, *18*, 550–U163.

- (65) Jankowski, W.; Saleh, T.; Pai, M. T.; Sriram, G.; Birge, R. B.; Kalodimos, C. G. Domain Organization Differences Explain Bcr-Abl's Preference for CrkL over CrkII. *Nat. Chem. Biol.* **2012**, *8*, 590–596.
- (66) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. Long-Timescale Molecular Dynamics Simulations of Protein Structure and Function. *Curr. Opin. Struct. Biol.* **2009**, *19*, 120–127.
- (67) Graves, R.; Baer, M.; Schreiner, E.; Stoll, R.; Marx, D. Conformational Dynamics of Minimal Elastin-Like Polypeptides: The Role of Proline Revealed by Molecular Dynamics and Nuclear Magnetic Resonance. *ChemPhysChem* **2008**, *9*, 2759–2765.
- (68) Oakley, M. T.; Johnston, R. L. Exploring the Energy Landscapes of Cyclic Tetrapeptides with Discrete Path Sampling. *J. Chem. Theory Comput.* **2013**, *9*, 650–657.
- (69) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562–12566.
- (70) Laio, A.; Gervasio, F. L. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71*, 126601–126622.
- (71) Leone, V.; Lattanzi, G.; Molteni, C.; Carloni, P. Mechanism of Action of Cyclophilin A Explored by Metadynamics Simulations. *PLoS Comput. Biol.* **2009**, *5*, e1000309.
- (72) Kutzner, C.; van der Spoel, D.; Fechner, M.; Lindahl, E.; Schmitt, U. W.; de Groot, B. L.; Grubmüller, H. Software News and Update - Speeding up Parallel GROMACS on High-Latency Networks. *J. Comput. Chem.* **2007**, *28*, 2075–2084.
- (73) Scott, W. R. P.; Hunenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.
- (74) 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: Struct., Funct. Bioinf.* **2006**, *65*, 712–725.
- (75) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (76) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, 020603–020606.
- (77) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; et al. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.
- (78) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. Substrate Binding Mechanism of HIV-1 Protease from Explicit-Solvent Atomistic Simulations. *J. Am. Chem. Soc.* **2009**, *131*, 11811–11818.
- (79) Babin, V.; Roland, C.; Darden, T. A.; Sagui, C. The Free Energy Landscape of Small Peptides as Obtained from Metadynamics with Umbrella Sampling Corrections. *J. Chem. Phys.* **2006**, *125*, 204909–204917.
- (80) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations. *Plos Comput. Biol.* **2009**, *5*, e1000452.
- (81) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics. *J. Comput. Chem.* **2009**, *30*, 1615–1621.
- (82) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules 0.1. The Method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (83) Aliev, A. E.; Bhandal, S.; Courtier-Murias, D. Quantum Mechanical and NMR Studies of Ring Puckering and cis/trans-Rotameric Interconversion in Prolines and Hydroxyprolines. *J. Phys. Chem. A* **2009**, *113*, 10858–10865.
- (84) Chen, J.; Edwards, S. A.; Gräter, F.; Baldauf, C. On the cis to trans Isomerization of Prolyl-Peptide Bonds under Tension. *J. Phys. Chem. B* **2012**, *116*, 9346–9351.
- (85) Schiene, C.; Fischer, G. Enzymes that Catalyze the Restructuring of Proteins. *Curr. Opin. Struct. Biol.* **2000**, *10*, 40–45.
- (86) Shaw, P. E. Peptidyl-Prolyl cis/trans Isomerases and Transcription: Is there a Twist in the Tail? *EMBO Rep.* **2007**, *8*, 40–45.
- (87) Hamelberg, D.; McCammon, A. Mechanistic Insight into the Role of Transition-State Stabilization in Cyclophilin A. *J. Am. Chem. Soc.* **2009**, *131*, 147–152.
- (88) Agarwal, P. K.; Geist, A.; Gorin, A. Protein dynamics and enzymatic catalysis: Investigating the peptidyl-prolyl cis–trans isomerization activity of cyclophilin A. *Biochemistry* **2004**, *43*, 10605–10618.
- (89) Trzesniak, D.; van Gunsteren, W. F. Catalytic Mechanism of Cyclophilin As Observed in Molecular Dynamics Simulations: Pathway Prediction and Reconciliation of X-ray Crystallographic and NMR Solution Data. *Protein Sci.* **2006**, *15*, 2544–2551.
- (90) Vohringer-Martinez, E.; Duarte, F.; Toro-Labbe, A. How Does Pin1 Catalyze the cis–trans Prolyl Peptide Bond Isomerization? A QM/MM and Mean Reaction Force Study. *J. Phys. Chem. B* **2012**, *116*, 12972–12979.
- (91) Li, G. H.; Cui, Q. What is so Special about Arg 55 in the Catalysis of Cyclophilin A? Insights from Hybrid QM/MM Simulations. *J. Am. Chem. Soc.* **2003**, *125*, 15028–15038.
- (92) Hur, S.; Bruice, T. C. The Mechanism of cis-trans Isomerization of Prolyl Peptides by Cyclophilin. *J. Am. Chem. Soc.* **2002**, *124*, 7303–7313.
- (93) Fischer, S.; Michnick, S.; Karplus, M. A Mechanism for Rotamase Catalysis by the Fk506 Binding-Protein (Fkbp). *Biochemistry* **1993**, *32*, 13830–13837.
- (94) Piotukh, K.; Gu, W.; Kofler, M.; Labudde, D.; Helms, V.; Freund, C. Cyclophilin A Binds to Linear Peptide Motifs Containing a Consensus that is Present in Many Human Proteins. *J. Biol. Chem.* **2005**, *280*, 23668–23674.
- (95) Howard, B. R.; Vajdos, F. F.; Li, S.; Sundquist, W. I.; Hill, C. P. Structural Insights into the Catalytic Mechanism of Cyclophilin A. *Nat. Struct. Biol.* **2003**, *10*, 475–481.
- (96) Meng, Y.; Roux, B. Locking the Active Conformation of C-Src Kinase through the Phosphorylation of the Activation Loop. *J. Mol. Biol.* **2013**, *426*, 423–435.
- (97) Deng, N.; Zheng, W.; Gallicchio, E.; Levy, R. M. Insights into the Dynamics of HIV-1 Protease: A Kinetic Network Model Constructed from Atomistic Simulations. *J. Am. Chem. Soc.* **2011**, *133*, 9387–9394.
- (98) Xia, J.; Deng, N.; Levy, R. M. NMR Relaxation in Proteins with Fast Internal Motions and Slow Conformational Exchange: Model-Free Framework and Markov State Simulations. *J. Phys. Chem. B* **2013**, *117*, 6625–6634.