THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

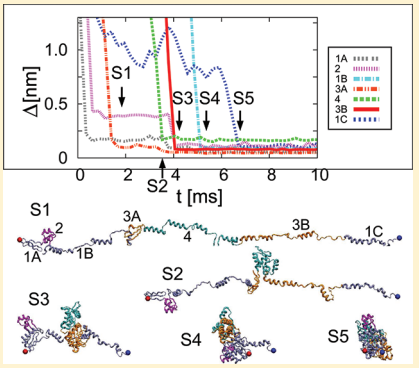
Exploring the Role of Topological Frustration in Actin Refolding with Molecular Simulations

Ji Young Lee, Li Duan, Tyler M. Iverson, and Ruxandra I. Dima*

Department of Chemistry, University of Cincinnati, Cincinnati, Ohio 45221, United States

Supporting Information

ABSTRACT: Actin plays crucial roles in the life of the cell while being notorious for its inability to reach a functional conformation without the help of assistant proteins. In eukaryotes, for example, the cytosolic chaperonin containing TCP-1 (CCT) and prefoldin (PFD) are required for actin folding assistance and prevention of protein aggregation in the crowded cellular environment. The folding of non-native actin is known to occur in a number of steps, but the reasons underlying its folding difficulty are unknown. Because a full, atomistic-level, investigation of the kinetics and thermodynamics of folding of such a large molecule is beyond computational reach, we focused our investigation on the role of topological frustration on the folding of actin. Namely, we studied the (re)folding of actin using simulations of a variant self-organized polymer model (SOP-DH) starting from a stretched state, leading to results that correlate well with experimentally driven conclusions and allowing us to make a number of testable predictions. Primarily, our simulations reveal that the successful refolding of the C-terminus end of actin occurs through a zipping process in which the α-helices wind up turn by turn upon formation of their native tertiary contacts. In turn, an early formation of the helical structure in this region of the chain has deleterious effects for actin's refolding fitness. Moreover, the C-terminus refolding is a very rare event in our simulations, in agreement with the large activation barrier predicted on the basis of experimental studies of actin unfolding in EDTA. We also discovered that subdomain 4 has a low refolding probability, which can help explain why many of the non-native actin target binding sites for CCT and PFD are located within this subdomain.



INTRODUCTION

Actin is one of the most abundant and highly conserved proteins in all eukaryotes. This cytoskeletal protein plays important roles in cellular processes such as muscle contraction, cell motility, cell division and cytokinesis, vesicle and organelle movement, and the establishment and maintenance of cell junctions and cell shape.[1−3] Monomeric actin, which spans 375 amino acids, adopts a globular shape (G-actin) consisting of two domains separated by a deep cleft which is the binding site for a nucleotide, either ADP or ATP.[4] The two domains, called the small (S) and large (L) domains, are further subdivided into subdomains with functional relevance. Namely, the S domain is subdivided into subdomains 1 and 2 and the L domain is subdivided into subdomains 3 and 4 (Figure 1). Out of these, only subdomains 2 (33−69) and 4 (181−269) are contiguous along the sequence. In contrast, subdomain 1 is divided into three segments (1A, 1−32; 1B, 70−144; 1C, 338−375) and contains both ends of the chain, while subdomain 3 consists of two regions of the chain (3A, 145−180; 3B, 270−337). Moreover, the DNase I-binding loop (DB loop, residues 40−48) in subdomain 2 adopts a disordered loop structure in actin with ATP (G-ATP), while it forms a short α-helix in G-actin with ADP (G-ADP).[5]

Despite its crucial functional relevance in the life of the cell, G-actin is one of the most well-known examples of a poor folder, i.e., a protein unable to fold autonomously. This makes

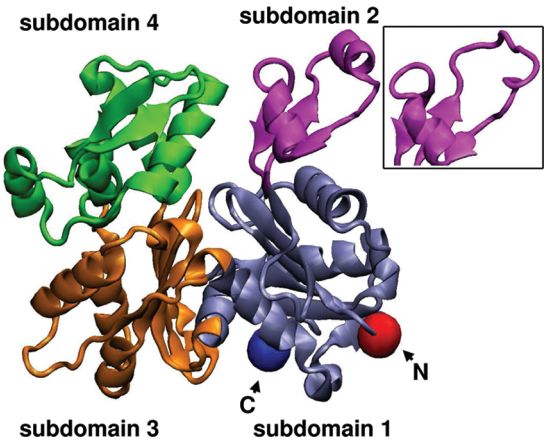


Figure 1. Crystal structure of G-ADP from the PDB (1J6Z).[5] The subdomains 1, 2, 3, and 4 are shown in iceblue, magenta, orange, and cyan, respectively. The inset on the right side shows the DB loop structure in G-ATP (1NWK).[25] This and all the conformational snapshots have been produced using the VMD[38] and PovRay packages.

Received: September 27, 2011
Revised: January 5, 2012
Published: January 16, 2012

actin a major target for the eukaryotic CCT and PFD. Experiments revealed that the nascent actin chain is captured during translation by PFD, which, upon completion of synthesis, transfers the non-native actin protein to the CCT. To add to an already complicated folding scenario, additional helper proteins are also required for the *in vivo* folding of actin such as CAP and the phosducin-like protein PLP2p in yeast.[6]

As a representative assistant protein, the cytosolic chaperonin CCT assists actin in folding in its native conformation.[7] CCT is a double barreled oligomer, and it is composed of eight different subunits.[8] In their cryo-electron microscopy study, Llorca et al.[9] reported that the small domain of actin (primarily, subdomain 2) binds to CCTδ and subdomain 4 from the large domain to CCTβ or CCTε. The biochemical study of Hynes et al.[10] confirmed the previous suggestion that CCTδ interacts with subdomain 2 from the small domain of actin and that either CCTβ or CCTε can interact with subdomain 4 from actin's large domain. More recently, Rommelaere and collaborators showed, using full Ala scanning mutagenesis coupled with pulse-chase time-course experiments of selected mutants, that the discrete binding determinants for CCT recognition by actin are the I-site (positions 30−34 in subdomain 1A), the N-site (positions 135−139 and 170−174 in subdomains 1B and 3A, respectively), the M-site (positions 240−254 and 265−274 from subdomain 4), and the C-site (positions 340−349 from subdomain 1C).[11] Out of these positions, Rommelaere and collaborators proposed that the region 245−249 from subdomain 4 is the initial recognition region by CCT. Hansen et al.[12] reported that PFD remains bound to the relatively unfolded actin polypeptide until its post-translational delivery to cytosolic chaperonin (CCT). Rommelaere et al.[11] also showed that the PFD binding sites in actin are site I consisting of positions 60−89 (from subdomains 1B and 2) and site II composed of positions 170−183 and 194−199 (from subdomains 3A and 4).

This large body of experimental evidence clearly indicates that the interaction between actin and helper proteins is crucial for the proper acquisition of its folded functional state. Moreover, it has been shown that, unlike its binding to PFD, binding of actin to CCT is an obligatory step in the folding of actin. However, only a rudimentary picture of the folding process is known.[6,11] The central finding is that the folding of actin involves at least one intermediate state that does not include the 1C subdomain part of the chain, which folds only in the end. The folding of the 1C subdomain is presumed to take place inside the CCT cavity as a result of the ATP hydrolysis in the CCT subunits.[6] Moreover, it has been proposed, based on cryo-EM and biochemical studies, that the non-native state delivered by PFD to the CCT consists of partial or even fully folded S and L domains, but in an open configuration (with the hinge regions connecting them being extended) and not containing the 1C subdomain.[11] However, it is unclear whether the formation of the S and L domain is initiated on the PFD given that, at least *in vitro*, PFD is not an obligatory helper protein. Alternatively, the folding of the two domains can be initiated cotranslationally as is the case with a variety of other proteins with a mixed α-helical and β-sheet fold as in actin.[13,14] For example, Kosolapov and Deutsch recently showed that, in addition to the established folding of α-helical protein stretches inside the ribosomal tunnel, long-range interactions that stabilize a β-hairpin spanning 22 amino acids can form near the exit port of the tunnel.[15] It is also unclear whether the S and L domains are properly folded when binding to the CCT. Experiments also suggest that the folding of actin involves more than one transition step and the time lag measured between the release of actin from ATP bound CCT versus release from ATP bound to the loaded CCT complex preincubated with AMP-PNP of ~250 s[6] is certainly long enough to accommodate a number of conformational transitions.

To shed light on the regions of actin that require help from assistant proteins, as well as on the sequence of events that can lead to the proper folding of actin starting from a possible cotranslational folding step to the eventual complete folding (in the CCT cavity), we carried out computer simulations of the folding of G-ATP and G-ADP based on a coarse-grained description of the chain, the self-organized polymer (SOP) model.[16] Previously, SOP led us to unravel the details of the force-unfolding scenarios in actin, the green fluorescent protein (GFP), and tubulin in agreement with experimental measurements.[16−19] Furthermore, a variant of this model, the SOP-SC model, has been successfully employed to follow the collapse kinetics and to determine the chevron plots of denaturant-induced unfolding and refolding of the src SH_3 domain.[20] Due to the large size of actin and the long folding time (>100 s), a full atomistic, explicit solvent, exploration of the folding dynamics of actin is beyond the reach of computational investigations. However, as shown by our previous study of the refolding of GFP from a rod-like configuration,[16] minimalist models provide important insight into the main kinetic traps encountered during the folding of slow or poor folders leading to (semi)quantitative predictions regarding the refolding reaction. Moreover, the recent investigation into the refolding of the chemically denatured src SH_3 domain[20] revealed a high quantitative agreement between simulations and experiments with regards to the thermodynamics of folding and that predicted concentration-dependent folding rates were in agreement with experimental measurements.

The SOP model based simulations allowed us to investigate the role of topological frustration in the refolding of G-actin initiated from the fully stretched rod-like structure. This model, similar to other topology-based or Go-like models, does not include energetic frustration resulting from formation of stable non-native interactions. Therefore, we cannot detect the presence of a potential intermediate stabilized by such non-native interactions. The choice of the starting state was inspired by our wish to look at the conformational changes initiated in the G-actin chain upon translation. As mentioned above, folding of actin requires the action of CCT. In our study, we mimicked the effect of CCT indirectly. Namely, because the C-terminal end of the chain is the last one to emerge from the ribosome and experiments revealed that it remains attached to CCT for the duration of the folding in the CCT cavity,[6] we too kept this end of the chain fixed in the majority of our simulation runs. The rationale behind this treatment of CCT is that the main goal of our study is to investigate the susceptible regions in G-actin that require the support of chaperones for correct folding to a functional state.

Our study provides molecular-level structural details related to the folding of actin. Despite the above-mentioned limitations, we found that most of our results correlate well with the available experimental data. We found that the refolding of the chain always starts at its N-terminus and that the C-terminus region of the chain (subdomain 1C) refolds last. Remarkably, attempts to speed up the refolding in the C-term part of the chain led to misfolding not only in this region but in other subdomains as well. Our simulations also reveal

that, in the absence of PFD, G-actin follows one of two refolding pathways. One pathway results in the early partial formation of the two domains (S and L) similar to what has been proposed on the basis of a number of experimental studies.[6,11,21] The other pathway, which is less probable than the first, leads to the early formation of only a partial S domain, with the L domain folding right before the last step which, as mentioned before, corresponds to the refolding of the 1C subdomain from the C-terminus end. As evidence for this second pathway is scarce, we propose that one possible role for PFD during the early folding steps in G-actin is to drastically reduce the likelihood of this alternative folding scenario.

METHODS

Self-Organized Polymer (SOP) Model. We used a topology-based model for actin in which each amino acid is represented by its C_α atom.[16,18,19] The total energy function for a conformation, specified in terms of the coordinates $\{r_i\}$ (i = 1, 2, ..., N), where N is the total number of residues, is

$$\begin{aligned} V_T &= V_{FENE} + V_{NB}^{ATT} + V_{NB}^{REP} \\ &= -\sum_{i=1}^{N-1} \frac{k}{2} {R_0}^2 \log\left(1 - \frac{(r_{i,i+1} - r_{i,i+1}^0)^2}{{R_0}^2}\right) \\ &\quad + \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \varepsilon_h \left[\left(\frac{r_{ij}^0}{r_{ij}}\right)^{12} - 2\left(\frac{r_{ij}^0}{r_{ij}}\right)^{6}\right] \Delta_{ij} \\ &\quad + \sum_{i=1}^{N-2} \varepsilon_l \left(\frac{\sigma_{i,i+2}}{r_{i,i+2}}\right)^6 + \sum_{i=1}^{N-3} \sum_{j=i+3}^{N} \varepsilon_l \left(\frac{\sigma}{r_{ij}}\right)^6 \\ &\quad \times (1 - \Delta_{ij}) \end{aligned} \tag{1}$$

The distance between two neighboring interaction sites i and i + 1 is $r_{i,i+1}$, and its initial value in the native structure is $r_{i,i+1}^0$. The finite extensible nonlinear elastic (FENE) potential (first term in eq 1) describes the backbone chain connectivity. We use the Lennard-Jones potential (second term in eq 1) to account for the interactions that stabilize the native state. If the noncovalently linked residues i and j for $|i - j| > 2$ are within a cutoff distance R_c in the native state (i.e., $r_{ij}^0 < R_c$), then Δ_{ij} = 1. If $r_{ij}^0 > R_c$, then Δ_{ij} = 0. A uniform value for ε_h = 2 kcal/mol, which specifies the strength of the nonbonded interactions, is assumed. All non-native interactions (third and fourth terms in eq 1) are repulsive. There are six parameters in the energy function (see Table S1 in the Supporting Information) which were taken from our previous study on the mechanical unfolding of actin.[17]

Simulations. We used Brownian dynamics simulations at T = 300 K[22] to generate the mechanical folding trajectories. To estimate the simulation time scale, we used $h = 0.08\tau_H$, where $\tau_H = (\zeta\varepsilon_h/k_BT)\tau_L$, with τ_L = 2 ps and ζ = 50 (the unitless friction coefficient which accounts for the high friction regime) for the overdamped limit. These choices led to an integration time step of $h \sim 27$ ps. We have followed the AFM experimental procedure to monitor refolding dynamics by quenching force from a high to a low value.[23] The dynamics of the folding transition is monitored from an extended low-entropy structure (rod-like initial state) to a folded low-entropy final state (native structure), by setting a force of 7 pN on the N-term residue, while the C-term residue is kept fixed. The 7 pN force represents 2−4% of the first rupture force for the actin monomer (∼200−300 pN) measured in our previous unfolding simulations.[17] Moreover, recent experiments[6] that analyzed the interaction between G-actin and CCT showed the C-term end of G-actin (position 374) remains attached to CCT until its final release upon ATP hydrolysis and opening of the CCT lid. Thus, running the simulation with the C-term end of the chain fixed accounts for a crucial aspect of actin's interaction with chaperones.

To investigate the role of topological frustration in the refolding of G-actin, we performed a search through the dihedral angle potential space employing the original SOP model and the SOP-DH model (described below). For each choice of parameters set, we generated 8−16 mechanical refolding trajectories for both G-ADP and G-ATP by using the coordinates of the corresponding crystal structures (Protein Data Bank (PDB)[24] codes: 1J6Z[5] and 1NWK[25]) as target (native) states in our simulations. Since there are missing residues (40−51) in the DB loop of 1NWK, we reconstructed this region using the corresponding part of the G-ATP structure from 1ATN.[4] Figure 1 depicts 1J6Z, where G-ADP (G-ATP) has an α-helix (a loop) structure for the DB loop in subdomain 2. The various types of simulations performed are summarized in Table S2 of the Supporting Information.

Angle Potentials: The SOP-DH Model. *Bond Angle Potential.* For the bond angle between three successive residues i, i + 1, and i + 2, we used the established harmonic potential $V_{bond\ angle} = \sum_{i=1}^{N-2} (k_\theta/2)(\theta_i - \theta_0)^2$, where $k_\theta = 20\varepsilon_h/(rad)^2$. θ is the bond angle in the current conformation, and θ_0 is the bond angle obtained from the native PDB structure.

Dihedral Angle Potential. The dihedral angle potential, measuring the energetic cost of the rotation between planes formed by four successive residues, depends on the secondary structure element encompassing the four chain positions.[26,27] Namely, for α-helices, we used

$$\begin{aligned} V_{helix} &= \sum_{i=1}^{N-3} [A_i(1 - \cos\phi) + B_i(1 + \cos 3\phi) \\ &\quad + C_i(1 - \sin\phi)] \end{aligned} \tag{2}$$

where $A_i = 1\varepsilon_h$, $B_i = 1.6\varepsilon_h$, and $C_i = 2\varepsilon_h$. ϕ is the dihedral angle. For β-sheets and loops, we used

$$\begin{aligned} V_{sheet,loop} &= \sum_{i=1}^{N-3} [A_i(1 + \cos\phi) \\ &\quad + B_i(1 + \cos 3\phi) + C_i \sin\phi] \end{aligned} \tag{3}$$

where $A_i = 1.2\varepsilon_h$, $B_i = 1.2\varepsilon_h$, and $C_i = \pm 0.6\varepsilon_h$ for a β-sheet and $A_i = 0\varepsilon_h$, $B_i = 0.2\varepsilon_h$, and $C_i = \pm 0.3\varepsilon_h$ for a loop. Moreover, following the above studies, for C_i, we choose a positive (negative) value if ϕ of the PDB structure is less (greater) than 0.

Data Analysis. To monitor the progress of the folding reaction, we followed the time series of a number of chain parameters such as the N-term to C-term distance (R), the fraction of native contacts (Q), the radius of gyration (R_G), the root-mean-square deviation (Δ), and the overlap function (χ). To eliminate the thermal noise from the analysis of these parameters, we report averages for each parameter over a time period of ∼2 ms. The overlap function is defined as $\chi(t) = [2/(N^2 - N)] \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \Theta(R_0 - |r_{i,j} - r_{i,j}^0|)$. Here, $r_{i,j}$, $r_{i,j}^o$, and R_0

dx.doi.org/10.1021/jp209340y | J. Phys. Chem. B 2012, 116, 1677−1686

Table 1. Refolding Failure for the Full Actin Molecule and Each of Its Subdomains during the Force-Quench Refolding of G-ADP and G-ATP[a]

		subdomain							
		whole	1A	2	1B	3A	4	3B	1C
G-ADP	D0	7	5	4	4	2	5	4	5
	DH1	7	0	0	0	0	0	4	6
	DH0	4	1	2	2	0	0	0	2
	DH10	5	0	1	0	0	3	1	1
	DH01	7	0	1	2	0	1	4	2
	DH1′	12	3	7	3	0	2	6	9
	DH0A	4	1	1	1	0	0	3	0
	DH0B	3	0	0	1	0	2	0	
	DH0C	5	1	4	4	0	1	0	
	sum	54	11	20	17	2	14	22	25
G-ATP	D0	8	4	3	6	7	7	7	6
	DH1	7	0	3	1	0	2	3	7
	DH0	8	0	1	1	0	6	4	2
	DH10	7	0	2	0	1	4	3	2
	DH01	8	0	2	0	0	3	0	6
	DH1′	14	1	4	1	0	4	1	13
	DH0A	7	0	3	1	0	5	2	2
	DH0B	7	2	4	2	0	4	3	
	DH0C	4	0	2	0	1	3	0	
	sum	70	7	24	12	9	38	23	38

[a]D0, DH1, DH0, DH10, DH01, DH1′, DH0A, DH0B, and DH0C are defined in the text and in Table S2 of the Supporting Information. Eight trajectories are obtained for each type of simulation, except for DH1′ where we ran 15 trajectories for G-ADP and 16 for G-ATP and DH0C for G-ATP for which we ran 6 trajectories. The integer in each column denotes the total number of trajectories resulting in misfolding of actin or its subdomains. The sum is the sum of the numbers for each column.

are the same as in eq 1 and $\Theta(x)$ is the Heaviside step function. N is the number of residues in a particular subdomain, since we calculated the overlap function for each subdomain. $\chi(t) = 1$ means that the structure at time t is identical to the PDB structure, while a 0 or a small value means that the structure is highly deformed compared to the one in the PDB.[24]

RESULTS

The exact configuration of G-actin in the absence of the nucleotide is unknown and ATP binding to actin is not a prerequisite for its release from CCT.[11] Thus, in all our simulations, we followed the dynamics of folding of G-actin into each of the nucleotide configurations, G-ATP and G-ADP, present in the PDB.[24]

To follow the dynamics of the chain, we determined the best reaction coordinate(s) for the refolding reaction.[28] We expect that a good refolding reaction coordinate shows changes at the moments corresponding to the main transitions in the chain structure, and it acquires a native-like (target PDB) value only if the chain gets properly folded. Because in single molecule spectroscopy studies the distance between the ends of the chain is the usual reaction coordinate, we investigated whether this parameter is indeed a good reaction coordinate for the refolding of actin from the stretched state. As elaborated in ref 28 and references therein, the reduction of the high-dimensional protein structure space to a one-dimensional model is a substantial simplification, which may be justified in the case of protein folding with a funneled energy landscape, but it is unlikely to apply in the case of actin. Indeed, we found that the end-to-end distance is not a good reaction coordinate for any of the refolding simulation set-ups (see Figures S1−S5 in the Supporting Information). Inspection of the conformational states of actin, along the various trajectories, when the end-to-end distance reaches its value (∼2 nm) in the native state shows that, in general, the chain conformation is far from native. In many cases, these are conformations in which the individual subdomains are either misfolded or not yet fully folded. For example, in trajectory 1 from Figure S5 in the Supporting Information (red continuous curve), R reaches its native value within 10 ms from the beginning of the trajectory. However, subdomains 1C and 3B are not folded at that time. Importantly, while subdomain 1C eventually refolds after ∼30 ms, subdomain 3B remains trapped into a misfolded state for the duration of the trajectory (60 ms). Inspired by these results and the known difficulties in characterizing folding pathways using only one reaction coordinate, we relied on multiple reaction coordinates: the full rmsd, the subdomain rmsd, and the corresponding χ values. For the above example, Figures S6, S7, and S8 in the Supporting Information, depicting the time evolution of these parameters, illustrate the time when the parameters reach their native-like values. Inspection of the related chain snapshots indicates that, when all these parameters reach their native values, the entire protein chain acquires its native structure. Because, as discussed in more detail in the Supporting Information, the rmsd and χ values show transitions at the same time, in Table 1, we report the refolding success measured on the basis of the behavior of the rmsd (full chain and each subdomain) along the various trajectories.

Because the above-discussed experimental investigations into the interactions between actin and chaperones revealed that a substantial degree of flexibility is present in actin until the final stages of folding on the CCT,[6,7,11] we followed the refolding dynamics of actin from a stretched state, while allowing for various levels of chain flexibility by varying the contribution of the dihedral angle potential to the overall energy of the protein.

Refolding Simulations Using the Original SOP Model. Similar to our previous investigations into the refolding dynamics of GFP,[16] we followed the force-quench refolding of both G-ATP and G-ADP using the original SOP model. We denote this set of simulations as **D0** (dihedral potential with all parameters set to 0). This approach allowed us to follow the refolding dynamics when full flexibility toward the acquisition of secondary structure is given to the chain positions. A detailed analysis of our results is presented in the Supporting Information. The main finding is that the folding of actin, starting from a rod-like configuration resembling the newly synthesized chain from the ribosome, results in a very low refolding probability (12.5 and 0% for G-ADP and G-ATP, respectively). This low probability was due to the fact that subdomains become trapped into mirror-image states upon refolding initiated from the fully stretched state.

Refolding Simulations Using the SOP-DH Model. *G-ADP.* The results of our simulations for the refolding of actin to the G-ADP state using the **DH1** setup are depicted in Figure 2A and in Figures S5, S6, and S7 of the Supporting Information. Just as in the **D0** simulations, only one trajectory reached the correct native structure; i.e., the probability of refolding is still 12.5%. This was trajectory 6 which refolds within 20 ms to reach Δ = 0.16 nm (see Table 2). Thus, for G-ADP, the inclusion of the full dihedral angle potential in the energy function does not improve the odds of actin refolding even if no mirror-refolded structures appear. We found that the main reason for the refolding failure is the inability of subdomains 3B and 1C to reach their native structures (see Table 1 and Figures S6 and S7 of the Supporting Information). This is in stark contrast with the behavior of the chain during the **D0** refolding trajectories where the refolding failure could result from misfolding in any of the subdomains.

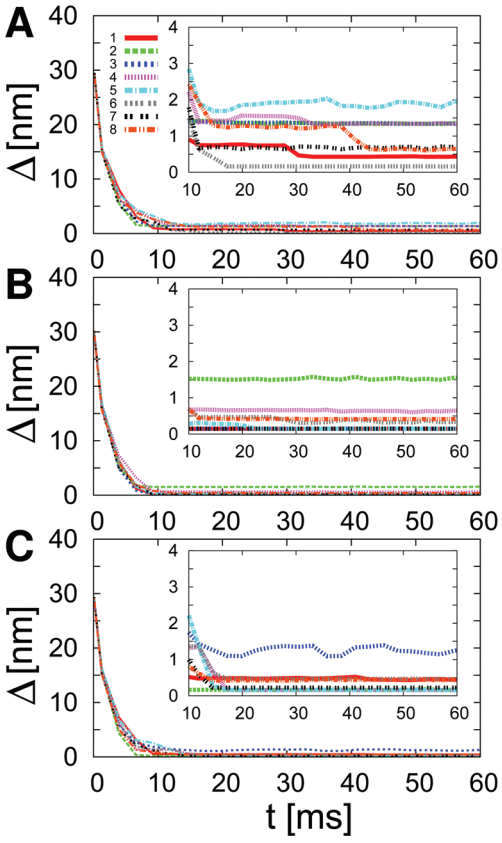


Figure 2. Root mean square deviation (Δ) for the force-quench refolding of G-ADP when considering the angle potentials along the eight trajectories. The data for **DH1**, **DH0**, and **DH10** is shown in parts A, B, and C, respectively. Each zoom-in graph has the same units as the overall graph.

Because both the 3B and 1C domains are composed of α-helices, we performed refolding simulations using the **DH0** setup; i.e., we included dihedral angle potentials only for β-sheets and loops but not for the α-helices. This choice should result in increased flexibility in the α-helical stretches of the chain leading to a delay in the folding of helices until some degree of chain collapse occurs, resulting in formation of native-like contacts. This is indeed the case. As seen in Figure 2B, 50% of the trajectories now refold to the G-ADP state because subdomain 1C refolds well in ∼75% of the trajectories and subdomain 3B is always well refolded (see Table 1 and Figure S8 of the Supporting Information). Therefore, allowing for an increased flexibility toward the formation of the α-helical stretches in the sequence leads to correct refolding of the C-term end of the chain (subdomains 3B and 1C). Moreover, the large domain (subdomains 3A, 4, and 3B) is well refolded in all the trajectories (Figure S8 of the Supporting Information). Because no mirror-image states are formed, we conclude that the presence of dihedral angle potentials for sheets and loops can prevent the improper collapse of the actin molecule.

We found that subdomains 3B and 1C are the only ones to misfold in the **DH1** simulation. To determine the origin of this behavior, we complemented the **DH0** type simulations with probes of actin's refolding ability when only the dihedral angle potentials for the helices in subdomains 3B and 1C are turned off. This **DH10** setup led to a 38% rate of success in refolding (see Figure 2C), which, in view of the modest number of trajectories, is likely to be comparable with the success for the **DH0** setup. For the **DH10** setup, we found that subdomain 1C cannot refold well in only one trajectory and subdomain 3B in another trajectory (Figure S9 in the Supporting Information), enforcing the idea that an early formation of helical stretches in these subdomains, prior to the overall collapse of the chain, is likely to lead to poorly folded conformations. While this conclusion aligns well with the above results for **DH0** refolding, an unexpected finding for the **DH10** setup is that this selective turning off for dihedral angles in only a part of the chain leads to a relatively large (38%) misfolding probability in subdomain 4 (see Table 1 and Figure S9 of the Supporting Information).

To quantify the overall effect of the early formation of persistent α-helical structure in the C-term region of the actin chain, we also carried out refolding simulations using the **DH01** setup corresponding to turning off the dihedral angle potential for all helical stretches in the protein, with the exception of those in subdomains 3B and 1C. As expected, in this case, as shown in Table 1, the rate of refolding drops from ∼50% found for the previous set-ups to only 12.5%, the same as in the **DH1** case. The analysis of the behavior of the rmsd for each subdomain reveals that the main reason for failure is the inability of subdomain 1C to refold (Table 3). For example, in trajectories 1, 2, and 4, the refolding of subdomain 1C takes much longer than in the trajectories obtained in the **DH10** case, thus confirming that an increased flexibility in the helical stretches from the C-terminal part of the chain is obligatory for the rapid and correct folding of G-ADP.

G-ATP. For the **DH1** simulations for G-ATP (Figure 3A, Table 1, and Figure S10 in the Supporting Information), similar

dx.doi.org/10.1021/jp209340y | J. Phys. Chem. B 2012, 116, 1677−1686

Table 2. Root Mean Square Deviation (Δ) Values for G-Actin and Its Three Fragments for All Trajectories[a]

		trajectory number								
		1	2	3	4	5	6	7	8	avg
G-ADP	D0	1.73	1.55	0.28	1.49	0.54	0.52	1.02	1.81	1.12
	DH1	0.43	1.36	1.33	1.34	1.96	0.16	0.70	0.65	0.99
	DH0	0.15	1.56	0.15	0.65	0.15	0.34	0.15	0.41	0.44
	DH10	0.44	0.17	1.24	0.16	0.16	0.46	0.22	0.45	0.41
	DH01	0.20	0.15	1.86	0.15	0.21	0.24	0.15	1.97	0.62
	DH1′	0.16	0.16	1.29	0.84	1.39	0.29	0.17	1.33	0.70
	DH1′	1.33	0.16	0.86	1.81	0.41		1.51	2.08	1.17
	DH0A	0.23	0.49	0.20	0.14	0.15	0.15	0.16	0.15	0.21
	DH0B	0.47	0.15	0.15	0.41	0.16	0.15	1.81	0.15	0.43
	DH0C	0.92	0.21	0.21	0.55	0.37	0.21	1.00	0.85	0.54
G-ATP	D0	1.47	1.32	1.30	1.53	1.54	0.57	1.54	1.33	1.32
	DH1	0.19	0.72	0.89	0.98	2.25	0.82	1.67	1.88	1.18
	DH0	0.49	1.95	0.54	0.61	1.73	0.52	0.54	0.55	0.87
	DH10	0.54	0.54	1.43	0.62	1.20	0.54	0.27	0.19	0.67
	DH01	0.35	0.87	0.73	0.70	1.94	0.50	0.92	0.73	0.84
	DH1′	0.16	0.16	1.29	0.84	1.39	0.29	0.17	1.33	0.70
	DH1′	1.39	0.74	0.77	0.82	0.56	0.74	0.23	0.73	0.75
	DH0A	1.26	1.00	0.51	0.56	0.62	1.08	0.19	0.66	0.74
	DH0B	0.54	0.19	0.56	0.43	0.46	0.95	0.31	1.01	0.56
	DH0C			0.85	0.58	0.45	0.25	0.57	0.23	0.49

[a]Average values over eight trajectories, except for DH0C in G-ATP for which we only ran six trajectories, are shown in the avg column. All values are taken from the end of the 60 ms run time. The units are nm.

Table 3. Root Mean Square Deviation (Δ) Values for the Whole Actin Molecule and Its Four Subdomains Reached during the Force-Quench Refolding of G-ADP and G-ATP[a]

		D0	DH1	DH0	DH10	DH01	DH1′	DH1′	DH0A	DH0B	DH0C
G-ADP	whole	1.12	0.99	0.44	0.41	0.62	0.70	1.17	0.21	0.43	0.54
	subdomain1	1.04	1.31	0.50	0.31	0.58	0.90	1.38	0.18	0.17	0.57
	subdomain2	0.42	0.12	0.23	0.16	0.15	0.20	0.40	0.15	0.11	0.34
	subdomain3	0.53	0.23	0.08	0.11	0.17	0.15	0.25	0.13	0.34	0.08
	subdomain4	0.72	0.19	0.15	0.33	0.21	0.18	0.20	0.16	0.33	0.24
G-ATP	whole	1.32	1.18	0.87	0.67	0.84	0.61	0.75	0.74	0.56	0.49
	subdomain1	1.00	1.50	0.49	0.54	0.98	0.69	0.90	0.63	0.39	0.14
	subdomain2	0.50	0.40	0.29	0.32	0.34	0.34	0.35	0.39	0.45	0.36
	subdomain3	1.17	0.24	0.18	0.17	0.11	0.08	0.11	0.18	0.20	0.11
	subdomain4	0.85	0.34	0.61	0.49	0.39	0.34	0.28	0.55	0.47	0.56

[a]The values were obtained at 60 ms as average values over all the given types of trajectories, and all units are in nm.

to the G-ADP case, we found that only one trajectory leads to a well-refolded structure characterized by a small Δ value of 0.19 nm (see Table 2). Thus, neither G-ATP nor G-ADP are able to refold well in the **DH1** case. However, in contrast with the G-ADP case when only subdomains 3B and 1C misfold, in the G-ATP case, most of the subdomains have problems refolding (see Figure S10 of the Supporting Information). The only exceptions are subdomains 1A and 3A, which refold well in all trajectories. This increased rate of failure throughout the sequence is likley due to the more flexible character of G-ATP compared with G-ADP.[17]

Following our approach for G-ADP, we also performed **DH0** simulations for G-ATP. The change in the overall chain rmsd from Figure 3B indicates that the average Δ is now 0.87 nm, i.e., smaller than the 1.18 nm obtained for the **DH1** simulations (see Table 2). However, the **DH0** setup does not lead to any well-refolded trajectory for G-ATP (see Figure S11 in the Supporting Information). This is in stark contrast with the G-ADP case where 50% of the trajectories were refolding well. The main reason for the failure in the G-ATP case is the substantial rate of misfolding (75%) in subdomain 4, but we found that subdomains 1C, 3B, and 2 also misfold in a number of trajectories (Figure S11 of the Supporting Information and Table 3).

Figure 3C depicts the results of our **DH10** simulations for G-ATP, which indicate that this setup leads to a further reduction of the overall rmsd to 0.67 nm. Despite this improvement, the probability of refolding is still as low as that for the **DH1** setup (1 trajectory out of 8, i.e., 12.5% rate). In these simulations, subdomain 1C is well refolded (75% rate of success), similar to the **DH0** case (see Figure S12 in the Supporting Information). However, as illustrated in Tables 2 and 3, subdomain 4 is again the most difficult to refold (50% rate of success), followed by subdomain 3B (62% rate of success), ultimately leading to a low overall refolding rate for the actin molecule.

Similar to our approach for the G-ADP refolding, to quantify the effect of the early formation of persistent α-helical structure in the C-term region of the actin chain, we also carried out refolding simulations using the **DH01** setup. In this case, as shown in Table 1, the rate of success in refolding is the same

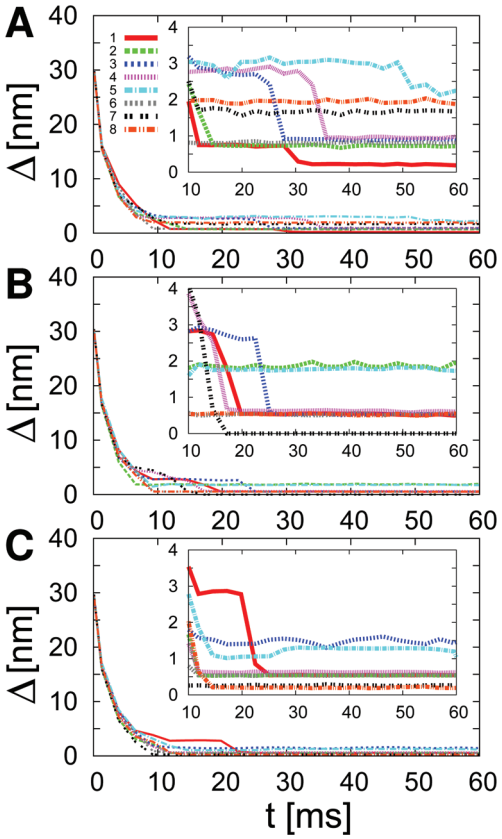


Figure 3. Root mean square deviation (Δ) for force-quench refolding of G-ATP when considering the angle potentials along the eight trajectories. The data for **DH1**, **DH0**, and **DH10** is shown in parts A, B, and C, respectively.

modest 12.5% observed for the **DH1** and **DH10** set-ups. Moreover, the analysis of the behavior of the rmsd for each subdomain reveals that subdomain 1C has problems in refolding (Tables 1 and 3).

These results strongly suggest that, unlike for the G-ADP state, an increased flexibility of the α-helical stretches from the subdomains 3B and 1C is not sufficient for the rapid and correct refolding of G-ATP.

In all of the above simulations, we followed the setup from AFM-based refolding experiments by fixing the C-terminal end of the chain. This in turn leads to a reduction in the entropy of the end. To evaluate the role played by the reduction in the entropy on the refolding dynamics of actin, we also conducted refolding simulations for both G-ATP and G-ADP in the **DH1** dihedral angle potential setup but without fixing any position along the chain, which we denote the **DH1′** setup. This is akin to the folding of the chain in bulk. As seen in Table 1, the rate of success is unchanged (12.5%) compared with the **DH1** setup, which implies that changes in the entropy of the C-terminal end of G-actin do not influence the refolding probability.

Refolding Simulations for Selected Fragments of G-Actin: DH0A, DH0B, and DH0C. To mimic the environment encountered by the actin chain while being translated in the ribosome, we followed, using the **DH0** setup, the force-quench induced refolding of three fragments of G-ADP lacking C-terminal positions. Namely, the fragments are the following: (1) the 1-356 region of the actin chain, i.e., lacking a part of the 1C subdomain (this is denoted by **DH0A**), (2) the 1-337 region, i.e., lacking the whole 1C subdomain (denoted by **DH0B**), and (3) the 1-295 fragment lacking part of the 3B subdomain and the whole 1C subdomain (denoted by **DH0C**). The rationale for selecting this set of fragments stems from the presence of rare codons at the end positions for each fragment. It is also important to note that, while the other two fragments end in just one rare codon position, the **DH0C** fragment has four rare codon positions at its C-term end. The results of this set of simulations, shown in Table 1, indicate that the **DH0A** and **DH0B** fragments from G-ADP have high refolding probability (50%), very similar to the one found for the whole chain during the **DH0** type simulations. In contrast, the **DH0C** fragment, which is missing the largest C-term part, has considerably lower refolding probability (25%). Moreover, in the **DH0B** case, the average rmsd value for the molecule is as large as 0.67 nm (see Table 3). We conclude that the fragment leading to the best refolding rate and overall well-refolded structures for G-ADP is **DH0A**. In stark contrast, for the G-ATP state, the fragment with the best refolding probability is **DH0C** (38%). This strongly suggests that, even if the C-terminal part of the chain (subdomains 3B and 1C) has the most difficulty refolding, their presence facilitates the proper refolding of the rest of the chain into the G-ADP state while it impedes the correct folding of the chain into the G-ATP state.

Refolding Pathway Diversity. Analysis of the refolding trajectories for both G-ATP and G-ADP from Table 1 revealed the presence of two pathways (see Table S3 in the Supporting Information). The fundamental difference between the two pathways is illustrated in Figures 4 and 5, which depict the

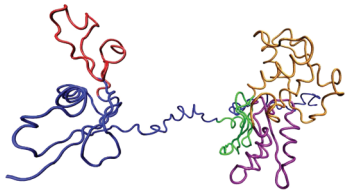


Figure 4. Main intermediate state for refolding of actin along pathway 1.

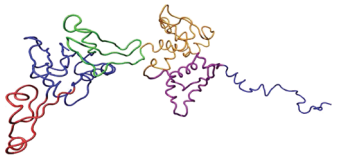


Figure 5. Main intermediate state for refolding of actin along pathway 2.

major intermediate along each pathway. In pathway 1, the intermediate state (Figure 4) consists of the formation of a large contact surface between subdomains 3A (in green) and 3B (in magenta) before 3A binds to the folded part of subdomain 1A, 1B (in blue). Thus, in pathway 1, the small and large domains appear to refold independently of each other and the refolding of the two domains occurs simultaneously. In contrast, in the intermediate state from pathway 2, depicted in Figure 5, subdomain 3A binds to subdomains 1A and 1B prior to formation of its native contacts with the 3B subdomain. This delays the refolding of subdomain 3 compared to pathway 1 and indicates that the small and large domains do not refold independently.

dx.doi.org/10.1021/jp209340y | J. Phys. Chem. B 2012, 116, 1677–1686

Representative Well-Refolded Case. In Figure 6, we present the succession of configurations that leads to the successful

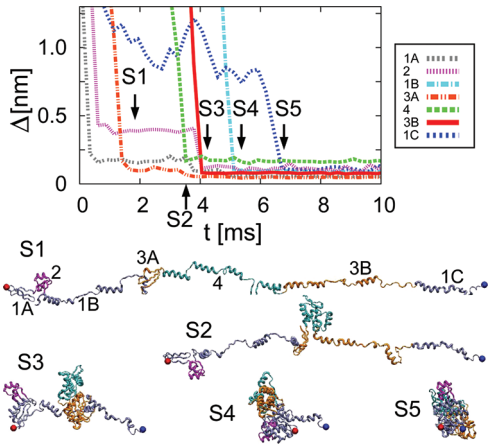


Figure 6. Representative well-refolded pathway (trajectory 2) of DH10 for G-ADP. The evolution of the root-mean-square deviation (Δ) for each subdomain part is shown. The five refolding events (S1, S2, S3, S4, and S5) are indicated by arrows in rmsd change, and their corresponding structures are shown below.

refolding of G-actin according to pathway 1 from above, which is the dominant refolding route (Table S3 in the Supporting Information). The states have been selected from trajectory 2 of the **DH10** setup for G-ADP, when the whole actin molecule is refolded within 7 ms. As shown in the upper panel of Figure 6, the refolding starts at the N-terminal end of the chain with subdomain 1A (gray dashed line). This is followed very soon by the folding of subdomains 2 (magenta dashed line) and 3A (orange dashed line) as depicted in the S1 state. Next, subdomain 4 refolds in the S2 state and subdomain 3B refolds in the S3 state, leading to the formation of the native state of the large domain. In this S3 state, the partially refolded small and large domains are separated by the still unfolded subdomain 1B. Also, subdomain 1C is still unfolded. The last two steps correspond to the refolding of subdomains 1B and 1C in states S4 and S5, repectively, resulting in the acquisition of the native structure of the whole actin molecule. Because the C-terminal part of the chain is the most likely region to misfold, we analyzed in detail the process of refolding in subdomain 1C (see Figure 7). Figure 7A shows the behavior of this subdomain along the trajectory analyzed above when it refolds well. The helix marked by an arrow is very flexible because it does not present regular helix turns. As depicted in the three bottom panels in Figure 7, this flexible fragment acquires its native helical form turn by turn upon folding in the cleft between subdomains 1B and 3. In contrast, along a trajectory that leads to a misfolded G-ADP structure corresponding to the **DH1** setup and illustrated in Figure 7B, the helix has regular turns and it is already very rigid before formation of its tertiary contacts. This rigidity prevents the correct formation of its tertiary contacts, resulting in overall misfolding. Thus, our results indicate that, in the C-terminal part of actin, the formation of the tertiary fold is a prerequisite for the formation of the secondary structure.

A successful refolding of G-actin along pathway 2 starts the same way as in pathway 1, i.e., from its N-term end with the refolding of subdomains 1A, 2, and 3A (in state S1). The next step in refolding (S2) now corresponds to the refolding of

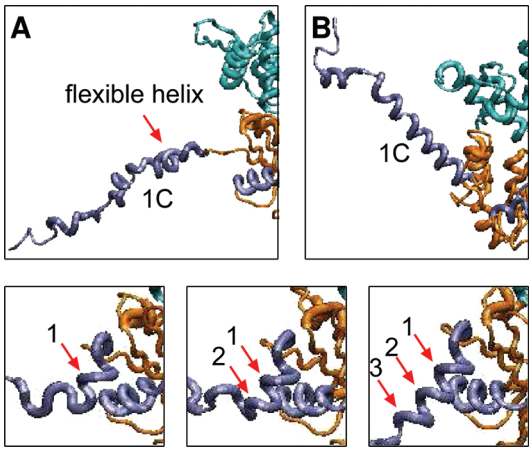


Figure 7. Helix structure formation in subdomain 1C during the (A) DH10 type simulations and (B) DH1 type simulations for G-ADP. The folding process of the helical region (the flexible helix marked by an arrow in part A) along a successful DH10 trajectory is shown on the bottom.

subdomain 3B. Subdomain 1B refolds in step S3, just prior to subdomain 4 which is refolded at step S4. The last event corresponds, just as above, to the refolding of subdomain 1C at step S5. Thus, the main difference between the refolding of G-actin along the two pathways is the delay of the refolding of subdomain 4. This subdomain refolds in step S2 for pathway 1 but only in S4 for pathway 2, thus resulting in a delay (up to state S4) in the correct refolding of the large domain of G-actin.

■ DISCUSSION AND CONCLUSIONS

Our SOP-DH model based approach is along the lines of recent methodological developments designed to track the role of secondary structure elements (SSEs) in guiding protein folding.[29−31] In general, these methods have been applied to the folding of relatively small proteins, while our method addresses important folding aspects in a much larger protein characterized by a long folding time and the inability to fold properly in the absence of molecular chaperones. Thus, the agreement between our findings and a large body of experimental results gives additional confidence in the use of SSE-based approaches for identifying major (topological) hurdles in the folding process. Also, our results allow us to make a number of predictions regarding bottlenecks in the refolding of G-actin.

A summary of our findings is presented in Figure 8. We found that the early formation of helical stretches along the chain (the **DH1** setup) has deleterious effects on the refolding fitness of subdomain 1. Namely, for G-ADP (G-ATP), the average rmsd for this subdomain is 1.31 (1.50) nm, which is much larger than the corresponding 0.1−0.2 (0.2−0.4) nm for the other three subdomains. Allowing for an increase in the flexibility of the α-helices (the **DH0** or **DH10** set-ups) leads to a dramatic reduction in the rmsd of subdomain 1 to 0.49−0.50 nm (in the **DH0** case). Thus, the simulation setup that leads to the best results in both G-ATP and G-ADP is **DH10**. On the basis of these results, we predict that the folding of the C-terminal part of G-actin evolves through the simultaneous formation of the secondary and tertiary structures resulting in the zipping up of α-helices one turn at a time. Our simulations reveal that this is a rare event, which signals the presence of a large barrier. Thus, the correct refolding of the C terminus will

dx.doi.org/10.1021/jp209340y | J. Phys. Chem. B 2012, 116, 1677−1686

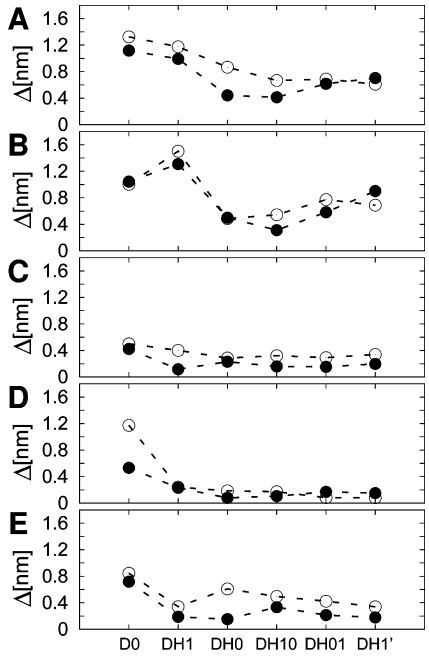


Figure 8. Root mean square deviation (Δ) values for force-quench refolding of G-ADP and G-ATP. The data for **D0**, **DH1**, **DH0**, **DH10**, **DH01**, and **DH1**′ is shown. Filled and empty circles correspond to results for G-ADP and G-ATP, respectively. The values were obtained at 60 ms as an average over the eight trajectories, and the unit is nm. The rmsd for the whole actin molecule is in part A, while the rmsd values for each of its subdomains 1, 2, 3, and 4 are shown in parts B, C, D, and E, respectively.

be facilitated by an apport of energy. Our proposal correlates well with recent experimental studies,[6] which found that the addition of ATP to CCT allows the C-term end of G-actin to fold properly on the intermediate folded structure. The authors proposed that the rearrangement of the C-terminus end of G-actin occurs as ATP is hydrolyzed, i.e., upon addition of energy in the system. Willison and collaborators[32] also found that the activation energy between the native G-actin state and an unfolding intermediate, which cannot refold spontaneously, is indeed large (∼100 kJ/mol), as corroborated by our results.

In G-ATP, besides the C-terminal end subdomains 3B and 1C, which tend to misfold in the majority of the trajectories, subdomains 2 and 4 have the largest misfolding probability: 30% for subdomain 2 and 44% for subdomain 4 (see Table 1 and discarding the data for the **D0** setup). Thus, our simulations indicate that controlled folding of these two subdomains is one requirement for the global folding of G-actin. This finding is important in light of experimental investigations which found that subdomains 2 and 4 contain many of the CCT binding sites of actin reviewed in the Introduction. For example, experimental studies[9,33,34] found that subdomain 4 interacts directly with CCT in order to induce refolding of actin, as well as it contains binding determinants for actin on prefoldin (PFD). On the basis of this data and our results which reflect the inherent topological frustration in the G-actin chain, we hypothesize that one reason for the early attachment of subdomain 4 on PFD and of subdomains 2 and 4 on CCT is to allow them to fold under controlled conditions which prevent misfolding resulting from a too rapid collapse of the chain.

Our refolding pathway 1 described above (see Figure 6) agrees with the experimental folding model of Neirynck et al.[11,21] Starting from an alanine scan of actin, these authors reported the following multistep model for the folding of actin through the interaction with CCT: (A) Actin is transferred from PFD to ADP-CCT in such a way that residues 30−34 and 245−249 contact CCT. The two small and large domains are nearly folded (but separate), except that the (unstructured) C-term is not located in subdomain 1. (B) ATP binding to CCT induces conformational changes in CCT, leading to the release of the small actin domain, which contacts the large domain, resulting in the closure of actin. (C) CCT releases the large domain of actin and simultaneously rebinds to the 135−139 and 170−174 and/or 340−349 actin sites. (D) The folding process ends upon formation of the correct set of contacts between the C-term end and subdomain 1 (e.g., the 360−364 region interacts with 125−129). The A state is very similar to the S3 state in our simulation in that both the large and small domains are folded, with the exception of the C-terminus end, and are separated by a hinge region.[11,35] The B state is similar to the S4 state from our simulations in which the residue stretch 135−139 (subdomain 1B) is refolded well due to the overall collapse of the chain and formation of the contact interface between the small and large domains. The last state (S5) from our simulations corresponds to step D from the experimental model, in which subdomain 1C refolds well upon formation of native tertiary contacts with the 125−129 region of the chain.

Our results enable us to make additional predictions. First, we found that the G-ATP configuration is almost impossible to reach within the 60 ms time scale of our simulations, likely due to its increased flexibility compared with G-ADP. This is akin to the known cellular behavior of G-actin, which leads us to propose that the nucleotide-free state of G-actin, which is associated with CCT, is closer to G-ATP than to G-ADP. This proposal correlates well with the experimental findings that the nucleotide free actin is able to refold when it is stabilized in high concentrations of sucrose to prevent denaturation[36] and that the G-ATP form is "floppy" like the nucleotide-free form. In general, G-ATP actin collapses and stiffens upon polymerization and hydrolysis to the ADP state.[37] Second, because our model which lacks energetic frustration still leads to a dismal folding probability for the G-ATP state, this strongly suggests that topological frustration plays a very important role in the folding of actin. This proposal is well supported by the fact that actin requires a set of chaperones to acquire its functional form. Lastly, our simulation results suggest that there are two folding scenarios in G-actin. However, as discussed before, only the experimental evidence for the main intermediate along pathway 1 is strong. This suggests that folding according to the second scenario is not biologically favorable. Because formation of this intermediate occurs early in the folding process, i.e., prior to delivery to CCT, we hypothesize that one role of PFD is to drastically reduce the probability of the second folding scenario, by preventing the formation of this intermediate. Moreover, the existence of the two folding scenarios from our simulations might account for the yeast ACT1/rabbit α-actin incompatibility.[7] For example, a substantial change in the probability of each scenario for the folding of mammalian skeletal α-actin in the presence of yeast CCT can be responsible for the inability of yeast CCT to fold rabbit α-actin. This hypothesis can be tested using our SOP-DH approach when X-ray structures of the mutants of rabbit α-actin, such as Asn-297/Ile-297, which have been shown to account for species folding specificity,[7] become available.

dx.doi.org/10.1021/jp209340y | J. Phys. Chem. B 2012, 116, 1677−1686

■ ASSOCIATED CONTENT

Supporting Information

Additional details regarding the choice of the reaction coordinate for the refolding reaction are presented. Two tables listing the values of the parameters used in our model are provided. A table listing the type of pathway followed by each of the simulation runs is also provided. Figures depicting the time series of various reaction coordinates along various trajectories are also included. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Ruxandra.Dima@uc.edu. Phone: +1 513 5563961. Fax: +1 513 5569239.

■ ACKNOWLEDGMENTS

This work was partially supported in part by the National Science Foundation grant MCB-0845002. We thank two anonymous reviewers for their help in improving the quality of the manuscript.

■ REFERENCES

(1) Mofrad, M. R. K.; Kamm, R. D. *Cytoskeletal mechanics: Models and Measurements*; Cambridge University Press: Cambridge, MA, 2007.
(2) Carballido-López, R. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 888−909.
(3) Doherty, G. J.; McMahon, H. T. *Annu. Rev. Biophys.* **2008**, *37*, 65−95.
(4) Kabsch, W.; Mannherz, H. G.; Suck, D.; Pai, E. F.; Holmes, K. C. *Nature* **1990**, *347*, 37−44.
(5) Otterbein, L. R.; Graceffa, P.; Dominguez, R. *Science* **2001**, *293*, 708−711.
(6) Stuart, S. F.; Leatherbarrow, R. J.; Willison, K. R. *J. Biol. Chem.* **2011**, *286*, 178−184.
(7) Altschuler, G. M.; Dekker, C.; McCormack, E. A.; Morris, E. P.; Klug, D. R.; Willison, K. R. *FEBS Lett.* **2009**, *583*, 782−786.
(8) Brackley, K. I.; Grantham, J. *Exp. Cell Res.* **2010**, *316*, 543−553.
(9) Llorca, O.; McCormack, E. A.; Hynes, G. M.; Grantham, J.; Cordell, J.; Carrascosa, J. L.; Willison, K. R.; Fernandez, J. J.; Valpuesta, J. M. *Nature* **1999**, *402*, 693−696.
(10) Hynes, G. M.; Willison, K. R. *J. Biol. Chem.* **2000**, *275*, 18985−18994.
(11) Neirynck, K.; Waterschoot, D.; Vandekerckhove, J.; Ampe, C.; Rommelaere, H. *J. Mol. Biol.* **2006**, *355*, 124−138.
(12) Hansen, W. J.; Cowan, N. J.; Welch, W. J. *J. Cell Biol.* **1999**, *145*, 265−277.
(13) Zhang, G.; Ignatova, Z. *Curr. Opin. Struct. Biol.* **2011**, *21*, 25−31.
(14) Wilson, D. N.; Beckmann, R. *Curr. Opin. Struct. Biol.* **2011**, *21*, 274−282.
(15) Kosolapov, A.; Deutsch, C. *Nat. Struct. Mol. Biol.* **2009**, *16*, 405−411.
(16) Hyeon, C.; Dima, R. I.; Thirumalai, D. *Structure* **2006**, *14*, 1633−1645.
(17) Lee, J. Y.; Iverson, T. M.; Dima, R. I. *J. Phys. Chem. B* **2011**, *115*, 186−195.
(18) Mickler, M.; Dima, R. I.; Dietz, H.; Hyeon, C.; Thirumalai, D.; Rief, M. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20268−20273.
(19) Dima, R. I.; Joshi, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 15743−15748.
(20) Liu, Z.; Reddy, G.; O'Brien, E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 7787−7792.
(21) Rommelaere, H.; Waterschoot, D.; Neyrinck, K.; Vandekerckhove, J.; Ampe, C. *Structure* **2003**, *11*, 1279−1289.
(22) Veitshans, T.; Klimov, D. K.; Thirumalai, D. *Fold. Des.* **1996**, *2*, 1−22.
(23) Fernandez, J. M.; Li, H. *Science* **2004**, *303*, 1674−1679.
(24) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(25) Graceffa, P.; Dominguez, R. *J. Biol. Chem.* **2003**, *278*, 34172−34180.
(26) Klimov, D. K.; Betancourt, M. R.; Thirumalai, D. *Fold. Des.* **1998**, *3*, 481−496.
(27) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2544−2549.
(28) Best, R. B.; Paci, E.; Hummer, G; Dudko, O. K. *J. Phys. Chem. B* **2008**, *112*, 5968−5976.
(29) Ozkan, S. B.; Wu, G. A.; Chodera, J. D.; Dill, K. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11987−11992.
(30) Balaraman, G. S.; Park, I. H.; Jain, A.; Vaidehi, N. *J. Phys. Chem. B* **2011**, *115*, 7588−7596.
(31) Camilloni, C.; Broglia, R. A.; Tiana, G. *J. Chem. Phys.* **2011**, *134*, 045105.
(32) Altschuler, G. M.; Klug, D. R.; Willison, K. R. *J. Mol. Biol.* **2005**, *353*, 385−396.
(33) Llorca, O.; Martin-Benito, J.; Ritco-Vonsovici, M.; Grantham, J.; Willison, G. M. H. K. R.; Carrascosa, J. L.; Valpuesta, J. M. *EMBO J.* **2000**, *19*, 5971−5979.
(34) Rommelaere, H.; Neve, M. D.; Neirynck, K.; Peelaers, D.; Waterschoot, D.; Goethals, M.; Fraeyman, N.; Vandekerckhove, J.; Ampe, C. *J. Biol. Chem.* **2001**, *276*, 41023−41028.
(35) McCormack, E. A.; Llorca, O.; Carrascosa, J. L.; Valpuesta, J. M.; Willison, K. R. *J. Struct. Biol.* **2001**, *135*, 198−204.
(36) De La Cruz, E. M.; Mandinova, A.; Steinmetz, M. O.; Stoffler, D.; Abi, U.; Pollard, T. D. *J. Mol. Biol.* **2000**, *295*, 517−526.
(37) Altschuler, G. M.; Willison, K. R. *J. R. Soc., Interface* **2008**, *5*, 1391−1408.
(38) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.