Nucleotide-dependent conformational states of actin

Jim Pfaendtner[a,b], Davide Branduardi[b], Michele Parrinello[b], Thomas D. Pollard[c], and Gregory A. Voth[a,1]

[a]Center for Biophysical Modeling and Simulation and Department of Chemistry, University of Utah, Salt Lake City, UT 84112-0850; [b]Computational Science, Department of Chemistry and Applied Biosciences, Eidgenössische Technische Hochschule (ETH) Zürich, Università della Svizzera italiana Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland; and [c]Departments of Molecular Cellular and Developmental Biology, Cell Biology and Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8103

Edited by Ken A. Dill, University of California, San Francisco, CA, and approved May 28, 2009 (received for review February 27, 2009)

The influence of the state of the bound nucleotide (ATP, ADP-Pi, or ADP) on the conformational free-energy landscape of actin is investigated. Nucleotide-dependent folding of the DNase-I binding (DB) loop in monomeric actin and the actin trimer is carried out using all-atom molecular dynamics (MD) calculations accelerated with a multiscale implementation of the metadynamics algorithm. Additionally, an investigation of the opening and closing of the actin nucleotide binding cleft is performed. Nucleotide-dependent free-energy profiles for all of these conformational changes are calculated within the framework of metadynamics. We find that in ADP-bound monomer, the folded and unfolded states of the DB loop have similar relative free-energy. This result helps explain the experimental difficulty in obtaining an ordered crystal structure for this region of monomeric actin. However, we find that in the ADP-bound actin trimer, the folded DB loop is stable and in a free-energy minimum. It is also demonstrated that the nucleotide binding cleft favors a closed conformation for the bound nucleotide in the ATP and ADP-Pi states, whereas the ADP state favors an open confirmation, both in the monomer and trimer. These results suggest a mechanism of allosteric interactions between the nucleotide binding cleft and the DB loop. This behavior is confirmed by an additional simulation that shows the folding free-energy as a function of the nucleotide cleft width, which demonstrates that the barrier for folding changes significantly depending on the value of the cleft width.

cytoskeleton | DNase binding loop | filament | protein folding

Actin is an abundant protein that is found in a variety of cell types and is important in a diverse set of cellular processes. Within eukaryotic cells, actin filaments confer shape and structure to the cytoskeleton, and the directed polymerization of the branched network of actin filaments is responsible for cell motility (1). The structure and properties of the branched network of actin filaments in the cytoskeleton are regulated by ATP hydrolysis within the well-defined nucleotide-binding pocket of individual actin monomers (2). In actin, ATP hydrolysis proceeds via a fast reaction step leading to the formation of ADP and bound phosphate (Pi). The ADP-Pi phase of actin is long-lived due to the slow release time of the Pi group (3). Among other things, ATP hydrolysis leads to softening of actin filaments (4, 5) and preferential binding of depolymerization factors (6). It follows that there is a substantial correlation between the properties of actin and the state of the bound nucleotide. Considerable research has therefore been devoted to quantifying the link between the state of the bound nucleotide and the structure of actin monomers (5, 7–10).

An interesting hypothesis has been proposed (7) concerning the relationship between actin structure and the state of the bound nucleotide. Actin monomers with bound ADP and labeled with tetramethylrhodamine (TMR) were found to adopt an α-helical conformation in residues 40–48, i.e., the so-called DNase-I binding loop (DB loop), whereas this region is disordered in ATP-bound actin (7). This finding led to the proposal that the DB loop folds on ATP hydrolysis. However, Rould et al. (8) found that the DB loop was disordered in another crystal form of both ATP- and ADP-actin and concluded that fortuitous contacts stabilized the α-helical DB loop in the crystals of rhodamine-ADP-actin. They concluded that the origin of nucleotide-dependent properties lies in a controlled opening of the binding cleft itself. This view is also supported by a previous study by Belmont et al. (11), and the proposed γ-phosphate timing mechanism (12, 13) is related to kinesins and G proteins in general.

A further consideration is that ATP hydrolysis is >40,000 times faster inside of the actin filament compared with isolated monomer (14). Therefore, the ADP and ADP-Pi forms of actin are believed to be much more relevant to the dynamics of the actin trimer and filament. If the DB loop does in fact undergo structural rearrangement on hydrolysis, it is far more likely to be important in the filament compared with the monomer. It is also conceivable that the hydrophobic pocket in adjacent actin monomers to which the DB loop binds could provide the "fortuitous contacts" that promote the proposed folding event (7). Unfortunately, there are no published crystal structures of the actin trimer or filament. Moreover, direct observation of the folding of α helices is challenging to observe both from the point of view of experiment and computation. Any computational investigation of these events should likely be fully atomistic in character, owing to the fact that solvent effects are likely to play a role both in the stability of the binding pocket and the small α-helical DB loop. Whereas experimental techniques suffer from the lack of a high-resolution crystal structure of actin filaments (15), sufficiently detailed all-atom computational techniques such as molecular dynamics (MD) are not able to sample long enough time scales to observe key folding events even within the actin monomer (10), because folding occurs on the microsecond timescale or longer. Additionally, various MD studies of actin or actin-related proteins have reported the nucleotide binding cleft in open and closed positions over simulation times ranging from 2–50 ns (5, 9, 16), but these studies have not explicitly simulated the opening and closing process.

Recently, however, a number of methods have been introduced (17–23) to overcome the issue of sampling rare events in molecular simulations. Specific to the study of protein folding, a number of approaches have been used including replica exchange (24), transition path sampling (25), and massively distributed computing (26). In particular, here we focus on metadynamics (18, 27), wherein accelerated sampling is accomplished by biasing the system through the addition to the potential energy function a sum of repulsive Gaussian terms based on certain descriptors also called collective variables (CVs) that approximately describe the real "reaction coordinate." Therefore, if a set of CVs are able to discriminate between two or more states of interest, metadynamics provides a computationally

Author contributions: J.P., T.D.P., M.P., and G.A.V. designed research; J.P. and D.B. performed research; J.P., D.B., M.P., T.D.P., and G.A.V. analyzed data; and J.P., D.B., M.P., T.D.P., and G.A.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: voth@hec.utah.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0902092106/DCSupplemental.

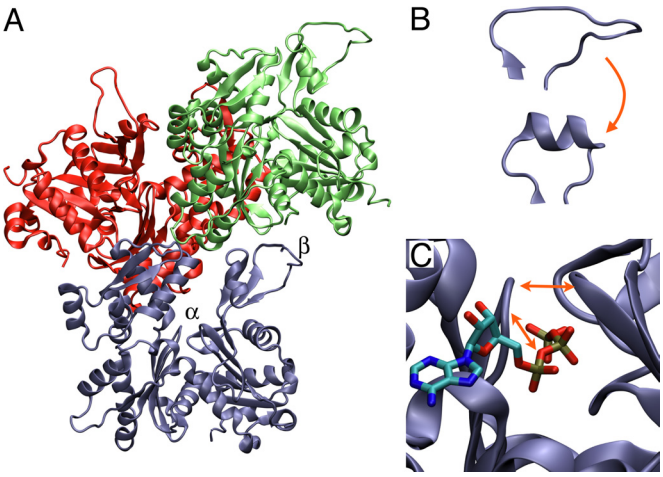

Fig. 1. Systems and biological events investigated using MD and metadynamics. (*A*) The actin trimer with the binding pocket and DB loop labeled as α and β, respectively. The monomer shown in blue is the monomer to which metadynamics was applied. (*B*) Shown is the proposed conformational change of the DB loop. (*C*) Shown is the nucleotide binding pocket. The arrows in *C* illustrate the collective variables used to study the binding pocket.

efficient route to exploring those states in a single simulation. Moreover, a defining feature of metadynamics lies in the fact that under certain conditions the history-dependent potential of Gaussian functions provides a good estimate of the free-energy of the system projected into the CVs (28, 29). Metadynamics has been successfully used so far to study several aspects of protein folding (30–32).

Herein an investigation of nucleotide state-dependent effects on the conformational states of the actin monomer and trimer is presented. Application of metadynamics to MD simulations of actin shows that the state of the nucleotide bound to the cleft influences both the width of the cleft and the conformation of the DB loop. The effect of the bound nucleotide on the relative free-energy differences between the stable conformational states has also been estimated. Metadynamics simulations have been performed on the ATP, ADP-Pi, and ADP states of actin in both the monomer and trimer. The metadynamics simulations also reveal a distinct allosteric relationship between the width of the bound nucleotide and the conformation of the DB loop.

Results

Opening and Closing of the Actin Nucleotide Binding Cleft. Metadynamics simulations were first performed to investigate the relationship between the width of the nucleotide binding cleft and the state of the bound nucleotide as depicted schematically in Fig. 1. The results for monomeric actin are shown in Fig. 2 as a function of the bound nucleotide. The most stable state, i.e., the region with the lowest free-energy, is the darkest color. The plots show a clear trend in the cleft width as a function of the bound nucleotide. In the ATP-bound state, the most stable state is a closed conformation, with strong contact between the nucleotide's P_β atom and the protein backbone. After ATP hydrolysis (ADP-Pi phase), the phosphate portion of the nucleotide loses its contact with the protein, and the most stable state becomes slightly more closed with respect to the distance defined in the methods section (i.e., the COM distance between residue nos. 14–16 and 156–158). The binding cleft of the ADP-Pi form becomes more closed than the ATP form because of the loss of contacts between phosphate groups and the protein backbone, as seen by the shift in the value of the contact number. In contrast, in the ATP form, the β- and γ-phosphates are positioned between each side of the binding cleft. Finally, after the

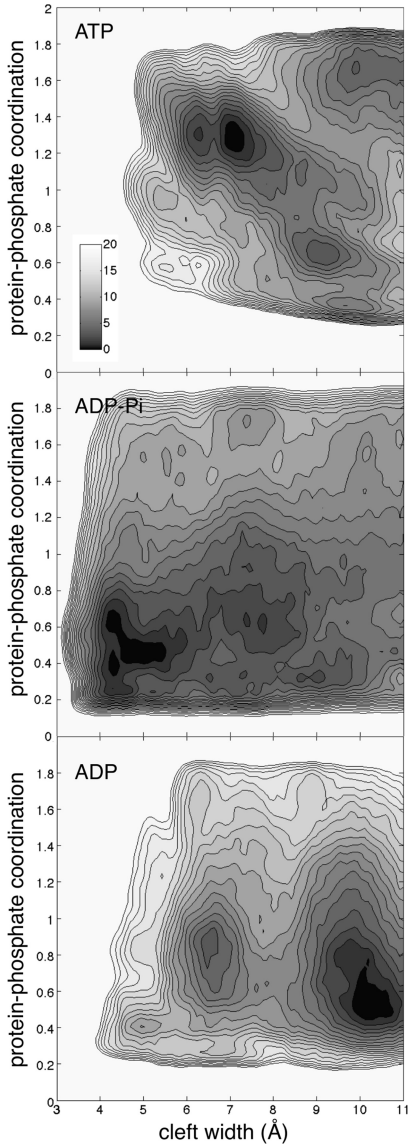

Fig. 2. Free-energy surfaces for the opening and closing of the nucleotide binding cleft in monomeric actin. The isolines are drawn using a 1 kcal/mol spacing, and the energy scale is in kcal/mol. Based on block-averaging analysis, the uncertainty is ≈1 kcal/mol. The collective variables used in the opening and closing simulations are described in *Results*.

release of the Pi group, the cleft is more stable in an open conformation. It can be seen that bound Pi group is responsible for the stabilization of the closed state, given that both the ADP-Pi and ADP phases are most stable when there is no contact between the β-phosphate group of the nucleotide and the binding cleft. The metadynamics simulations completely and reversibly explore the phase space defined by the full-range of the opening and closing CVs and, in the case of ATP and ADP monomeric actin, a few other metastable minima are located. Although the other minima found in these simulations are higher in free-energy by only a few kcal/mol, these regions are separated from any other metastable minima by relatively large barriers. It is particularly important that the ADP-Pi state is found in a closed conformation as was proposed in ref. 33, although not yet supported by high-resolution experimental structures or computer modeling.

It is also interesting to compare the most stable states found by metadynamics simulation with the experimental values avail-

able from high-resolution crystal structures. Table S1 contains a detailed comparison of the value of each CV in 2 ATP-bound monomeric actin structures (PDB entries 1ATN and 1NWK; refs. 48 and 49, respectively), ADP-bound actin (PDB entry 1J6Z; ref. 7), and those CV that would be obtained from a representative structure of actin-related protein (Arp) 3 (PDB entry 2P9S; ref. 34). Arp3 is a subunit of the Arp2/3 complex that has very similar structure to monomeric actin. It forms the first subunit of a branched daughter filament in the actin cytoskeleton (35), and there is a well-established link between its open and closed forms and the bound nucleotide (9, 16). As seen in the table, the ATP-bound form from metadynamics is more open than the crystal structures, but the difference is within the reported resolution. The most stable state from metadynamics of actin in the ADP form is more open than the corresponding experimental value. We note that for the clearest comparison possible, our metadynamics simulations of ADP monomeric actin were performed with the DB loop region unfolded. To determine whether there is any major correlation between the secondary structure of the DB loop and the opening and closing of the binding cleft, the same simulation was performed on the 1J6Z crystal structure that shows the DB loop in the folded state. The comparison between the 2 forms is given in Fig. S1. The results show an open form of the ADP state is the most stable form for both conformations of the DB loop, indicating that the most favored states are similar in both forms.

The same simulations were also performed for the actin trimer with the 3 different types of bound nucleotide. ATP hydrolysis in actin is much faster in the filament (14), so elucidation of nucleotide-dependent opening and closing effects must include an extension to these multiactin structures. The free-energy profiles for the trimer are given in Fig. S2. It is noted that within the trimer, metadynamics was only performed on 1 of the actin monomers (identified in Fig. 1). Overall, opening and closing of the nucleotide binding cleft is not greatly affected by the presence of 2 bound actin subunits. The similarities in the cleft width trends for the monomer and trimer indicate that the cleft width is a signature of the state of the bound nucleotide, but not a means by which the propensity for ATP hydrolysis can be determined. It is noted that the opening and closing simulations in the trimer were performed within the Holmes filament model as discussed in *SI Text*.

Folding of the DB Loop Region Through Metadynamics. Using the path-based CVs outlined in *Methods*, the nucleotide state-dependent folding of the DB loop region was next investigated. Following the procedure outlined in *SI Text*, an optimized reference path was first obtained. We then used the reference path to study folding of the DB loop and nucleotide state-dependent effects. For the monomeric actin system, the free-energy profiles obtained for all 3-nucleotide states are given in Fig. 3, along with representative structures for the folded, unfolded, and relevant transition states. Because no unusual features in the distance (z) CV were observed on the 2D free-energy surfaces, the profiles are a projection of both CVs onto the collective variable progress (s) only. It was found that all folding events observed in this study follow a characteristic nucleation-condensation mechanism for folding (36). The first barrier to overcome in the folding process (Fig. 3, point *TS1*) is a rotation of MET44 around the backbone leading to the formation of a general coiled shape, which can be seen as a nucleation step where a dense packaging is retrieved. After this step another barrier (Fig. 3, point *TS2*), corresponding to a concerted rotation of residues GLY46 and MET47, must be overcome to achieve the correct fold matching the published crystal structure of DB loop in the folded state, i.e., the condensation step. The condensation step is required to achieve the full network of hydrogen bonds following the $i - (i + 4)$

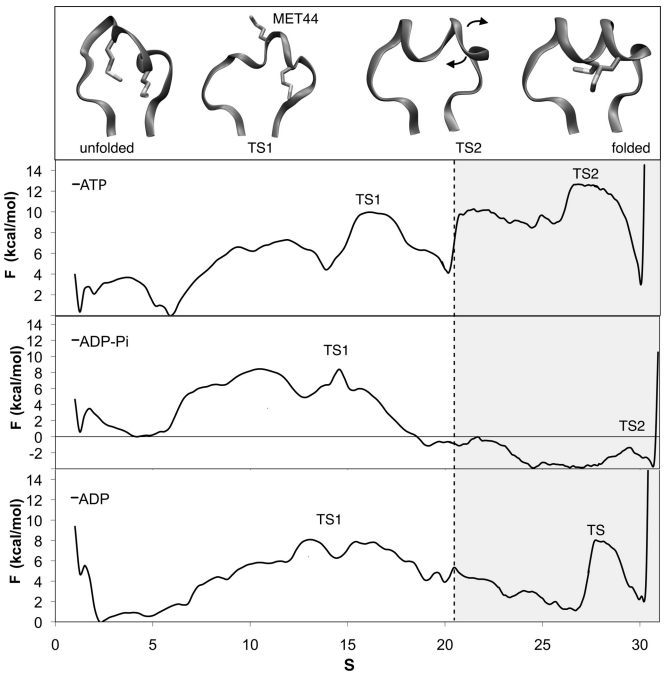

Fig. 3. Free-energy profiles for folding the DB loop region in monomeric actin as a function of the bound nucleotide. The profiles are a projection of both path collective variables onto the s collective variable. The transition states in the nucleation-condensation model discussed in *Results* are labeled *TS1* and *TS2* and shown (*Upper*) in ribbon representation along with representative unfolded and folded structures. (*Lower*) The shaded region marks the approximate location of the fine-grained/coarse-grained border in the path collective variables. The approximate error is 0.6 kcal/mol.

pattern for α-helices. Overall, the match between the folded DB loop obtained from metadynamics and that published in the 1J6Z crystal structure is very good. For example, by comparing only the DB loop region, the RMSD (C_α atoms) between simulation and experiment is $<$0.6 Å, whereas by comparing the entire protein (ADP-bound), the agreement of our structure with the crystal structure is 1.55 Å, which is approximately the reported experimental resolution. Although the final point on our reference path is the folded DB loop region itself, it is not guaranteed that the DB loop will fold during a metadynamics simulation. It is an indication of the accuracy of our proposed reaction coordinate (reference path) that the simulation reached the folded state during the simulation.

As seen in Fig. 3, there is a clear dependence of the folding free-energy profile on the state of the bound nucleotide. In ATP monomeric actin, folding proceeds through a high-energy intermediate, although the folded state ultimately resides in a stable free-energy minimum. The ADP-Pi state has a thermodynamic driving force to pass through the nucleation step. However, the actual folded state is a shallow local free-energy minimum. Finally, the folded basin in that ADP state is deep and well defined.

The folding of the DB loop region was also investigated within the actin trimer. The same procedure as in monomeric actin was followed to obtain the folding free-energy profiles for each of the 3 types of bound nucleotide. The profiles for folding in the trimer are shown in Fig. 4 for both the Holmes (37) and Oda (38) filament models (further details are provided in *SI Text*). As shown Fig. 4, the state of the bound nucleotide also clearly influences the DB loop folding free-energy profile as in monomeric actin. Although all 3 systems still possess an inherent nucleation-condensation type folding character, the 3 profiles show distinct behavior from one another. In particular, it is

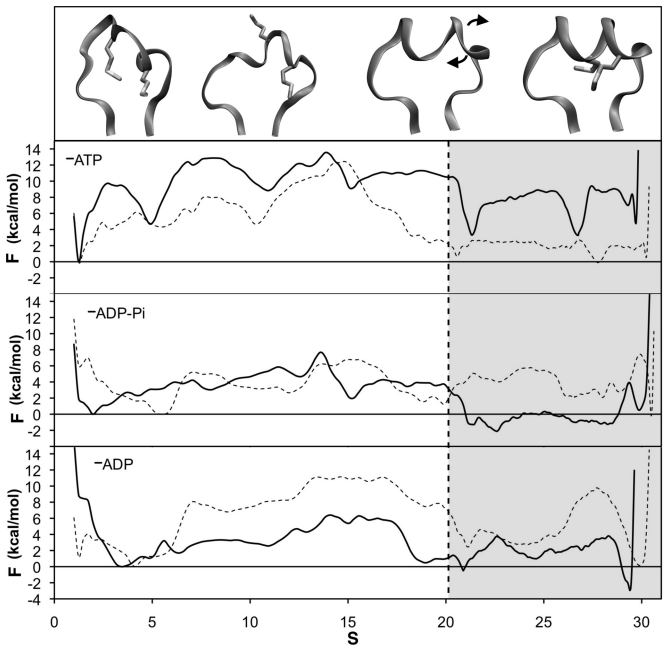

Fig. 4. Free-energy profiles for folding the DB loop region in the actin trimer as a function of the bound nucleotide for the Holmes (dashed line) and Oda (solid line) filament models. The profiles are a projection of both path collective variables onto the *s* collective variable. (*Upper*) Representative unfolded, transition state, and folded structures are given. (*Lower*) The shaded region marks the approximate location of the fine-grained/coarse-grained border. The approximate error is 0.6 kcal/mol.

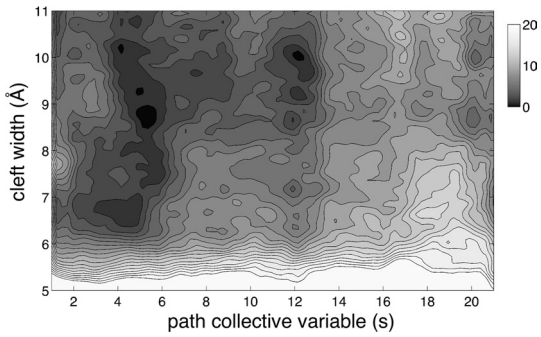

Fig. 5. Free-energy profile for combined metadynamics simulation by using both the cleft width collective variable and the path collective variable (*s*). The isolines are drawn using a 1-kcal/mol spacing, and the energy scale is in kcal/mol. The collective variables are described in *Results*.

interesting to see that in the ADP-Pi phase (also considered the most populated state of filamentous actin; ref. 39) the barrier for the folding nucleation event is reduced by nearly a factor of 2. Additionally, it can be seen that although the folding event is possible in the ATP trimer, the folded state is not stabilized to an appreciable degree. By contrast, the folded state within the ADP trimer is well-defined and stabilized by a deep well. The folding free-energy profiles for the 2 types of filament model have generally similar features with a few key differences. In the ATP form, the folded state in the Oda model is less stabilized than the unfolded state, although the overall barrier for folding is still high. In the ADP-bound Oda trimer, the barrier for the nucleation event is lower by ≈5 kcal/mol, and the folded state is more stabilized than the Holmes model. Oda et al. (38) reported that the folded DB loop in the actin filament could explain some of the low-resolution portions of the diffraction pattern in understanding the actin filaments structure, so it is particularly interesting to see the increased thermodynamic efficacy of folding of the DB loop within the Oda actin trimer.

The free-energy profiles presented in Figs. 3 and 4 demonstrate that the state of the bound nucleotide is able to modulate the folding free-energy profile. To better identify and quantify the allosteric interactions between the DB loop conformation and the bound nucleotide, an additional metadynamics simulation was performed. Using both the cleft-width collective variable and the path collective variable *s* (which measures the progress along the folding pathway) a 2D free-energy surface was obtained as shown in Fig. 5. This calculation proved to be very computationally expensive, therefore only the ADP actin monomer was used and only the coarse region (i.e., the nucleation region) of the folding pathway was investigated. As seen in Fig. 5, the barrier height for nucleation (i.e., the approximate region between $s = 10$ and $s = 20$) changes greatly depending on the width of the binding cleft, with the lowest energy crossing point for nucleation occurring at a cleft width in the open position.

To further understand the basis for the nucleotide dependent folding, principal component analysis (PCA) was performed on each of the 3 structures in the unfolded configuration. Details of the PCA calculations are provided in *SI Text*. The results from these calculations show a strong correlation between the features of the lowest eigenvector of the covariance matrix (a signature of the dominant motion being sampled) and the bound nucleotide. Fig. S3 shows the lowest eigenvector as a function of the bound nucleotide and the residue number of the protein. In the ATP state, the lowest eigenvector has a peak in the DB loop region (labeled in Fig. S3) and is relatively flat everywhere else. By contrast, the ADP-Pi and ADP states have a much smaller peak in the DB loop region, comparable to the other peaks in the graph. Moreover, the features of the binding cleft (labeled in Fig. S3) are identifiable in the ADP-Pi and ADP states. These results reveal correlation between the dynamics of the DB loop region and the state of the bound nucleotide. This correlation is also consistent with a previous coarse-grained study that found the dynamics of the DB loop region in ATP monomeric actin to be distinct from the rest of the protein (40).

To verify that the observed basins and transition states were not artifacts of the path procedure, we performed a committor analysis (41) on the second transition state, i.e., the condensation portion of the pathway. This analysis was done with a small model DB loop system (described in *SI Text*) designed to reproduce the essential features of the constrained folding process. As seen in Fig. S4, the point of 50% commitment, i.e., the definition of a dynamical transition state, falls approximately within k_BT of the saddle point, indicating that the actual dynamical transition state is very close to the transition state observed in metadynamics. Technical details regarding the committor calculation are provided in *SI Text*.

Additionally, the small model DB loop system was used to perform 5 metadynamics folding simulations with alanine substituting for key residues within the DB loop. The purpose of these simulations was to provide experimental targets for point substitutions, and to verify our findings regarding the key residues in the folding process. The results from the ALA substitutions are given in Table S2, which shows the ratio of the free-energy difference between the folded and unfolded states of each mutant compared with the wild-type. Both of the methionine residues in the DB loop (positions 44 and 47) are considered to be important because of their general propensity for α helical conformation (42) and to their specific role in the nucleation process in the folding of the DB loop. Moreover, the long side chains from these residues are able to interact with the $n + 2$ subunit in the actin filament. It is therefore interesting to

see that the M44A substitution has the largest effect on the folding process, increasing the free-energy of the folded state almost 4-fold compared with the wild-type. Glycine is the least favorable residue to be found in an α-helix (42); it is somewhat surprising, therefore, to see that the G42A and G46A substitutions increase the free-energy of the folded state compared with the wild-type. The 2 GLY residues in the DB loop region (positions 42 and 46) evidently confer some required flexibility to the DB loop region that is required for the constrained folding process.

Discussion

Based on the findings presented in *Results*, we can draw some key insights into the nucleotide-dependent structural properties of actin and multi-actin structures. Specifically, we are able to use the atomically detailed free-energy picture of the opening and closing and DB loop folding events within the monomer and trimer to address several key questions.

What Is the Native State of the DB Loop Region in Monomeric ADP-Bound Actin? Given the controversy over the structure of the DB loop region, it was a goal of this work to use metadynamics to try to determine the most probable structure. Because the simulations performed in this work were carried out in the full solvent environment and did not have (or require) labeling with TMR, the folding results can be considered to represent that of actin in conditions not affected by the labeling environment. It was shown (cf., Fig. 3) that the folding of the DB loop within TMR-free ADP monomeric actin is thermodynamically feasible. Moreover, the folded state is well defined and has a stable minimum on the free-energy surface. It is noteworthy that the folding free-energy profile for the ADP state in fact has several stable minima. This finding suggests an equilibrium exists between the 3 states and helps explain why the region is disordered in some monomeric actin crystal structures. It was proposed (8) that the helical form of the DB loop in ADP actin was the result of specific crystal packing interactions between actin monomers. However, given that we find a stable DB loop region with similar free-energy to the unfolded state, we propose that the DB loop region in monomeric ADP-bound actin will be found in both the folded and unfolded states under realistic conditions. Additionally, the present computational mutation studies (cf., Table S2) suggest a series of experiments that could be performed in the future to further understand the conformational properties of the DB loop.

Does ATP Hydrolysis Effect the Structure of the DB Loop in Actin? A central challenge since the publication of the first ADP-bound actin structure with a folded DB loop region (7) has been to determine the origin of this apparent nucleotide dependence of the DB loop conformational states. As seen in Figs. 3 and 4, a clear dependence of the folding free-energy profile on the bound nucleotide of actin has been demonstrated. Based on the PCA analysis performed in this work (cf., Fig. S5), we found a prevalent correlation between the state of the bound nucleotide and the dynamics of the DB loop region. Furthermore, an additional metadynamics calculation (Fig. 5) revealed that the barrier height of the early stages of folding (i.e., nucleation) depends on the width of the cleft of monomeric actin. This finding provides a plausible allosteric interaction between the DB loop and the bound nucleotide. Namely, the state of the nucleotide regulates the equilibrium cleft width (cf., Fig. 2), which in turn regulates the barrier height for nucleation in the folding pathway (cf., Fig. 5). These results support the idea that the bound nucleotide in actin can function as an allosteric regulator of the DB loop conformation.

What Is the Relationship Between the Conformation of the DB Loop and Actin–Actin Interactions? The discovery of the folded DB loop within ADP actin led to the idea that the folding of this region changes actin–actin interactions in the filament, thereby providing the basis for nucleotide-dependent properties in actin (7). The feasibility of this hypothesis was demonstrated by Chu and Voth who used MD calculations (5) and coarse-grained modeling (43) to show how the persistence length of filamentous actin can decrease after ATP hydrolysis and DB loop folding. Herein we have shown that nucleotide-dependent folding of the DB loop region is indeed energetically feasible in the actin trimer. It is noteworthy that the DB loop folding free-energy profile in ADP-Pi actin, the most populated state of actin because of fast hydrolysis and slow Pi release, has a significantly lower free-energy barrier for the nucleation step of the DB loop folding process. In contrast, the DB loop nucleation step in the ATP trimer has a large barrier, and, moreover, the folded state in the ATP trimer has no appreciable free-energy basin. Furthermore, we observe that in the ADP actin trimer, the folded DB loop is stabilized in a well-defined free-energy basin. This observation is consistent with previous work (5, 43, 44) suggesting that, in actin filaments, the helical form of the DB loop leads to a softer filament with a reduced persistence length. The results imply that the interactions in the filament between the DB loop in a neighboring actin subunit change the free-energy landscape for folding compared with the isolated monomer.

Methods

All MD calculations were performed by using NAMD 2.6 (45) in conjunction with the AMBER99SB (46) force field. Standard procedure was used and a detailed description of the MD system preparation and simulation protocol is provided in *SI Text*. The CVs used in the metadynamics studies in this work were chosen according to the specific problem. The 2 regions of actin where metadynamics was applied are illustrated in Fig. 1 *B* and *C*. The quality of results obtained via metadynamics is related to the ability of CVs to discriminate between the states of interest. Previous studies of actin used the distance between the C_α atoms of residues that span the nucleotide binding cleft to describe open and closed conformations [the residue pairs 15/157 and 14/158 were used in actin (9), and a similar convention was used for Arp2 (16) and Arp3 (9, 16)]. To study the opening and closing of the nucleotide binding cleft, 2 CVs were used. Inspired by the previous studies the open and closed conformations of actin, we used the center-of-mass (COM) distance of the 2 lobes (residues 14–16 and 156–158) of the protein constituting the binding cleft. Additionally, we used metadynamics to study the propensity of the phosphate groups of the bound nucleotide to remain bound to the protein. Therefore, the number of contacts between the beta phosphate atom and the protein was also used as a CV. These CVs are calculated from the geometry of the system in a straightforward manner. Further details regarding their application are provided in *SI Text*.

Protein folding is a challenging problem to study with biasing methods because of the difficulty in developing CVs that both describe the unfolded and folded states and provide sufficient bias to travel between the folded and unfolded states. To simulate the folding of the DB loop region, we have used the recently introduced path-based CVs (47):

$$s(R) = \frac{\sum_{i=1}^{P} i e^{-\lambda(R - R_i)^2}}{\sum_{i=1}^{P} e^{-\lambda(R - R_i)^2}}$$

$$z(R) = -\frac{1}{\lambda} \ln \left(\sum_{i=1}^{P} e^{-\lambda(R - R_i)^2} \right)$$

These CVs function by comparing the configuration of the DB loop (R), at any moment during the simulation, to a reference pathway that describes the folding process (R_i). The CVs are unique in their ability to sample s and z along some reference path composed of P frames (the quantity λ is a metric of the

interframe spacing). The index i in Eqs. **1** and **2** refers to the ith frame along the reference pathway. By convention, $i = 1$ refers to the unfolded state, and $i = P$ (30 in the present study) refers to the folded state. At any given moment during the simulation the coordinates R can be used to determine a unique value for s, the progress along the folding pathway, which ranges from 1 to P. Additionally, the distance away from the reference pathway ($R = R_i$ or by definition $z = 0$) is determined by Eq. **2**. Thus, the variables s and z together define a tube of configuration space that is centered on the reference pathway. The object of metadynamics with the path-based CVs, therefore, is to sample this tube of configuration space and provide a free-energy profile for the folding pathway.

Additionally, we discovered in the course of establishing the reference path that the folding of the DB loop proceeds through a series of large (with respect to the interframe RMSD chosen as the measure for distance between frames) movements of the backbone and side chains followed by a series of very short-range motions of 2 key residues. In other words, the first part of the folding process is characterized by comparatively large changes in the RMSD between 2 adjacent frames $R_i/R_i + 1$, whereas the second part of the folding process, characterized by fine motions, requires more frames to describe the small changes in the RMSD between adjacent frames. In practical terms, the required number of frames (i.e., resolution) to accurately sample the folding pathways was observed to change as a function of the progress along the pathway. Given that the computational cost for the calculation of the path-based CVs would be exceedingly high by using the highest resolution needed, we have developed a multiscale implementation of the path-based metadynamics approach according to the scales of the different features of this folding event. In this spirit, the system is divided into a coarse-grained and fine-grained region, and the simulation is performed with a large enough overlap between the 2 regions that the free-energy surfaces can be matched through rescaling of the s parameters by the ratio of the compression factors between the 2 paths. In this way, a free-energy surface for the entire folding event can be produced. This method proved to be very robust and worked well for all systems studied for the folding of the DB loop. A detailed description of how the reference path for folding was obtained and further details of the metadynamics simulations (including a description of the convergence criteria and error estimation) are given in *SI Text*.

ACKNOWLEDGMENTS. This work was supported in part by the National Science Foundation International Research Fellows Program (OISE-0700080 to J.P.) and by the National Institutes of Health Grant GM066311 (to T.D.P.). Computational support was provided by the National Science Foundation Teragrid resources at the Texas Advanced Supercomputing Center and the Swiss National Supercomputing Center.

1. Pollard TD, Blanchoin L, Mullins RD (2000) Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annu Rev Biophys Biomol Struct* 29:545–576.
2. Reisler E, Egelman EH (2007) Actin's structure and function: What we still do not understand. *J Biol Chem* 282:36133–36137.
3. Carlier MF, Pantaloni D (1986) Direct evidence for ADP-Pi-F-actin as the major intermediate in ATP-actin polymerization. Rate of dissociation of Pi from actin-filaments. *Biochemistry* 25:7789–7792.
4. Isambert H, et al. (1995) Flexibility of actin filaments derived from thermal fluctuations. Effect of bound nucleotide, phalloidin, and muscle regulatory proteins. *J Biol Chem* 270:11437–11444.
5. Chu JW, Voth GA (2005) Allostery of actin filaments: Molecular dynamics simulations and coarse-grained analysis. *Proc Natl Acad Sci USA* 102:13111–13116.
6. Pollard TD, Borisy GG (2003) Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112:453–465.
7. Otterbein LR, Graceffa P, Dominguez R (2001) The crystal structure of uncomplexed actin in the ADP state. *Science* 293:708–711.
8. Rould MA, Wan Q, Joel PB, Lowey S, Trybus KM (2006) Crystal structures of expressed non-polymerizable monomeric actin in the ADP and ATP states. *J Biol Chem* 281:31909–31919.
9. Dalhaimer P, Pollard TD, Nolen BJ (2007) Nucleotide-mediated conformational changes of monomeric actin and Arp3 studied by molecular dynamics simulations. *J Mol Biol* 376:166–183.
10. Zheng X, Diraviyam K, Sept D (2007) Nucleotide effects on the structure and dynamics of actin. *Biophys J* 93:1277–1283.
11. Belmont LD, Orlova A, Drubin DG, Egelman EH (1999) A change in actin conformation associated with filament instability after Pi release. *Proc Natl Acad Sci USA* 96:29–34.
12. Sablin EP, Kull FJ, Cooke R, Vale RD, Fletterick RJ (1996) Crystal structure of the motor domain of the kinesin-related motor ncd. *Nature* 380:555–559.
13. Vale RD (1996) Switches, latches, and amplifiers: Common themes of G proteins and molecular motors. *J Cell Biol* 135:291–302.
14. Blanchoin L, Pollard TD (2002) Hydrolysis of ATP by polymerized actin depends on the bound divalent cation but not profilin. *Biochemistry* 41:597–602.
15. Aguda AH, Burtnick LD, Robinson RC (2005) The state of the filament. *EMBO Rep* 6:220–226.
16. Pfaendtner J, Voth GA (2008) Molecular dynamics simulation and coarse-grained analysis of the Arp2/3 complex. *Biophys J* 95:5324–5333.
17. Patey GN, Valleau JP (1975) A Monte Carlo method for obtaining the interionic potential of mean force in ionic solution. *J Chem Phys* 63:2334–2339.
18. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99:12562–12566.
19. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. *J Comput Chem* 16:1339–1350.
20. Ferrenberg AM, Swendsen RH (1988) New Monte-Carlo technique for studying phase-transitions. *Phys Rev Lett* 61:2635–2638.
21. Roux B (1995) The calculation of the potential of mean force using computer simulations. *Comput Phys Commun* 91:275–282.
22. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693.
23. Darve E, Pohorille A (2001) Calculating free energies using average force. *J Chem Phys* 115:9169–9183.
24. Zhou R (2006) Replica exchange molecular dynamics method for protein folding simulation. *Protein Folding Protocols*, *Methods in Molecular Biology*, eds Bai Y, Nussinov R (Humana Press, Totowa, NJ), Vol 350, pp 205–223.
25. Dellago C, Bolhuis PG (2007) Transition path sampling simulations of biological systems. *Atomistic Approaches in Modern Biology: From Quantum Chemistry to Mol Simuls*, *Topics in Current Chemistry*, ed Reiher M (Springer, Berlin), Vol 268, pp 291–317.
26. Snow CD, Nguyen N, Pande VS, Gruebele M (2002) Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* 420:102–106.
27. Barducci A, Bussi G, Parrinello M (2008) Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys Rev Lett* 100:020603.
28. Bussi G, Laio A, Parrinello M (2006) Equilibrium free energies from nonequilibrium metadynamics. *Phys Rev Lett* 96:090601.
29. Laio A, Rodriguez-Fortea A, Gervasio FL, Ceccarelli M, Parrinello M (2005) Assessing the accuracy of metadynamics. *J Phys Chem B* 109:6714–6721.
30. Piana S, Laio A (2007) A bias-exchange approach to protein folding. *J Phys Chem B* 111:4553–4559.
31. Bonomi M, Branduardi D, Gervasio F, Parrinello M (2008) The unfolded ensemble and folding mechanism of the C-terminal GB1 β-hairpin. *J Am Chem Soc* 130:13938–13944.
32. Bussi G, Gervasio FL, Laio A, Parrinello M (2006) Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc* 128:13435–13441.
33. Sablin EP, et al. (2002) How does ATP hydrolysis control actin's associations? *Proc Natl Acad Sci USA* 99:10945–10947.
34. Nolen BJ, Pollard TD (2007) Insights into the influence of nucleotides on actin family proteins from seven structures of Arp2/3 complex. *Mol Cell* 26:449–457.
35. Rouiller I, et al. (2008) The structural basis of actin filament branching by the Arp2/3 complex. *J Cell Biol* 180:887–895.
36. Nölting B (2006) *Protein Folding Kinetics* (Springer, Berlin).
37. Holmes KC, Popp D, Gebhard W, Kabsch W (1990) Atomic model of the actin filament. *Nature* 347:44–49.
38. Oda T, Iwasa M, Aihara T, Maeda Y, Narita A (2009) The nature of the globular- to fibrous-actin transition. *Nature* 457:441–445.
39. Robertson CI, Gaffney DP, Chrin LR, Berger CL (2005) Structural rearrangements in the active site of smooth-muscle myosin. *Biophys J* 89:1882–1892.
40. Zhang Z, et al. (2008) A systematic methodology for defining coarse-grained sites in large biomolecules. *Biophys J* 95:5073–5083.
41. Bolhuis P, Chandler D, Dellago C, Geissler P (2002) Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu Rev Phys Chem* 53:291–318.
42. Nelson DL, Cox MM (2000) *Lehninger Principles of Biochemistry* (Worth, New York), 3rd Ed.
43. Chu JW, Voth GA (2006) Coarse-grained modeling of the actin filament derived from atomistic-scale simulations. *Biophys J* 90:1572–1582.
44. Dominguez R (2004) Actin-binding proteins—a unifying hypothesis. *Trends Biochem Sci* 29:572–578.
45. Phillips JC, Braun R, Wang W, Gumbart J (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.
46. Hornak V, et al. (2006) Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins* 65:712–725.
47. Branduardi D, Gervasio FL, Parrinello M (2007) From A to B in free energy space. *J Chem Phys* 126:054103.
48. Graceffa P, Dominquez R (2003) Crystal structure of monomeric actin in the ATP state. *J Biol Chem* 278:34172–34180.
49. Kabsch W. Mannherz HG, Suck D, Pai EF, Holmes KC (1990) Atomic structure of the actin: DNase complex. *Nature* 347:37–44.