Structure
Article

The Atomistic Mechanism of Conformational Transition in Adenylate Kinase: A TEE-REX Molecular Dynamics Study

Marcus B. Kubitzki[1] and Bert L. de Groot[1,*]
[1]Computational Biomolecular Dynamics Group, Max Planck Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany
*Correspondence: bgroot@gwdg.de
DOI 10.1016/j.str.2008.04.013

SUMMARY

We report on an atomistic molecular dynamics simulation of the complete conformational transition of *Escherichia coli* adenylate kinase (ADK) using the recently developed TEE-REX algorithm. Two phases characterize the transition pathway of ADK, which folds into the domains CORE and LID and the AMP binding domain AMPbd. Starting from the closed conformation, half-opening of the AMPbd precedes a partially correlated opening of the LID and AMPbd, defining the second phase. A highly stable salt bridge D118-K136 at the LID-CORE interface, contributing substantially to the total nonbonded LID-CORE interactions, was identified as a major factor that stabilizes the open conformation. Alternative transition pathways, such as AMPbd opening following LID opening, seem unlikely, as full transition events were not observed along this pathway. The simulation data indicate a high enthalpic penalty, possibly obstructing transitions along this route.

INTRODUCTION

Over the last years, experimentally determined structures of protein conformations have become increasingly available. Structures determined with atomic resolution using X-ray crystallography or NMR spectroscopy provide mostly static pictures of different conformational states of proteins. Transitions between such conformations form the functional basis for many proteins, as in motor proteins such as myosin and kinesin, the bacterial flagella, ATP synthase (Gerstein et al., 1994; Berg et al., 2002; Karplus and Gao, 2004), the chaperonin GroEL (Xu et al., 1997), and nuclear transport proteins (Zachariae and Grubmüller, 2006). However, questions related to the underlying transition pathway often remain open (Harrison, 2004). Despite recent advances in time-resolved X-ray crystallography (Moffat, 2003; Schotte et al., 2003), it remains a challenge to elucidate the pathways and mechanisms of protein conformational dynamics.

Among the family of nucleoside triphosphate (NTP) kinases, *Escherichia coli* adenylate kinase (ADK) is a structurally well-studied protein exhibiting large conformational motions crucial for its catalytic function (Müller et al., 1996). It has become a benchmark for studying conformational transitions in proteins (Henzler-Wildman et al., 2007). ADK is a monomeric ubiquitous enzyme that plays a key role in energy maintenance within the cell, controlling cellular ATP levels by catalyzing the reaction $Mg^{2+}\cdot ADP + ADP \leftrightarrow Mg^{2+}\cdot ATP + AMP$. Structurally, the enzyme consists of three domains: the large central "CORE" domain, an AMP binding domain referred to as "AMPbd," and a lid-shaped ATP binding domain termed "LID," which covers the phosphate groups at the active center (Müller et al., 1996). In an unligated structure of ADK the LID and AMPbd adopt an open conformation, whereas they assume a closed conformation in a structure crystallized with the transition-state inhibitor Ap_5A (Müller and Schulz, 1992). Here, the ligands are contained in a highly specific environment required for catalysis. Recent ^{15}N nuclear magnetic resonance spin relaxation studies (Shapiro and Meirovitch, 2006) have shown the existence of catalytic domain motions in the flexible AMPbd and LID domain on a nanosecond timescale, whereas the relaxation in the CORE domain is on a picosecond timescale (Tugarinov et al., 2002; Shapiro et al., 2002).

Molecular dynamics (MD) simulation techniques are widely used to investigate molecular motions at atomic detail on a nanosecond timescale (van Gunsteren and Berendsen, 1990; Hansson et al., 2002; Karplus and McCammon, 2002). For ADK, several computational studies have addressed its large conformational flexibility (Temiz et al., 2004; Lou and Cukier, 2006; Whitford et al., 2007). Maragakis and Karplus (2005) used a coarse-grained plastic network model to generate a minimum energy path between the open and closed crystal structures. Their study predicts that the motion of the LID region precedes the motion of the AMPbd. In a recent study, essential dynamics simulations have been performed to determine whether the closed conformation is accessible to the open unligated enzyme (Snow et al., 2007). The study suggests that the open conformation is at a slightly lower free energy than the closed conformation, as a consequence of hinge motions responsible for domain closure. Very recently, a series of umbrella sampling simulations also suggested a barrier-less transition from the closed to the open conformation in the absence of ligands, with a substantial free-energy difference between the two conformations (Arora and Brooks, 2007).

Although considerable progress has been made toward the understanding of these large conformational changes in ADK, a detailed characterization of the open to closed transition pathway at full atomic resolution has thus far not been achieved. All studies presented so far rely on a preselected pathway along which the transitions were induced. We recently established the highly efficient TEE-REX algorithm (Kubitzki and de Groot,

2007), incorporating the idea of essential dynamics (Amadei et al., 1993) into the replica exchange framework (Sugita and Okamoto, 1999). Tested extensively on two small molecular systems, the algorithm was shown to considerably enhance conformational sampling as compared to conventional MD. Here we employed TEE-REX to facilitate the transition between the open and closed structures of ADK, achieving a full description of the transition pathway and its underlying atomic mechanics without introducing directed bias. In particular, two phases characterize the observed transition pathways between both conformations. Starting from the closed unligated state, a correlated opening of both the AMPbd and LID domains follows a half-opening of the AMPbd during phase 1. Simultaneously, a highly stable salt bridge D118-K136 forms at the LID-CORE interface, contributing significantly to the overall nonbonded interactions of these two domains and thereby stabilizing the open state. Although possible, transition pathways other than the observed ones seem unlikely due to a high enthalpic penalty.

RESULTS

Eight TEE-REX simulations were carried out, four starting from the equilibrated open (named hereafter TRo 1–4) and four from the closed crystal structures (TRc 1–4). Additionally, two reference MD simulations, one of 92 ns and one of 109 ns length, were started from the closed (MDc) and open (MDo) conformation, respectively (see Experimental Procedures and Table 1 for computational details). Two sets of eigenvectors were excited in the TEE-REX simulations. For TRo 1-2 and TRc 1-2, a subset was selected that includes the difference vector between the open and closed X-ray structures, whereas in TRo 3-4 and TRc 3-4 a subset was excited that was solely derived from local fluctuations observed in MD simulations.

Table 1. TEE-REX and MD Simulation Details

Simulation	Starting Structure	Parameters
MDo	Open	109 ns, T = 300K
MDc	Closed, ligand removed	92 ns, T = 300K
TRo 1	Open	42 ns, {κ1,…,κ5}, max{Tes} = 550K
TRo 2	Open	50 ns, {κ1,…,κ5}, max{Tes} = 650K
TRo 3	Open	50 ns, {μ1,…,μ5,ν1,…,ν5}, max{Tes} = 550K
TRo 4	Open	47 ns, {μ1,…,μ5,ν1,…,ν5}, max{Tes} = 650K
TRc 1	Closed, ligand removed	50 ns, {κ1,…,κ5}, max{Tes} = 550K
TRc 2	Closed, ligand removed	50 ns, {κ1,…,κ5}, max{Tes} = 650K
TRc 3	Closed, ligand removed	50 ns, {μ1,…,μ5,ν1,…,ν5}, max{Tes} = 550K
TRc 4	Closed, ligand removed	50 ns, {μ1,…,μ5,ν1,…,ν5}, max{Tes} = 650K

Shown simulation parameters are simulation length (ns); used essential subspace {es}; maximum {es} excitation temperature (K). TEE-REX simulations TRo 1-2 and TRc 1-2 excite the eigenvector directly connecting the open and closed structures (see Experimental Procedures for a detailed definition of used essential subspaces {κi} and {μi,…,νj}).

Conformational Transition

To analyze whether any complete transition between the open and closed conformation was observed, a two-dimensional root-mean-square deviation (rmsd) measure of all backbone atoms was used (see Figure 1), allowing for easy identification of conformational transitions.

In a first step, rmsd curves as a function of time are calculated for both MDc (black) and MDo (red) simulations with respect to their starting structures (see Figure 1A) as well as with regard to the experimental structure that was not the starting structure of the respective MD simulation (i.e., the closed structure for the MDo simulation starting from the open X-ray structure and vice versa; see Figure 1B). Taking the averaged equilibrium fluctuations of 0.3 nm as the transition threshold (Figure 1A), none of the MD simulations shows a complete transition between the two conformations. The MDc simulation approaches the open conformation up to 0.33 nm rmsd briefly during the 92 ns of simulation time (Figure 1B), in accordance with measured rates of ~52 ns for ADK domain motions (Shapiro and Meirovitch, 2006). Therefore, both MD ensembles can be considered localized around both conformations. Constructing the average structure for each MD simulation thus provides a reasonable representation of the open and closed conformation under physiological conditions. In a second step, following the same procedure as above, rmsd curves as a function of time are calculated for both MD simulations using the average <MDc> and <MDo> structures as a reference (Figures 1C and 1D). These two distance measures are combined in a third step to yield the two-dimensional rmsd measure (Figure 1E). Here, the x axis is given by the rmsd with respect to <MDo>, whereas the rmsd with respect to <MDc> defines the y axis. Time information is now implicitly shown in the distribution of the respective MD ensembles. The position of the vertical and horizontal lines, defining the regions of open and closed conformation, are determined by calculating the average distance of each MD ensemble to its average structure (Figure 1C). Within this two-dimensional representation, each trajectory directly connecting the open and closed regions is sampling structures connecting both conformations. In this sense, ensembles possessing this characteristic show transition events.

Within this two-dimensional rmsd representation, both MD ensembles do not show a complete transition event (Figure 1E). Starting from the closed crystal structure (green asterisk), the MDc ensemble exhibits a clear propensity toward the open state (blue diamond). However, no such tendency toward the closed state is seen for the MDo ensemble. Unlike both MD simulations, several complete and independent transition events are observed for TEE-REX simulations TRo 1, TRo 2 (Figures 2A and 2B), and TRc 3 (Figure 3C), as evident from the fact that these simulations directly connect both regions in rmsd space representing the open and closed conformation. Interestingly, the modes excited in the TRc 3 simulation did not include the difference X-ray mode connecting the open and closed experimental structures, whereas both TRo 1 and TRo 2 simulations did excite this mode. The fact that three out of eight TEE-REX simulations

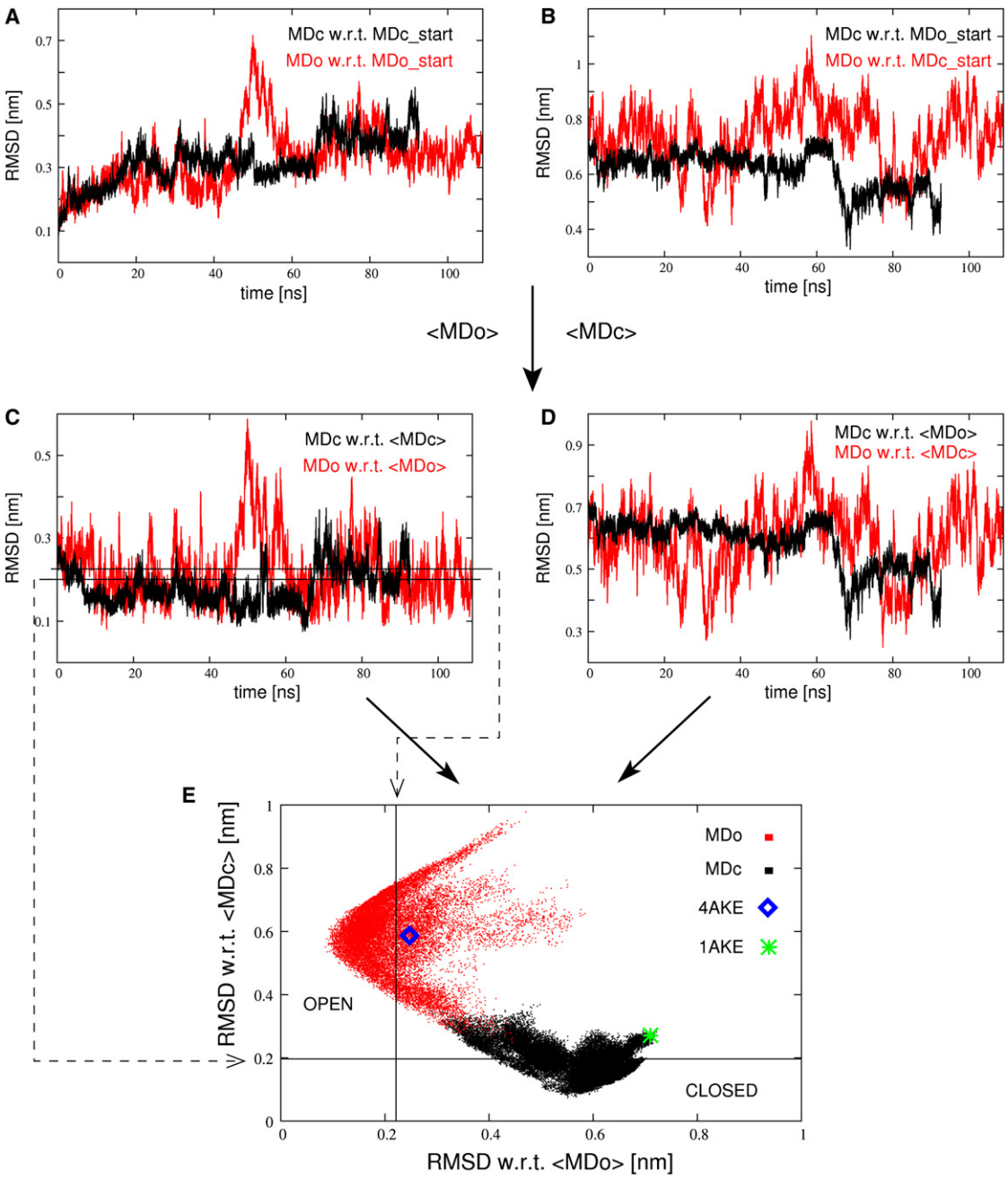

Figure 1. Construction of the Two-Dimensional Root-Mean-Square Deviation Space Used to Measure Conformational Transitions in ADK

In a first step, rmsd's of the MDc (black) (MDo, red) simulation are calculated with regard to the respective MDc (Mdo) starting structure (A) and with respect to the starting structure of the MDo (MDc) simulation (B). Next, this procedure is repeated using as reference the average structure <MDc> (<MDo>) of the MDc (Mdo) (C) and the average structure <MDo> (<MDc>) of the MDo (MDc) simulation (D). Combining (C) and (D) results in the two-dimensional rmsd space (E), displaying the rmsd with respect to the <MDo> structure (x axis) and the rmsd with respect to the <MDc> structure (y axis). Average rmsd distances with respect to the <MDc> and <MDo> structures (C) define the open and closed regions (horizontal and vertical lines in [D]). Conformational transitions are counted for ensembles that cross both lines.

show a complete transition event illustrates the statistical significance of the results. A detailed investigation of the three transition ensembles reveals that within the 50 ns of simulation time, three distinct subensembles are found in the TRo 2 trajectory showing transition events (Figure 2B, black, red, and blue). For the TRc 3 simulation, two distinct subensembles exist (Figure 3C, black and red), further emphasizing the statistical relevance of the results.

A calculation of rmsd curves with regard to the structure that is not the starting structure of the respective simulation (cp. Figure 1B) shows that throughout all MD and TEE-REX simulations a preference for the open conformation is observed (results

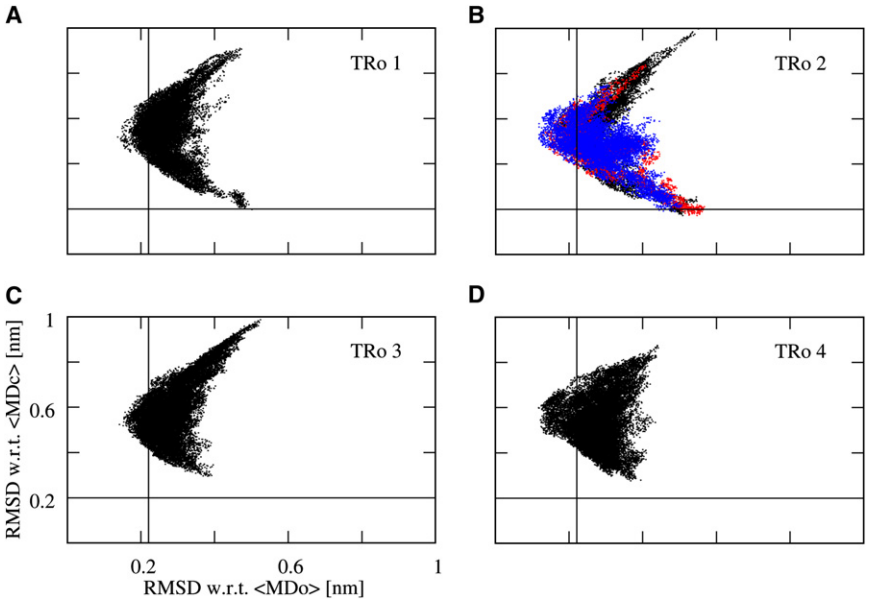

Figure 2. Projection of TRo Simulations into Rmsd Space
(A–D) Distinct TRo 2 ensembles exhibiting transition events (B) are colored (black: 0–19.5 ns; red: 19.5–30 ns; blue: 30–50 ns). For clarity, axis labels are only shown for (C).

not shown), consistent with a lower free energy for the open conformation (Snow et al., 2007; Arora and Brooks, 2007). However, the inability of several TRc and TRo simulations to reach the open and closed conformation, respectively, indicates a free-energy barrier additional to the monotonic profile suggested before (Snow et al., 2007; Arora and Brooks, 2007).

Pathway Characterization

In order to analyze the conformational transition in more detail, the TRc 3 ensemble (Figure 2B, black) was used as a representative pathway, as all transitions exhibited similar characteristics. A pathway was constructed from this ensemble using an rmsd distance measure (see Experimental Procedures for details and Movie S1 available online for the transition pathway).

Figure 4 shows the complete pathway (yellow line) overlaying the TRc 3 ensemble (black dots) as well as the MDo (red dots) and MDc (cyan dots) ensembles. Two principal component analyses (PCA) were carried out on the TRc 3 ensemble to define the x and y coordinates. From the first PCA, performed on all backbone atoms of the CORE domain and the AMPbd, the first eigenvector describing the AMPbd motion with respect to the CORE defines the x axis. Similarly, the first eigenvector of the second PCA, performed on all backbone atoms of the CORE and the LID domains and depicting the LID motion with respect to the CORE, defines the y axis. The crystal structure of the open conformation (Protein Data Bank [PDB] ID code: 4AKE) is indicated by a green square, together with four ribbon representations visualizing different structures along the pathway (Figure 4, structures A–D, magenta triangles). Not shown is the closed crystal structure (PDB ID code: 1AKE, ligand removed) because it almost coincides with the lower right triangle at one end of the pathway (Figure 4, structure D). Secondary structure assignments were taken from Müller and Schulz (1992).

A secondary structure analysis of the structures along the pathway was performed using DSSP (Kabsch and Sander, 1983) to check for structural stability during the transition. For the CORE (residues 1–29, 60–121, and 160–214) and AMPbd

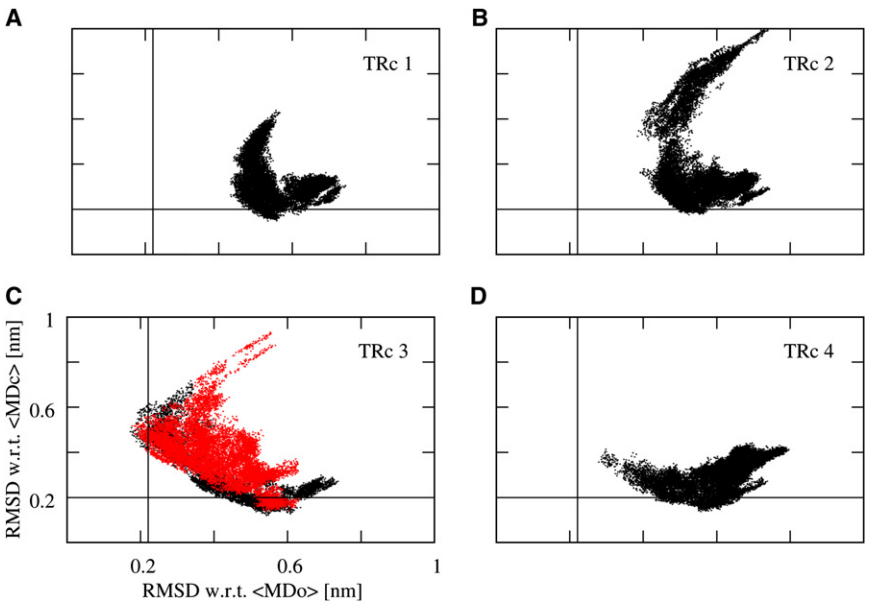

Figure 3. Projection of TRc Simulations into Rmsd Space
(A–D) Distinct TRc 3 ensembles showing transition events (C) are colored (black: 0–20 ns; red: 20–50 ns). For clarity, axis labels are only shown for (C).

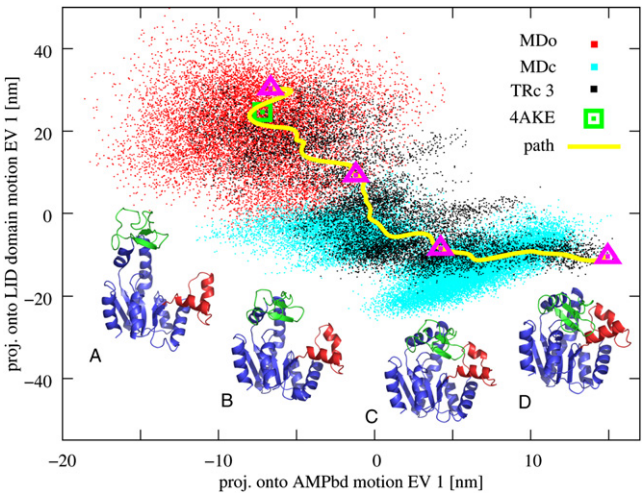

Figure 4. ADK Transition Pathway

Two-dimensional projection of the complete transition path (yellow line), overlaying the TRc 3 (black dots), MDo (red dots), and MDc (cyan dots) ensembles, as well as the crystal structure of the open conformation (PDB ID code: 4AKE; green square). Colored ribbon representations visualize different structures (insets A–D) along the pathway (magenta triangles). Within all representations, the CORE domain is drawn in blue, LID in green, and the AMPbd is depicted in red. The x and y axes are defined via a principal component analysis on the TRc 3 ensemble. The x axis is given by the first eigenvector (EV) describing the AMPbd motion with respect to the CORE. Similarly, the first EV depicting the LID motion with regard to the CORE defines the y axis.

(residues 30–59), no significant change in secondary structure is seen despite the large conformational change of the latter. As for the LID (residues 122–159), both β sheets are stable, with only small conversions among residues constituting bends and turns of the domain. Overall, ADK strongly maintains its integrity, showing only minute changes in secondary structure. Thus, the system essentially behaves like a rigid body with flexible domains.

Unligated ADK can sample a wide range of conformations between the open and closed structure, offering ligands several favorable structures for binding. This result is indicative of the "conformational selection" view of ligand binding proposed for ADK (Zhang et al., 1998; Han et al., 2002; Eisenmesser et al., 2005; Tobi and Bahar, 2005; Henzler-Wildman et al., 2007). The sampled regions of configurational space suggest a preferential transition pathway of the unligated enzyme. The complete transition from the unligated closed (Figure 4, structure D) to the open conformation (structure A) can be characterized by two phases. During the first phase (Figure 4, structures C and D), the LID remains essentially closed whereas the AMPbd, comprising helices α2 and α3, assumes a half-open conformation (cp. Figure 4, structure C). In doing so, α2 bends toward helix α4 of the CORE by 15° with respect to α3. This opening of the AMP binding cleft could facilitate an efficient release of the formed product. For the second phase, a partially correlated opening of the LID domain together with the AMPbd is observed. Halfway toward the open conformation (structure A), an intermediate half-open structure (transition state B) at the contact interface of both MDc and MDo ensembles is visited. During ~100 ns of simulation time, both MD simulations were unable to pass beyond this interface in either direction, suggesting a substantial free-energy barrier along the pathway.

A domain motion analysis using DynDom (Hayward and Berendsen, 1998) of the second phase of the transition shows that the LID opening motion cannot be described by a pure hinge-bending motion. Only in the last part of phase 2 (going from B to A) the LID motion follows a pure hinge-bending motion of ~30°, with the hinge axis given by residues L115, I116 and R167, L168. At the beginning of phase 2, a combination of a hinge-bending motion and an outward translation away from the AMPbd characterizes the pathway toward transition state B.

For phase 1, an opening motion of the AMPbd is found, with bending residues S30, T31, and K69-R71. An interesting observation in this respect is the formation of a highly stable salt bridge D118-K136, connecting the LID and CORE domains. Estimating the total nonbonded interaction between the LID and CORE, it was found that this salt bridge contributes substantially to the interaction of the two domains. From a comparison of 14 PDB structures from yeast, maize, human, and bacterial adenylate kinase, 11 structures feature such a salt-bridge motif at the LID-CORE interface.

Alternative Pathways

The observed transition pathway of the TRc 3 ensemble corresponds to a particular sequence of events: starting from the closed conformation, a subsequent half-opening of the AMPbd is followed by a partially correlated opening of the LID/AMPbd complex. It is an interesting question whether this pathway is the only possibility or whether alternative pathways are also possible.

An investigation of all TEE-REX simulations, conducted within the subspace shown in Figure 4, reveals that simulations TRc 1 and TRc 2 sample the first part of an alternative transition pathway: starting from the closed state, the LID opens independently from the AMPbd, with the latter remaining between the closed and a half-open conformation. Inside the depicted PCA subspace of Figure 4, this motion corresponds to a basically vertical path parallel to the y axis, with the starting point being the closed conformation. To complete a transition along this alternative pathway, the AMPbd would be required to fully open up. Yet, both the TRc 1 and TRc 2 simulations fail to complete this last step, presumably indicating a substantial free-energy barrier in this part of configuration space.

A comparison of the final structures from the TRc 2 and TRc 3 simulations shows that conformational deviations are located in helix α3 of the AMPbd, LID, helix α5 of the CORE domain as well as helices α6 and α7 connecting the LID and CORE domain. To investigate structural features in more detail, the pathway obtained from the TRc 2 ensemble was compared to the transition pathway of the TRc 3 ensemble (cp. Figure 5). After initial relaxation, the TRc 2 pathway diverges from the TRc 3 case. In the former, the LID domain fully opens, with the AMPbd remaining closed. Here, the LID domain motion occurs independently from the AMPbd, in contrast to the observed transition in TRc 3, where opening of the AMPbd precedes LID movement. Whereas the LID assumes a half-open conformation in TRc 3, the AMPbd gains flexibility via bending of helix α2 toward α4 of the CORE, resembling the motion encountered in the transition pathway. However, a full opening of the AMPbd is prevented

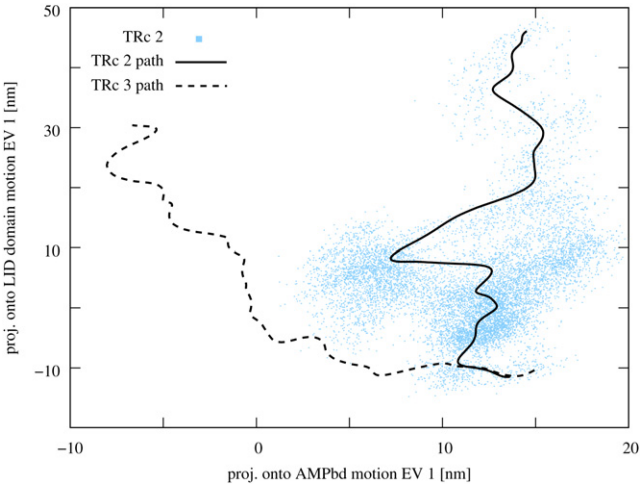

Figure 5. Alternative ADK Transition Pathways

Two-dimensional projection of the complete TRc 3 transition path (dashed line) and the TRc 2 pathway (solid line), overlaying the TRc 2 ensemble (gray dots).

by helix α3, keeping its relative position to helix α7 throughout the TRc 2 pathway. LID flexibility is comparable for both pathways, although end configurations assume a slightly different aperture angle with regard to the CORE domain. A secondary structure assignment of the TRc 2 pathway using DSSP shows larger fluctuations compared to the TRc 3 transition, the differences concentrating in the CORE and AMPbd (data not shown). For the latter, residues of helices α2 and α3 adjacent to the helix-connecting loop exhibit the largest structural variability.

DISCUSSION

Using the recently developed TEE-REX algorithm, we successfully simulated, to our knowledge for the first time, the complete domain conformational transition of *E. coli* adenylate kinase at full atomic resolution without introducing directed bias. In contrast to coarse-grained models (Temiz et al., 2004; Maragakis and Karplus, 2005; Whitford et al., 2007) originally designed for such tasks, TEE-REX combines the advantage of atomic detail with a highly efficient and ensemble-preserving algorithm.

From a series of eight TEE-REX simulations, complemented by two MD references, a possible transition pathway was found. Transitions were induced without a directed bias, because in the TRc 3 simulation, showing two full transition events, the temperature-enhanced subspace did not contain the transition mode connecting the open and closed conformations (see Table 1). The pathway can be characterized by two phases. Starting from the closed conformation, a half-opening of the AMPbd is followed by a partially correlated opening motion of the LID/AMPbd complex toward the open state. This sequence of events exceeds findings of a study by Maragakis and Karplus (2005). From a minimum free-energy path, calculated using a coarse-grained model of ADK, Maragakis and Karplus found that LID motion occurs independently of AMPbd motion.

Along the transition pathway, two prominent features were identified. First, during phase 1, opening of the AMPbd occurs via bending of the α2 helix toward α4 of the CORE domain by approximately 15° with respect to helix α3. This opening of the AMP binding cleft might be involved in facilitating an efficient release of the formed product after catalysis. However, because all simulations were carried out in the absence of any ligand, no conclusions can be drawn with respect to ligand behavior. Second, a stable salt bridge, D118-K136, connecting the LID and CORE domains forms that strongly contributes to the total enthalpic interaction between the two domains, suggesting a stabilizing function for the open conformation. The occurrence of such a salt-bridge motif in several adenylate kinase structures of different species supports the hypothesis. Breaking this salt bridge via mutation, such as D118A, should thus decrease the stability of the open state.

From our simulation data, a qualitative picture of the underlying free-energy landscape of unligated ADK can be drawn. All TEE-REX and MD simulations, starting from the closed crystal structure, show a preference for the open conformation, whereas no such preference for the closed state is seen for simulations starting from the open state. This finding is consistent with declining free-energy profiles deduced from both essential dynamics and umbrella sampling simulations that induced the transition from the closed to the open conformation (Snow et al., 2007; Arora and Brooks, 2007). However, the inability of both MD and several TRc and TRo simulations to reach the open and closed structure, respectively, indicates a free-energy barrier for the spontaneous pathway revealed here, in addition to the previously suggested monotonic profiles. In a recent study, Lou and Cukier (2006) found a stable conformation similar to the closed state using a distance replica exchange method. However, conclusions regarding the relative free-energy difference between the open and the closed states could not be drawn. Taking all TEE-REX and MD simulation data together, the following picture emerges supporting the conformational selection view of ligand binding: in equilibrium, unligated ADK can sample both—open and closed—conformations, as observed, for example, in the transition of the TRc 3 simulation from the closed to the open structure using no directed bias. Here, the closed state has a higher free energy with regard to the open state. Upon ligand binding, the closed structure is stabilized by protein-ligand interactions for catalysis. Recent experimental results on *Aquifex aeolicus* ADK further this view (Henzler-Wildman et al., 2007). From the behavior of both MD and TEE-REX simulations, the coarse location in configuration space of the additional free-energy barrier can be estimated, corresponding to a half-open conformation of ADK.

Transition pathways other than the one characterized seem possible, as an analysis of all TEE-REX simulations suggests. Although a complete transition was not observed, an independent opening of the LID domain with respect to the AMPbd was found. In both pathways the characteristic half-opening of the AMPbd and the stable salt-bridge motif are present, underlining their significance for the atomistic mechanics of the transition. Full opening of the AMPbd would complete the alternative route. However, this was not observed, possibly indicating a free-energy barrier in this part of configuration space. Together with the observed larger fluctuations in secondary structure elements, indicating high internal strain energies, the enthalpic penalty along this route possibly renders it unfavorable as a transition pathway of ADK. However, despite the fact that no full transition events

were observed along this pathway, it can presently not be ruled out that transitions also occur along this pathway.

EXPERIMENTAL PROCEDURES

All simulations were carried out using the MD software package GROMACS 3.3.1 (Lindahl et al., 2001), supplemented by the TEE-REX module (Kubitzki and de Groot, 2007). The OPLS all-atom force field (Jorgensen et al., 1996) was used for the protein and TIP4P was used as a water model (Jorgensen et al., 1983). All simulations were performed in the NPT ensemble. The pressure was coupled to a Berendsen barostat (Berendsen et al., 1984) with a coupling constant of 1.0 ps and an isotropic compressibility of 4.5×10^{-5} bar^{-1}. All bonds were constrained using the LINCS algorithm (Hess et al., 1997). An integration time step of 2 fs was used. Lennard-Jones interactions were calculated with a 10 Å cutoff. Electrostatic interactions were calculated explicitly at a distance smaller than 10 Å; long-range electrostatic interactions were calculated by particle-mesh Ewald summation (Darden et al., 1993) with a grid spacing of 0.12 nm and fourth-order B spline interpolation.

MD Setup

The temperature in all MD simulations was kept constant at T = 300K by coupling to an isotropic Berendsen thermostat (Berendsen et al., 1984) with a coupling time of 0.01 ps for the MDc and τt = 0.1 ps for the MDo simulations.

The MDc simulation system was set up as follows. From the protonated crystal structure (Müller and Schulz, 1992; PDB ID code: 1AKE), the two-substrate-mimicking inhibitor P1,P5-bis(adenosine-5′)pentaphosphate (Ap_5A) was removed. The protein was then solvated in a rhombic dodecahedral box with box vectors of 74.775 Å length. The system comprised 37,965 atoms. Four Na^+ ions were added to neutralize the system. The energy of the solvated system was minimized using the steepest descent algorithm. Subsequently, a 100 ps MD simulation at the target temperature was carried out using harmonic position restraints on the heavy atoms of the protein with a force constant of k = 1000 kJ mol^{-1} nm^{-2} to equilibrate water and ions. Next, a trajectory of 109 ns length was produced by a free (unbiased) MD simulation. Structures were recorded every 1 ps for subsequent analysis.

For the MDo simulation system, the protonated crystal structure (Müller et al., 1996; PDB ID code: 4AKE) was solvated in a rectangular box with a size of 63.309 Å × 83.52 Å × 77.031 Å. The system comprised 53,195 atoms. To neutralize the system, 38 Na^+ and 35 Cl^- ions were added. Energy minimization of the solvated system using the steepest descent algorithm was followed by a 500 ps MD simulation at the target temperature using harmonic position restraints on the heavy atoms of the protein with a force constant of k = 1000 kJ mol^{-1} nm^{-2} to equilibrate water and ions. Subsequently, a trajectory of 92 ns length was produced by a free MD simulation. Structures were recorded every 3 ps for subsequent analysis.

TEE-REX Algorithm and Setup

In addition to the replica framework (Sugita and Okamoto, 1999), the concept of an essential subspace {es} is central to the TEE-REX algorithm. A few collective coordinates (called eigenvectors), obtained by a principal component analysis (PCA), are sufficient to describe the major functional motions of a protein (Amadei et al., 1993). Within TEE-REX, a low-dimensional {es} is excited using a higher temperature while the remaining degrees of freedom are kept at the reference temperature. Frequent exchanges between all replicas of the system allow the unbiased reference replica to sample an extensive volume of the configuration space, while approximately preserving Boltzmann statistics.

Eight 20 ns TEE-REX simulations were performed, starting from the equilibrated open (TRo 1–4) and closed (TRc 1–4) starting structures used in the MDo and MDc simulations, respectively. Each TEE-REX simulation consisted of three replicas, having essential subspace temperatures of 300K and 320K for the reference (replica zero) and first excited replica, respectively. Excitation temperatures of 550K or 650K were used for the second replica.

Three different eigenvector sets were used in the construction of the essential subspaces {es}. A PCA was performed on the first 5 ns of the MDo simulation, taking all backbone atoms into account. The first five eigenvectors, describing 92% of all backbone fluctuations, defined the essential subspace $\{\mu 1,\ldots,\mu 5\}$. Repeating this procedure for the first 5 ns of the MDc simulation yielded the eigenvector set $\{\nu 1,\ldots,\nu 5\}$, describing 92% of all backbone fluctuations. A PCA on the combined MDo and MDc ensemble results in {es} = $\{\kappa 1,\ldots,\kappa 5\}$. The latter was used for simulations TRo 1-2 and TRc 1-2. Hereby, the eigenvector directly connecting the open and the closed structures is also excited. For the second set of eigenvectors, {es} = $\{\mu 1,\ldots,\mu 5,\nu 1,\ldots,\nu 5\}$, used in simulations TRo 3-4 and TRc 3-4, only eigenvectors describing the local fluctuations of the respective conformation as extracted from the respective simulations are excited. The used essential subspaces were coupled to a Berendsen thermostat with a coupling time of 0.05 ps, with all other degrees of freedom coupled with a coupling constant of 0.1 ps. Exchanges between replicas were attempted every 160 ps and were accepted with >95% probability. Structures were recorded every 2 ps for further analysis. A summary of all parameter combinations used in this work is given in Table 1.

Construction of the Transition Pathway

Construction of a pathway between the open and closed conformations of ADK for the TRc 3 simulation is based on the following idea: to get from point A to some distant point B requires increasing the distance from A while at the same time approaching the target point B (thereby neglecting detour routes). The backbone rmsd curves with respect to the open (rmsdo) and closed (rmsdc) structures are used as the distance measure. In a first step, rmsd differences rmsdo − rmsdc were calculated for each frame and sorted in decreasing order. A large distance in the rmsd space thereby corresponds to a structure close to the beginning of the path, whereas a low-rmsd distance points to a structure in the vicinity of the target conformation. Representative structures for a given number of evenly spaced waypoints in the constructed rmsd distance space were chosen to visualize the pathway (cp. Figure 4).

To probe for possible detour routes neglected in the algorithm described above, the TRc 3 ensemble was projected into the plane spanned by the two rmsd coordinates given in Figures 1A and 1B, that is, the rmsd with respect to the open and closed structures, respectively. A narrow distribution along the diagonal connecting both end states of the pathway is found, implying only very limited variability for ADK along the transition path (results not shown).

SUPPLEMENTAL DATA

Supplemental Data include a movie of the closed to open transition of unligated ADK and can be found with this article online at http://www.structure.org/cgi/content/full/16/8/1175/DC1/.

ACKNOWLEDGMENTS

The authors thank Uli Zachariae for very stimulating discussions, Jürgen Haas for kindly providing both MD trajectories of ADK, and Ira Tremmel for carefully reading the manuscript.

Received: August 22, 2007
Revised: April 28, 2008
Accepted: April 30, 2008
Published: August 5, 2008

REFERENCES

Amadei, A., Linssen, A.B.M., and Berendsen, H.J.C. (1993). Essential dynamics of proteins. Proteins *17*, 412–425.

Arora, K., and Brooks, C.L., III. (2007). Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism. Proc. Natl. Acad. Sci. USA *104*, 18496–18501.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. J. Chem. Phys. *81*, 3684–3690.

Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002). Biochemistry, Fifth Edition (New York: W.H. Freeman).

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald—an N log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Eisenmesser, E.Z., Millet, O., Labeikovsky, W., Korzhnev, D.M., Wolf-Watz, M., Bosco, D.A., Skalicky, J.J., Kay, L.E., and Kern, D. (2005). Intrinsic dynamics of an enzyme underlies catalysis. Nature *438*, 117–121.

Gerstein, M., Lesk, A.M., and Chothia, C. (1994). Structural mechanisms for domain movements in proteins. Biochemistry *33*, 6739–6749.

Han, Y., Li, X., and Pan, X.M. (2002). Native states of adenylate kinase are two active sub-ensembles. FEBS Lett. *528*, 161–165.

Hansson, T., Oostenbrink, C., and van Gunsteren, W.F. (2002). Molecular dynamics simulations. Curr. Opin. Struct. Biol. *12*, 190–196.

Harrison, S.C. (2004). Whither structural biology? Nat. Struct. Mol. Biol. *11*, 12–15.

Hayward, S., and Berendsen, H.J.C. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. Proteins *30*, 144–154.

Henzler-Wildman, K.A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M.A., Petsko, G.A., Karplus, M., et al. (2007). Intrinsic motions along an enzymatic reaction trajectory. Nature *450*, 838–844.

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

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

Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996). Development and testing of the OPLS-AA force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. *118*, 11225–11236.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.

Karplus, M., and Gao, Y.Q. (2004). Biomolecular motors: the F1-ATPase paradigm. Curr. Opin. Struct. Biol. *14*, 250–259.

Karplus, M., and McCammon, J.A. (2002). Molecular dynamics simulations of macromolecules: a perspective. Nat. Struct. Biol. *9*, 646–652.

Kubitzki, M.B., and de Groot, B.L. (2007). Molecular dynamics simulations using temperature-enhanced essential dynamics replica exchange. Biophys. J. *92*, 4262–4270.

Lindahl, E., Hess, B., and van der Spoel, D. (2001). GROMACS 3.0: a package for molecular simulation and trajectory analysis. J. Mol. Model. *7*, 306–317.

Lou, H., and Cukier, R.I. (2006). Molecular dynamics of apo-adenylate kinase: a distance replica exchange method for the free energy of conformational fluctuations. J. Phys. Chem. B *110*, 24121–24137.

Maragakis, P., and Karplus, M. (2005). Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase. J. Mol. Biol. *352*, 807–822.

Moffat, K. (2003). The frontiers of time-resolved macromolecular crystallography: movies and chirped X-ray pulses. Faraday Discuss. *122*, 65–77, discussion 79–88.

Müller, C.W., and Schulz, G.E. (1992). Structure of the complex between adenylate kinase from *Escherichia coli* and the inhibitor Ap5A refined at 1.9 Å resolution: a model for a catalytic transition state. J. Mol. Biol. *224*, 159–177.

Müller, C.W., Schlauderer, G., Reinstein, J., and Schulz, G.E. (1996). Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. Structure *4*, 147–156.

Schotte, F., Lim, M., Jackson, T.A., Smirnov, A.V., Soman, J., Olson, J.S., Phillips, G.N., Jr., Wulff, M., and Anfinrud, P.A. (2003). Watching a protein as it functions with 150 ps time-resolved X-ray crystallography. Science *300*, 1944–1947.

Shapiro, Y.E., and Meirovitch, E. (2006). Activation energy of catalysis-related domain motion in *E. coli* adenylate kinase. J. Phys. Chem. B *110*, 11519–11524.

Shapiro, Y.E., Kahana, E., Tugarinov, V., Liang, Z., Freed, J.H., and Meirovitch, E. (2002). Domain flexibility in ligand-free and inhibitor-bound *Escherichia coli* adenylate kinase based on a mode-coupling analysis of ^{15}N spin relaxation. Biochemistry *41*, 6271–6281.

Snow, C., Qi, G., and Hayward, S. (2007). Essential dynamics sampling study of adenylate kinase: comparison to citrate synthase and implication for the hinge and shear mechanisms of domain motion. Proteins *67*, 325–337.

Sugita, Y., and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. Chem. Phys. Lett. *314*, 141–151.

Temiz, N.A., Meirovitch, E., and Bahar, I. (2004). *Escherichia coli* adenylate kinase dynamics: comparison of elastic network model modes with mode-coupling ^{15}N-NMR relaxation data. Proteins *57*, 468–480.

Tobi, D., and Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. Proc. Natl. Acad. Sci. USA *102*, 18908–18913.

Tugarinov, V., Shapiro, Y.E., Liang, Z., Freed, J.H., and Meirovitch, E. (2002). A novel view of domain flexibility in *E. coli* adenylate kinase based on structural mode-coupling (15)N NMR relaxation. J. Mol. Biol. *315*, 155–170.

van Gunsteren, W.F., and Berendsen, H.J.C. (1990). Computer simulations of molecular dynamics: methodology, applications and perspectives in chemistry. Angew. Chem. Int. Ed. Engl. *29*, 992–1023.

Whitford, P.C., Miyashita, O., Levy, Y., and Onuchic, J.N. (2007). Conformational transitions of adenylate kinase: switching by cracking. J. Mol. Biol. *366*, 1661–1671.

Xu, Z., Horwich, A.L., and Sigler, P.B. (1997). The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature *388*, 741–750.

Zachariae, U., and Grubmüller, H. (2006). A highly strained nuclear conformation of the exportin Cse1p revealed by molecular dynamics simulations. Structure *14*, 1469–1478.

Zhang, H.J., Sheng, X.R., Niu, W.D., Pan, X.M., and Zhou, J.M. (1998). Evidence for at least two native forms of rabbit muscle adenylate kinase in equilibrium in aqueous solution. J. Biol. Chem. *273*, 7448–7456.

