Structural Polymorphism of Actin

Toshiro Oda [1], Shuichi Takeda [2], Akihiro Narita [2] and Yuichiro Maéda [2,3]

*1 - **Faculty of Health and Welfare,** Tokai Gakuin University, 5-68 Nakakirino-cho, Kakamigahara, Gifu 504-8511, Japan*
*2 - **Structural Biology Research Center,** Graduate School of Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan*
*3 - **Toyota Physical and Chemical Research Institute,** 41-1 Yokomichi, Nagakute, Aichi 480-1192, Japan*

***Correspondence to Toshiro Oda:** toda375@tokaigakuin-u.ac.jp, takeda.shuichi@f.mbox.nagoya-u.ac.jp, narita.akihiro@f.mbox.nagoya-u.ac.jp, ymaeda@cc.nagoya-u.ac.jp*
https://doi.org/10.1016/j.jmb.2019.05.048
Edited by James Sellers

Abstract

Information on the structural polymorphism of a protein is essential to understand the mechanisms of how it functions at an atomic level. Numerous studies on actin have accumulated substantial amounts of information about its polymorphism, and there are over 200 published atomic structures of different forms of actin using crystallography, fiber diffraction, and electron microscopy. To characterize all the reported structures, we proposed simple parameters based on the discrete rigid bodies within the actin molecule and identified four conformation groups by cluster analysis: the F-form in naked F-actin, the C-form in cofilactin, the O-form in profilin–actin, and the G-form in the majority of actin-containing crystal structures. The G-form group included the most variations, but each conformational variation was convertible via a thermal fluctuation, whereas the F- and C-forms were not accessible from the G-form. The convertibility and accessibility of the structures were evaluated using molecular dynamics simulations. Information about conformational conversion among each group is useful for understanding the mechanisms of actin function.

Introduction

Actin is a central player in dynamic cellular activities such as cell division and cell motility [1,2]. Elucidation of the basic functions of actin is essential to completely understand these events. Generally, conformational changes are often involved in protein functions. To understand the mechanism(s) of how a particular protein functions at an atomic level, it is essential to obtain information on the various conformations that it can adopt, that is, its polymorphism. Numerous studies have revealed that actin is highly polymorphic. Over 200 atomic structures of actin have been reported by crystallography, fiber diffraction, and electron microscopy, and to date, at least four groups of distinct actin conformations have been identified. The first conformation, the G-form, has been reported in many crystal structures of monomeric actin and is characterized by a closed nucleotide-binding cleft. The second group, the F-form, has been observed in structures of naked actin filaments [3–5] and actin filaments complexed with tropomyosin [6]. This conformation has a closed cleft, and, compared with the G-form, one domain is rotated in an anti-clockwise direction with respect to the other, generating a flatter overall shape (Figs. S1 and S2). The third group, the C-form, has been observed in cofilin-decorated actin filaments (cofilactin). This conformation also possesses a closed cleft, but has a distinctive clockwise domain rotation [7] (Figs. S1 and S2). The fourth conformation, the O-form, has been reported in phosphate-treated crystals of profilin–actin complexes and is characterized by a wide-open nucleotide-binding cleft [8,9] (Fig. S2). Other minor conformational variations have also been reported [10–12]. Structural changes in the actin molecule between these various forms will affect the environment of cleft-bound nucleotides

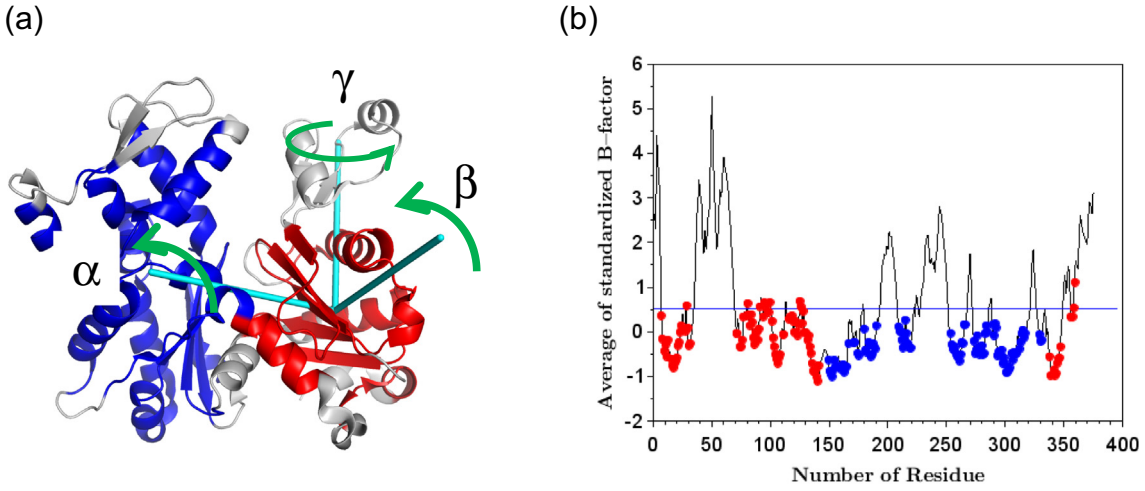

Fig. 1. Definition of the core domains. (A) Structure of actin (PDB accession number: 1J6Z). The outer domain (OD) and the inner domain (ID) cores are colored red and blue, respectively. The cyan lines represent the three Cartesian axes used for describing the OD core orientation. (B) Graph showing standardized *B*-factor averages for 79 actin crystal structures plotted against the residue number. The red and blue points represent the OD and ID domain core residues, respectively, and the blue line represents the mean value of averages of the standardized *B*-factors.

such as ATP, which in turn drives actin functions such as cycles of polymerization and depolymerization through hydrolysis of bound ATP and exchange of bound nucleotides.

Furthermore, local polymorphisms of loops within the actin structure have been reported. In particular, the D-loop has been studied extensively owing to its relatively easy accessibility by biochemical techniques such as restricted enzymatic cleavages. The D-loop is only visible, that is, with significant electron densities, when it is in contact with other molecules in a crystal, and its conformation is also dependent upon contact with molecules such as DNase I and thymosin β-4. Therefore, this polymorphism appears to be passive and may not directly drive actin functions relevant to energy consumption.

To discuss the global polymorphism of actin, it is necessary to provide a clear definition of its various forms. In 2010, to clearly distinguish between the G- and F-forms, we proposed a parameter expressing the rotation of one domain with respect to the other about an axis determined by the G- to F-actin transition [13] (Fig. S1). However, as described above, other actin conformations have recently been described. Therefore, a new set of descriptions are required to provide a clear distinction between each of the forms that will aid discussions of the structure–function relationships. In this report, we propose simple, visible, and sharply defined parameters based on discrete rigid bodies identified within the actin molecule [7] and discuss their potential applications.

Results

Definition of the inner and outer domains of actin

Actin has two major domains enclosing a nucleotide-binding cleft. These are termed the inner domain (ID) and the outer domain (OD) owing to their relative distances from the helix axis within an actin filament. The ID is composed of subdomain 3 (residues 145–180 and 270–337) and subdomain 4 (181–269), and the OD is composed of subdomain 1 (residues 1–32, 70–144, and 338–375) and subdomain 2 (33–69), as defined by Kabsch *et al*. [14].

The main chains in the periphery of the two major domains that include the D-loop, residues 200–206, and the C-terminus, are flexible and adopt several different conformations [13,15]. In contrast, it appears that the core of each major domain has an invariant structure and moves as a rigid body. Tanaka *et al*. [7] used eight available actin structures [PDB accession numbers: 1J6Z, 4PKH, 3DAW, 5JLF, fragmin–actin complex (unpublished data), 2PAV, 4PL7, 5YU8] to define the residues comprising the rigid bodies of the two major domains by setting a threshold value (<0.7 Å) for the standard deviation calculated for each Cα position. In this study, only the residues defined by Tanaka *et al*. that are visible in all 127 actin crystal structures at a resolution of at least 3.5 Å were included in the core domains. The OD core was specified as composing of residues

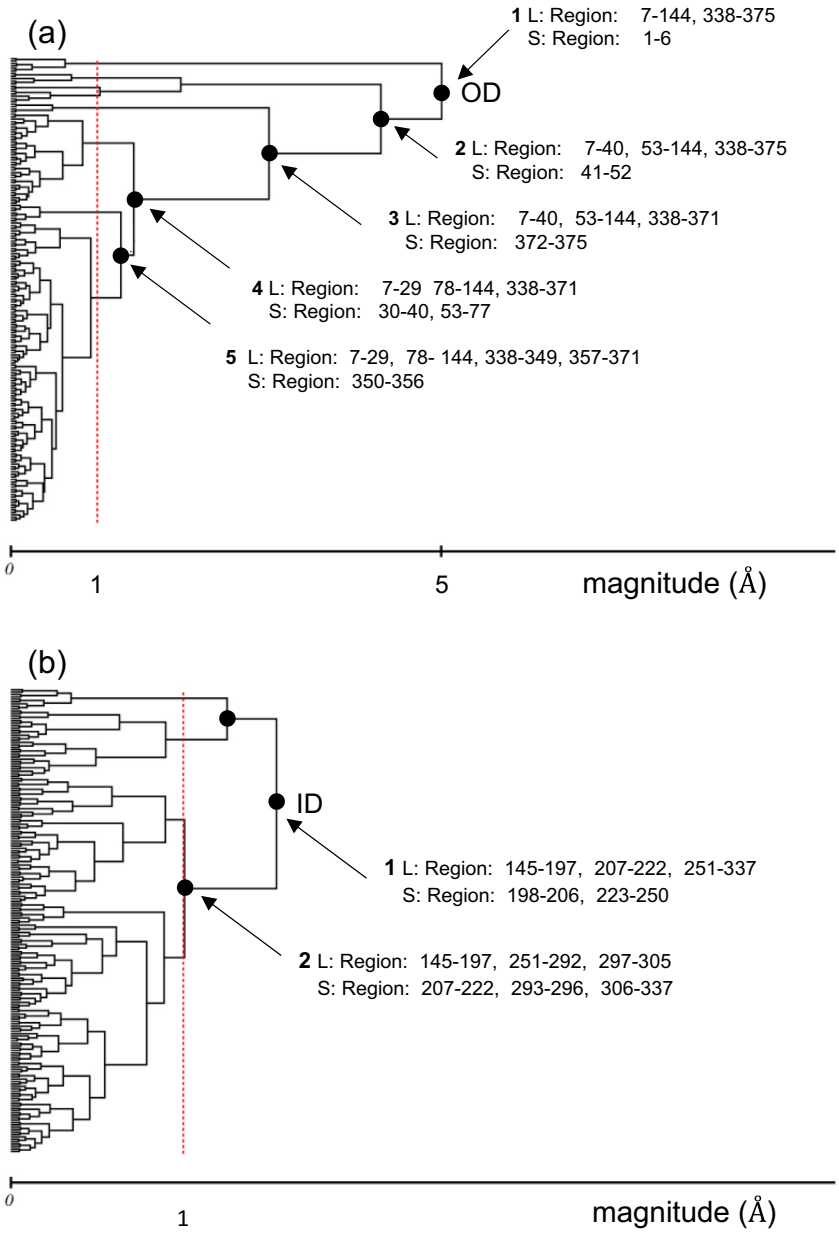

Fig. 2. Motion tree for the outer domain (A) and inner domain (B) using the MD trajectory of actin (PDB accession number: 1WUA). Each node shows integration of two rigid bodies [the L (large) and S (small) part]. The height from the bottom indicates the magnitudes of a relative motion of two rigid bodies integrated at the node. Red dashed lines indicate 1-Å level of distance between clusters.

7–31, 70–71, 76–108, 115–137, 338–347, and 355–364, and the ID core as residues 147–166, 171–194, 205–229, 237–238, 251–322, and 327–337 (Fig. 1A).

To characterize the domain cores by the internal motions of the two major domains, we created a "motion tree" to describe the structural dynamics using hierarchically defined rigid-body pairs on the basis of RMSD values [16]. Figure 2A shows a typical motion tree for the OD, which was calculated from frames of 200–500 ns in the trajectory of molecular dynamics (MD) simulation of actin (chain A of 1WUA); the details of the MD simulation are described in the later section. The root of the tree denotes the cluster composed all residues in the OD of actin. The root is divided into the N-terminus (residues 1–6) and the larger rest of the OD (residues 7–144, 338–375) at node 1, the height of which from the bottom represents a magnitude of the motion of N-terminus relative to the rest part. The rest part is subdivided, and small parts are sequentially removed: the D-loop (residues 41–52) at node 2, the C-terminus (residues 372–375) at node 3, and the neighbor region of the D-loop (residues 30–40 and 53–77) at node 4. The rest part obtained at node 4 (residues 7–29, 78–144, and 338–371) corresponds to the OD core described above. Further removal of small

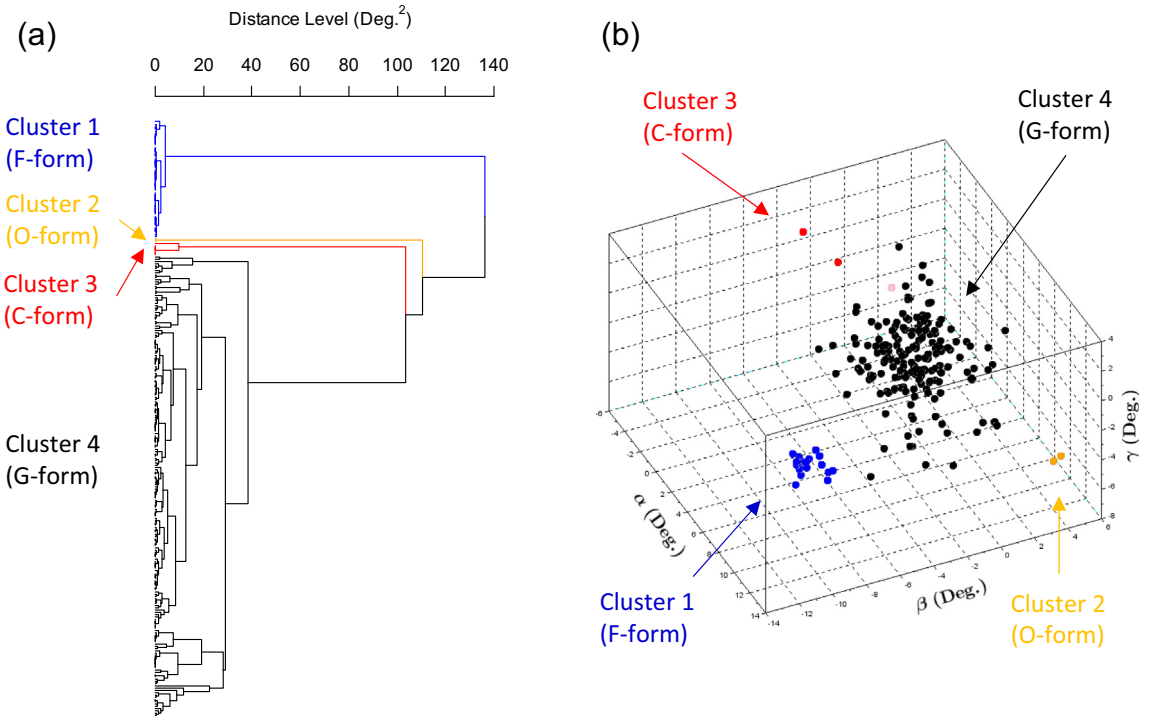

Fig. 3. Clustering of outer domain orientations. (A) Dendrogram showing classification of actin conformations using the group average method. Four clusters separated are colored [21]. Cluster 1: blue, Cluster 2: orange, Cluster 3: red, and Cluster 4: black. (B) Orientations of the outer domain (OD) core in the αβγ-space. F-form: blue, C-form: red, O-form: orange, and G-form: black. The pink dot represents the origin of the plot, indicating the OD orientation of 1J6Z.

parts only reduces the height of nodes slightly. This suggests that magnitudes of motion for parts constituting the OD core are comparable. Thus, motions of the OD core can be assumed to be representative of the global motions of the OD, excluding local motions of the N-terminus, D-loop, and C-terminus. Similarly, detected motions of the ID core would be unaffected by local motions of the 198–206 and 223–250 regions (Fig. 2B). Motion trees derived from the MD trajectories of other actin structures exhibited similar patterns.

Furthermore, we calculated the standardized *B*-factor for each Cα in 79 reported crystal structures of skeletal muscle α-actin at a minimum resolution of 3.0 Å. Each Cα present in the core domains had a lower *B*-factor value, which is consistent with lower mobility (Fig. 1B). Although a set of residues in the domain core was altered by a change of the RMSD threshold value and/or altering the sets of analyzed actin structures, the overall results were generally unaffected.

Determination of the orientation of the actin outer domain

As actin is composed of two major domains that are composed of a rigid body-like core, actin conformations can be specified by defining the relative configuration of the two cores, that is, the orientation of the OD core relative to the ID core. The orientation was determined using the following procedures (Fig. 1A). First, the ID core of a test structure was superimposed on the reference structure (1J6Z) by translation and rotation of the whole molecule. The structure 1J6Z is a commonly used reference structure of actin that is not complexed with ABPs. Second, the center of mass of the OD core of the superimposed structure was moved onto that of the reference structure by translation. The determined OD shift vector was minute and negligible. Third, the shifted OD core was superimposed on the OD of the reference structure by rotation around the mass center of the shifted OD core. The rotation matrix was calculated using Kabsch's algorithm [17,18]. Finally, three rotation angles (β, α, and γ) were calculated from the rotation matrix. These were defined by sequential rotations around the *Y*, *X*, and *Z* axes of the coordinate system fixed to the reference structure (Fig. 1A).

Actin molecule conformations can be classified into four groups

Hierarchical cluster analysis was performed on the orientations of the OD cores using the grouped

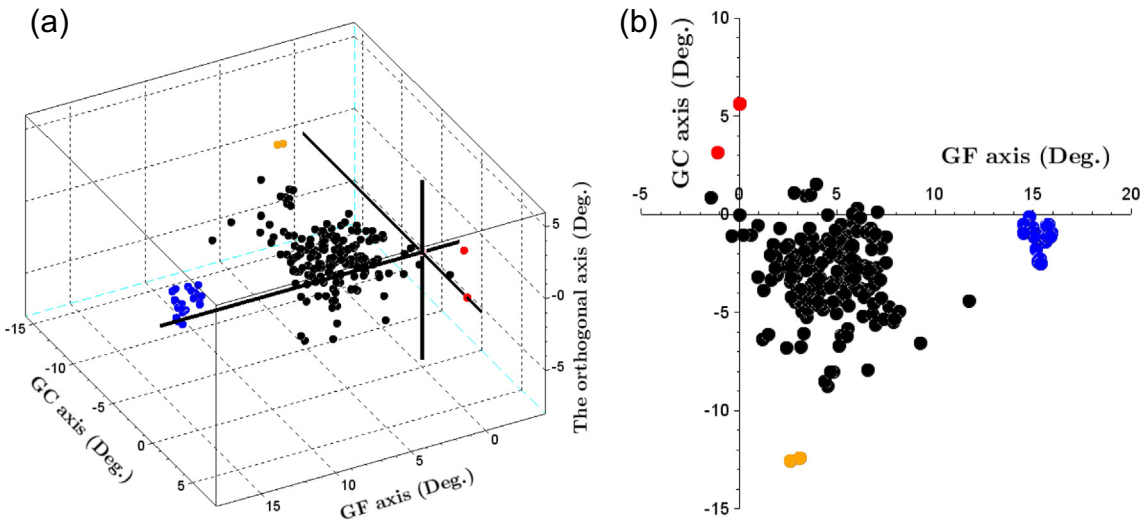

Fig. 4. Outer domain orientations in the GCF space. (A) Orientations of the OD core in the GCF-space specified by the GF, GC and orthogonal axes (black lines). F-form: blue, C-form: red, O-form: orange, and G-form: black. (B) 2D projection of panel A along the orthogonal axis.

average method in the R software package [19,20]. The distance measure used was the square of the Euclidean distance in the αβγ-space. The resulting dendrogram, shown in Fig. 3A [21], was categorized into four clusters using the "Dynamic Tree Cut" package for R to detect clusters on the basis of the dendrogram only [22]. Jaccard bootstrap means were calculated for each of the four clusters using the "cboot" function of R to assess whether the clusters represent a true structure [20,23]. The values for clusters 1–4 were 0.99, 0.84, 0.80, and 0.97, respectively, suggesting that these clusters were valid. In addition, each of the four clusters could be interpreted as one of the four actin forms already described. As shown in Fig. 3B, we observed large gaps separating the four actin clusters in the αβγ-space. Cluster 1 was centered at (α, β, γ) = (12, –10, 0.82) and consisted of F-actin conformations (F-form, which includes the structures 2ZWH, 5JLF, and 5ONV). The set of parameters used means that the actin OD in F-actin is superimposed onto the reference structure via sequential rotations of -10 degrees around the *Y* axis, 12 degrees around the *X* axis, and 0.81 degrees around the *Z* axis. Cluster 2 was located at (α, β, γ) = (9.5, 5.5, –6.8) and was characterized by a wide-open nucleotide-binding cleft, a conformation adopted in profilin–actin complexes (O-form: 1HLU and 3UB5). Cluster 3 was located at (α, β, γ) = (–3.5, –3.3, 2.0) and corresponded to the actin conformation in the cofilactin filament and the twinfilin–monomeric actin complex (C-form: 5YU8 and 3DAW). Cluster 4 was widely distributed in the area enclosed by the F-, C-, and O-forms and contained the majority of the actin conformations reported from crystal structures (G-form).

Interestingly, the direction of the F-form cluster from the origin of the coordinate system (1J6Z) was almost perpendicular to the direction of the C-form cluster. This suggests that the rotation axis of a hinge motion from the G- to the F-form is almost perpendicular to the rotation axis of a hinge motion from the G- to the C-form (Fig. S1). For ease of OD core characterization, the αβγ-space was rotationally transformed to the space specified by a new orthogonal unit vector set (the GCF-space; see supplementary note): the GC axis was directed to 5YU8 (C-form) from 1J6Z (origin), the GF axis was almost directed to the F-form cluster (which was perpendicular to the GC axis), and the corresponding orthogonal axis (Fig. 4A). The GF axis represents the degree of flattening (propeller motion [3]), and the GC axis denotes the degree of the cleft closure (scissors motion [7]). As a result of this transformation, the centers of the F-, C-, O-, and G-form clusters are (15, –1.4, 0), (–0.19, 5.2, 0.26), (2.9, –12, 1.5), and (4.5, –3.1, –0.58), respectively. The O-form cluster has a large negative value along the GC axis, indicating that the nucleotide-binding cleft is open (Fig. 4A and B). Furthermore, the two-dimensional projection with the GC axis and the GF axis (the 2D-GCF plot) appears to be sufficient to characterize each of the actin conformations (Figs. 4B, 5A–D, and S3–S6; the F-, C-, and O-forms are colored blue, red, and orange, respectively).

Bound ABP-dependent subclassification of the G-form cluster

As shown in Fig. 5A and B, OD orientations in the G-form cluster form small sub-groups that are dependent upon the bound ABP. For example, one

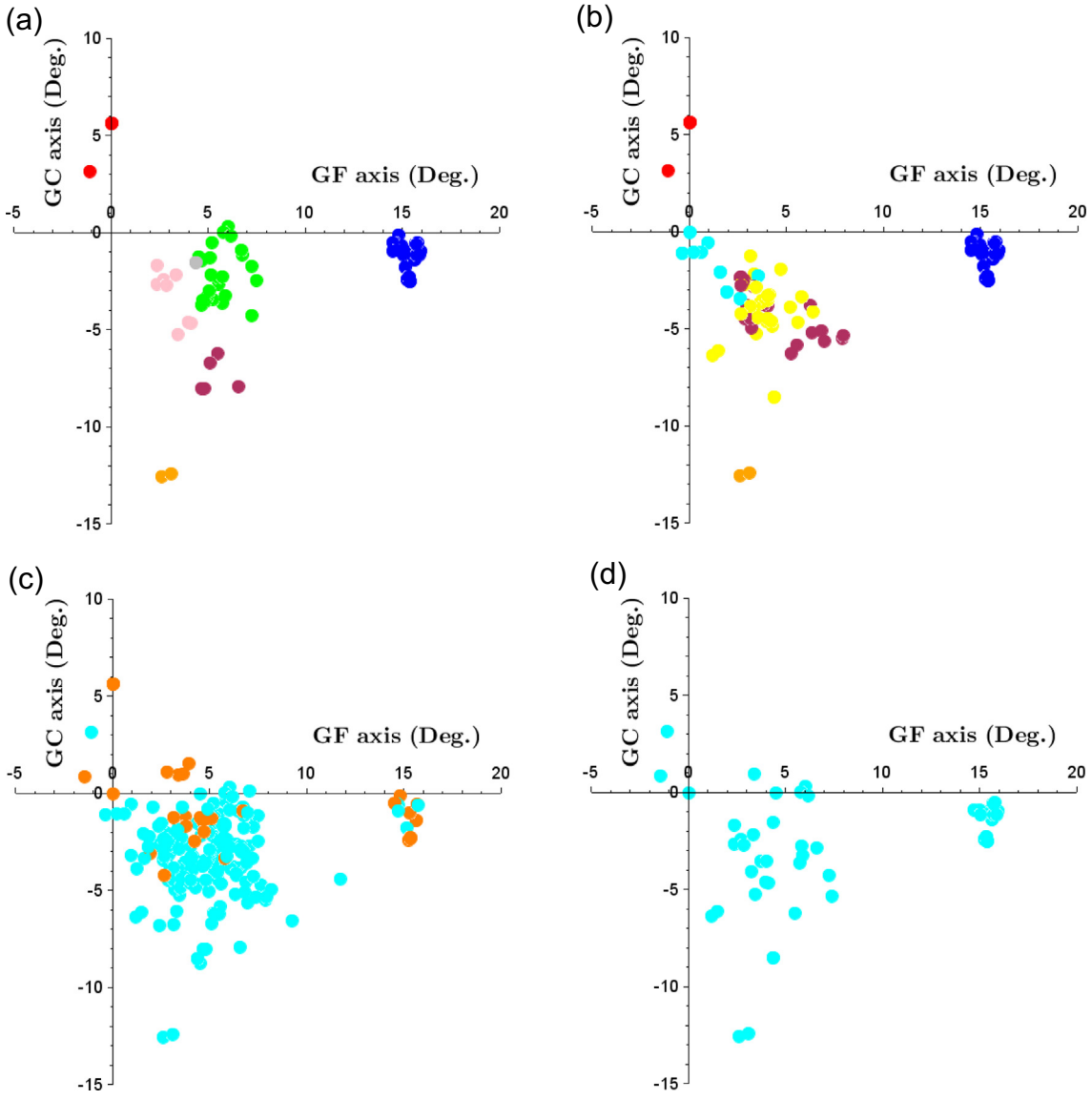

Fig. 5. Outer domain orientation of G-form in the GCF space. The figures are 2D map of panel B in Fig. 4 along the orthogonal axis. F-form: blue, C-form: red, and O-form: orange in (A) and (B). In panel A, actin in complex with gelsolin S1: green; with DNase I: pink; with profilin: maroon; with gelsolin S1 and DNase I in a ternary complex (*3j3c*): grey; In panel B, actin in complex with the WH2 domain: yellow; and with the RPEL domain: maroon. Uncomplexed actin is represented as cyan in panel B. In panel C, actin with bound ADP: orange; and with bound ATP/AMPPNP: cyan. In panel D, actin with a visible D-loop: cyan.

G-form sub-group (shown in maroon in Fig. 5A) contains the profilin–actin complex, which also tends to have an open ATP-binding cleft between the two major domains [11] that is analogous to the O-form cluster. In addition, bound gelsolin S1 (green data points in Fig. 5A) and DNase I (pink in Fig. 5A) clearly form distinct groups from one other, and actin molecules bound to gelsolin S1 are flatter. Intriguingly, the actin–gelsolin S1-DNase I ternary complex adopts the domain orientation of the gelsolin S1 group despite the typical D-loop conformation for DNase I binding (*3cjc*; grey in Fig. 5A). This suggests that gelsolin S1, rather than DNase I, dominantly dictates the actin conformation in the ternary complex. Furthermore, the actin–gelsolin S1 region (green in Fig. 5A) barely overlaps with regions comprising the actin–RPEL domain complex that includes the RPxxxEL actin-binding motif (maroon in Fig. 5B) [24] and actin-WH2 (Wiskott–Aldrich syndrome protein-homology domain 2)-like domain complexes (yellow in Fig. 5B) [25].

Crystal structures of uncomplexed, non-polymerizable actin molecules, either by amino acid replacements (e.g., A204E/P243K) and/or by

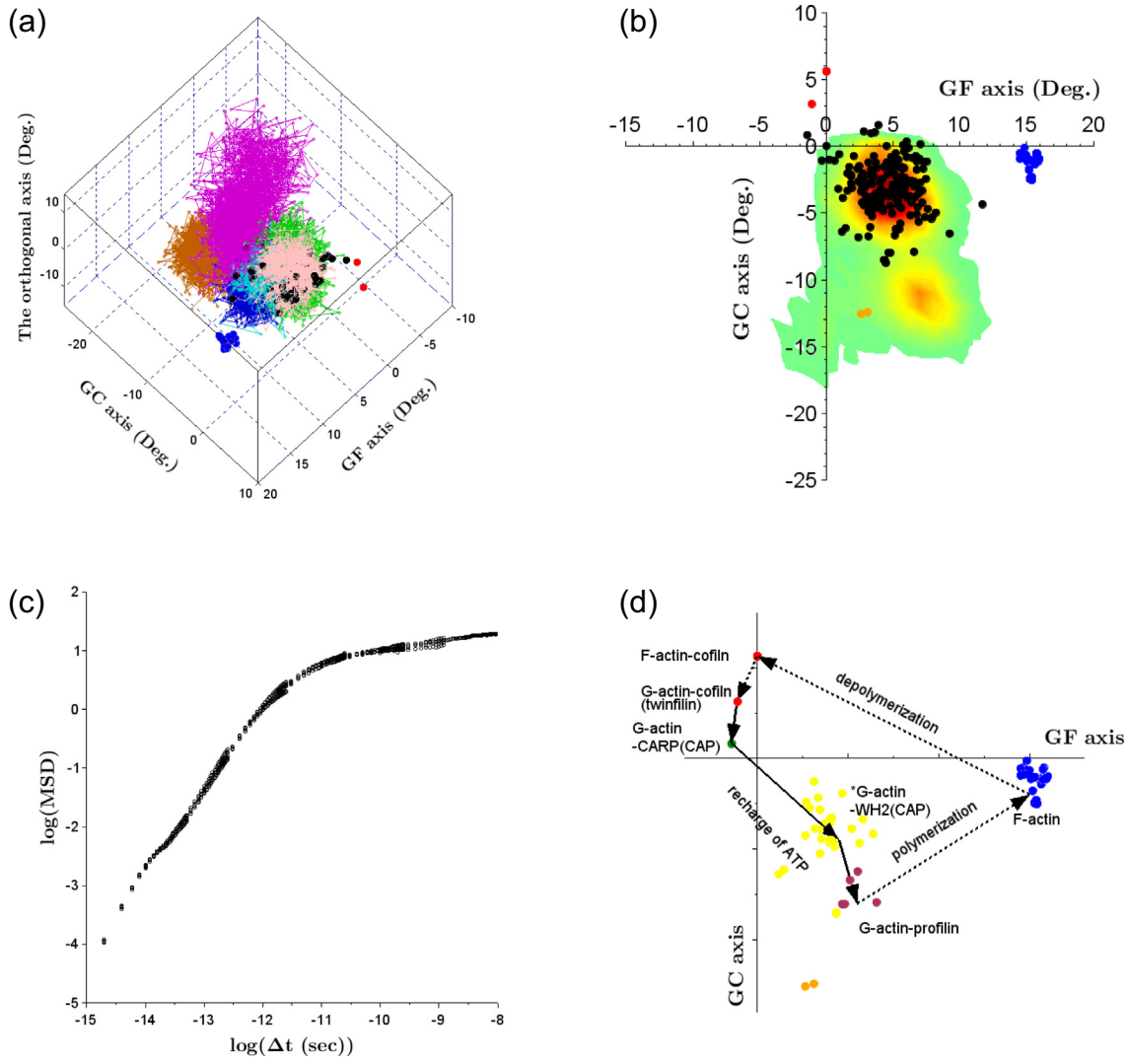

Fig. 6. Characterization of the outer domain fluctuating motion in the αβγ-space (or GCF-space). (A) 500-ns MD trajectories (6 in total) calculated from six G-form actin structures, which have been added to Fig. 4A: 2FXU (blue line and dots), 1WUA (green line and dots), 2PBD (cyan line and dots), 4K41 (pink line and dots), 2GWJ (magenta line and dots), and 4PKG (brown line and dots). Each dot represents the outer domain (OD) orientation sampled every 0.1 ns in the trajectory of 200–500 ns. (B) The two-dimensional (2D-GCF) surface plot for distribution of OD orientation sampled every 0.1 ns in 200–500 ns. The distribution was calculated from the numbers of points present in the grid defined by the GC and GF axes from panel A. The color map progresses from green (1) to red (437) in 32 steps. F-form: blue, C-form: red, O-form: orange, and G-form: black. (C) Logarithmic plot of MSD as a function of time, created from the MD trajectory of chain A of 4K41. (D) Pathway of actin ATPase cycle in the GCF-space. Actin F-form: blue, C-form: red, O-from: orange, actin in complex with profilin: maroon; with the WH2 domain: yellow; and with the CAP-CARP domain: deep-green. Arrows represent the pathway of exchanging ABPs. *G-actin-WH2(CAP) has not been crystalized.

chemical modifications such as ADP-ribosylation and TMR-modification (cyan in Fig. 5B) also appear to form a discrete group. In contrast, the majority of actin structures with bound ADP (orange in Fig. 5C) were distributed in the same area as those with bound ATP or AMPPNP (cyan in Fig. 5C); however, in contrast to bound ADP, bound ATP appears to allow conformations with an open cleft. In addition, actins that possess a D-loop with continuous electron density did not show a unique distribution of OD orientations (cyan in Fig. 5D). These findings indicate that actin conformations are primarily determined via their interaction with other proteins.

Relationship between thermal fluctuations and variations of the G-form cluster

To investigate the relationship between the conformations within the G-form cluster and thermal fluctuations of domain orientation, MD simulations

were performed on six isolated high-resolution crystal structures of actin at 300 K and 1 atm: 18 short runs for a duration of 100 ns and six long runs for a duration of 500 ns. The actin structures selected were chain A of 2FXU (a bistramide A–actin complex; 1.35 Å [26]), chain A of 1WUA (an Aplyronine A–actin complex; 1.45 Å [27]), chain A of 2PBD (a profilin–actin complex with the poly-PRO-GAB domain of VASP; 1.5 Å [28]), chain A of 4K41 (a kabiramide–actin complex; 1.4 Å [29]), chain A of 2GWK (SpvB ADP-ribosylated, uncomplexed actin; 2.0 Å [30]), and chain A of 4PKG (actin complexed with the N-terminal actin-binding domain of tropomodulin; 1.8 Å [31]). These structures were scattered in the G-form cluster region. After the systems were energy minimized [followed by short runs to equilibrate temperature and pressure with harmonic position restraints applied for the heavy atoms of actin (C, N, O, S, and Ca)], OD core orientations were plotted every 0.1 ns from trajectories of 200–500 ns (Fig. 6A). From the six MD trajectories in Fig. 6A, total distribution of the OD orientation was calculated (Fig. 6B), and distribution for each MD trajectory is shown in Fig. S5. To characterize the fluctuating motion of the OD core in the αβγ-space during the MD simulations, we performed a logarithmic plot of mean square displacement (MSD) as a function of time (Fig. 6C). MSD increased linearly with time up to the order of 10 ps, at which point the slope became smaller, suggesting that the displacement has an upper limit. Our results indicated that the OD orientation fluctuates within a confined space when viewed on a 1- to 500-ns timescale.

Conformational fluctuations were observed throughout almost the entire G-form cluster region (Fig. 6A and B). For quantitative analysis, we divided the αβγ-space into sets of 1.5 Degree voxels and counted the number of voxels in which at least one data point resides (Table S1). The entire G-form cluster region (shown in Fig. 6B) was covered by 55 voxels, and MD trajectories (data points sampled every 0.1 ns from one trajectory of 200–500 ns) were covered by 300 voxels on average. The data points selected from the MD runs were widely spread toward large negative values along the GC axis (the open cleft) (Fig. 6B). Voxels picked by six MD simulation long runs partially overlapped each other and the cumulative voxels covered 95% of the entire G-form cluster voxels. This suggests that various conformations within the G-form cluster are accessible via thermal fluctuations.

Accessibility across separated clusters is summarized in Table S1. Overlapping of MD trajectories within the O-form region was observed in 2 of the 6 MD simulation long runs as shown in Fig. 5S (and in 3 of the 18 short runs). In contrast, we did not observe any overlapping within the F- and the C-form regions. These results indicate that, within the 500-ns timescale tested, the O-form is accessible from the G-form by a thermal fluctuation, whereas the F- and the C-forms are not accessible from the G-form under equivalent conditions. However, the G-form region was accessible from the F- and C-forms (Figs. S4 and S6).

Discussion

Proteins can change their overall conformations by shear and hinge motions, which, in recent years, have been determined by the use of computer programs such as DynDom [32,33]. To date, over 200 actin structures have been determined, which exhibit a variety of hinge motions of the two main domains when compared against a representative G-actin reference structure (typically 1J6Z). This approach is excellent for categorizing the characteristics of each structure. However, the relationship between actin conformations and their relevant functions are still not understood. We previously proposed two alternative parameters to summarize each structure: the distance between Cαs of G15 and of D157 as an indicator of cleft closure and the rotational angle of two major domains around the G/F axis as a measure of flattening [13] (Fig. S1). However, these two parameters are not suitable for plotting in a conformational space to allow a collective overview of the available data. Xue *et al.* [34] used principal component analysis to analyze MD-computational and experimental actin structures, and plotted data points for each structure in a two-dimensional space of the first and second principal components. This clearly showed bound ABP-dependent and filament-specific actin conformations; however, a clear separation between the F- and G-forms was not evident. In contrast, our study shows discrete F-and G-form clusters in the αβγ-space of the orientation of the OD core that are separated by a large gap, which may reflect the existence of an energy barrier between the two clusters. Principle component analysis-based analysis of a set that mainly includes G-form actins is likely to be insufficient to accurately describe the F-form. To conclude, we believe that OD core orientation is a suitable parameter to delineate all reported actin conformations.

In this study, we showed that actin conformations can be described by two parameters (flattening and cleft closing) and determined ranges of allowed values for these parameters that represent all the reported actin conformations (Figs. 3B and 4A–B). The allowed space was the conformational space of the actin molecule, which could be used for further bioinformatical analyses such as energy landscape calculations despite some limitations. The conformational space was then divided into G-, C-, and F-form regions. O-form appears to be a satellite cluster of G-form as the O-form is accessible from G-form.

The G-form has a wide conformational space toward the open cleft with likely multiple energy minima; four of the six MD simulation long runs remained in the original G-form cluster area, while the nucleotide-binding cleft in the other two runs partially opened at early stage and retained the state (Fig. S5 and Table S1). The F- and C-forms were not accessible from the G-form via a thermal fluctuation when viewed on a 500-ns timescale (Figs. 6A and B, S4 and S5, and Table S1). This suggests that there is an energy barrier from the G- to F-forms and from the G- to the C-forms that cannot be overcome by a thermal fluctuation in this timescale. However, transition from the G- to the F-form would occur in the polymerization process: upon incorporation into the filament end, a monomer would overcome the energy barrier by interacting with two other actin molecules constituting the filament end. Nucleators such as the Arp2/3 complex mediate a similar effect. Conversely, of the four MD simulation long runs on the isolated F-form of actin, two proceeded to the G-form, while the other two remained unchanged (2 of 18 short runs proceeded to the G-form) (Figs. S4 and S6). However, both the two MD simulation long runs and six of the short runs on the isolated C-form proceeded to the G-form (Figs. S4 and S6). It is feasible that there is an energy barrier from the F- to the G-form and no barrier from the C- to the G-form: the F-form is a semi-stable state, while the C-form is an unstable state induced by cofilin binding.

Each ABP binds to its respective binding site on actin and fixes the two core domains to a particular unique configuration in the conformational space, which allosterically affects its binding affinity to another ABP. When an ABP–actin complex comes into contact with another ABP with a different binding site, two ABPs inducing different configurations would show a negative cooperativity. As shown in Fig. 5A, the configuration of the two cores of actin in complex with DNase I is vastly different from that with gelsolin S1 or profilin, and, indeed, a negative cooperativity of their binding has been observed [35,36]. All cellular processes involving actin are reliant upon the exchange of bound ABPs. Our GCF-space plot is convenient for summarizing bound ABPs in addition to actin conformations. For example, Fig. 6D shows an actin ATPase cycle pathway in the GCF-space: actin exchanges a bound ABP to profilin for polymerization, then to cofilin for depolymerization, further to CAP (cyclase-associated protein) for ATP-recharging, then to profilin again [37]. CAP is a multi-domain protein, the C-terminus domain of which has multiple actin binding domains including the WH2 and CARP domains [38]. The CAP catalyzes the exchange of actin-bound ADP to ATP during change of the actin binding sites from CARP to WH2 domains [37]. The GCF-space plot suggests that exchange of bound ABP from cofilin to the CARP domain as well as from the WH2 domain to profilin occur without actin undergoing a global conformational change. Conversely, an exchange of CARP to the WH2 domain would involve a conformational change; this corresponds to a key process for the exchange of ADP to ATP.

Finally, the F-actin D-loop strengthens inter- and intra-strand interactions and controls F-actin stability [3,6]; indeed, Try53 phosphorylation [11,39] or Mical-mediated oxidation [40] impairs actin polymerization. However, in crystal structures, D-loop conformations appear to be variable depending on the contact partners, and visibility of the D-loop has no correlation with the configuration of the two core domains, or what nucleotide is bound (Fig. 5C and D). Furthermore, the D-loop in naked F-actin is visible, while it is invisible in cofilin-coated F-actin as the subunits are differently arranged. This suggests that the D-loop can adopt multiple conformations that are principally determined by the shape and position of contact partners. However, a recent report has indicated that the D-loop and the C-terminus are responsible for the conformational difference between ATP–F-actin and ADP–F-actin [4]. In the present study, the global conformation (the relative configuration of the two major domains) of isolated actin was comprehensively analyzed by clustering and MD simulations. To determine a factor that dictates multiple D-loop conformations in actin function, it would be necessary to perform in-depth MD simulations of systems including actin complexed with its various binding partners.

Materials and Methods

Software

Software for calculations of actin OD orientations was run on Scilab 6.0 [41]. Hierarchical cluster analyses were performed using RStudio (R 3.4.3; Boston, MA, USA) [19]. MD simulations of protein were performed using GROMACS 2016.4 (GPU version) for Tegara workstation in our laboratory and NEC LX 108Th-4G in the Research Center for Computational Science, Okazaki, and using GROMACS version 5.0.4 (MPI version) for Fujitsu CX4000/2550 in Nagoya University Information and Communications department [42].

Analysis of *B*-factors

Each actin structure was selected in turn and the mean (m) and standard deviation (σ^2) of the Cα *B*-factors (bf) in the non-missing regions (comprising residues 7–31, 70–196, 205–229,

237, 238, 251–324, and 327–364) were calculated. Next, using the mean and the standard deviation, we standardized each Cα *B*-factor contained within the actin structure [(bf – *m*)/σ]. After performing this procedure for the 79 α-actin crystal structures with a minimum resolution of 3.0 Å, we averaged the standardized *B*-factors of the actin structures for each Cα.

MD simulations

All simulations were performed using the GROMACS package [42] using the CHARMM27/CMAP all-atom force field and the TIP3P water model [43]. The missing region of the initial crystal structure was added using the Modeller 9.0 package [44]. At the beginning of each simulation, the protein was solvated in a rectangular box with a minimum distance of 1.0 nm between any protein atom and the edges of the box. Potassium ions and chloride ions (100 mM) were added to compensate the net charge of the system. Electrostatic interactions were calculated using the particle-mesh Ewald (PMA) algorithm [45]. All bond lengths were constrained using the linear constraint solver (LINCS) algorithm [46]. The temperature and pressure were stabilized at 300 K and 1 atm using a v-rescale thermostat [47] and the Parrinello–Rahman barostat [48], respectively. The system was equilibrated under constant volume and temperature (NVT) conditions for 100 ps with position restraints on the protein heavy atoms. Next, the system was repeatedly equilibrated under constant pressure and temperature (NPT) conditions for 100 ps with position restraints on the protein heavy atoms while the force constants for the position restraints were gradually decreased from 1000 to 5 kJ mol^{-1} nm^{-2}. After the equilibrations were performed, the simulation was started. For G-form simulations, three independent simulations were obtained from one structure using different initial velocities generated by random seeds.

CRediT authorship contribution statement

Toshiro Oda: Conceptualization, Data curation, Software, Funding acquisition, Writing, review & editing. **Shuichi Takeda:** Data curation, Writing, review & editing. **Akihiro Narita:** Conceptualization, Data curation, Writing, review & editing. **Yuichiro Maéda:** Conceptualization, Writing, review & editing.

Acknowledgments

We thank Prof. Dr. Ota (Nagoya University) for comments on the manuscript, Dr. Koike for advice on Motion Tree analysis, and Emma Andrew, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript. This work was supported by JSPS (Japan Society for the Promotion of Science, Japan), KAKENHI (Grants-in-Aid for Scientific Research) Grant Number 17 K07373. Some computations were partially performed at the Research Center for Computational Science, Okazaki, Japan. In addition, other computations were partially performed at the Information Technology Center of Nagoya University, Nagoya, Japan.

Declaration of Competing Interest: The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.05.048.

Received 4 February 2019;
Received in revised form 17 May 2019;
Available online 8 June 2019

Keywords:
domain orientation;
rigid bodies;
protein assembly;
clustering;
conformational space

Abbreviations used:
G-actin, globular actin; F-actin, fibrous actin; D-loop, DNase I binding loop; OD, outer domain; ID, inner domain; Cα, α-carbon; MD, molecular dynamics; *B*-factor, temperature factor; ABP, actin binding protein; gelsolin S1, gelsolin sub-fragment 1; RPEL, RPxxxEL motif; WH2, Wiskott–Aldrich syndrome protein-homology domain 2; TMR-actin, tetra-methyl-rhodamine conjugated actin; VASP, vasodilator- stimulated phosphoprotein; SpvB, *Salmonella enterica* virulence-associated protein; MSD, mean square displacement; Arp, actin-related protein; CAP, cyclase-associated protein; CARP, CAP and X-linked retinitis pigmentosa 2 protein.

References

[1] T.D. Pollard, Regulation of actin filament assembly by Arp2/3 complex and formins, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 451–477, https://doi.org/10.1146/annurev.biophys.35.040405.101936.

[2] M.-F. Carlier, D. Pantaloni, Control of actin assembly dynamics in cell motility, J. Biol. Chem. 282 (2007) 23005–23009, https://doi.org/10.1074/jbc.R700020200.

[3] T. Oda, M. Iwasa, T. Aihara, Y. Maéda, A. Narita, The nature of the globular- to fibrous-actin transition, Nature. 457 (2009) 441–445, https://doi.org/10.1038/nature07685.

[4] F. Merino, S. Pospich, J. Funk, T. Wagner, F. Küllmer, H.-D. Arndt, P. Bieling, S. Raunser, Structural transitions of F-actin upon ATP hydrolysis at near-atomic resolution revealed by cryo-EM, Nat. Struct. & Mol. Biol. 25 (2018) 528–537, https://doi.org/10.1038/s41594-018-0074-0.

[5] S.Z. Chou, T.D. Pollard, Mechanism of actin polymerization revealed by cryo-EM structures of actin filaments with three different bound nucleotides, Proc. Natl. Acad. Sci. United States Am. (2019)https://doi.org/10.1073/pnas.1807028115.

[6] J. von der Ecken, M. Müller, W. Lehman, D.J. Manstein, P.A. Penczek, S. Raunser, Structure of the F-actin–tropomyosin complex, Nature. 519 (2015) 114–117, https://doi.org/10.1038/nature14033.

[7] K. Tanaka, S. Takeda, K. Mitsuoka, T. Oda, C. Kimura-Sakiyama, Y. Maéda, A. Narita, Structural basis for cofilin binding and actin filament disassembly, Nat. Commun. 9 (2018), 1860. https://doi.org/10.1038/s41467-018-04290-w.

[8] J.C. Porta, G.E.O. Borgstahl, Structural basis for profilin-mediated actin nucleotide exchange, J. Mol. Biol. 418 (2012) 103–116, https://doi.org/10.1016/j.jmb.2012.02.012.

[9] J.K. Chik, U. Lindberg, C.E. Schutt, The structure of an open state of beta-actin at 2.65 Å resolution, J. Mol. Biol. 263 (1996) 607–623, https://doi.org/10.1006/jmbi.1996.0602.

[10] U.B. Nair, P.B. Joel, Q. Wan, S. Lowey, M.A. Rould, K.M. Trybus, Crystal structures of monomeric actin bound to cytochalasin D, J. Mol. Biol. 384 (2008) 848–864, https://doi.org/10.1016/j.jmb.2008.09.082.

[11] K. Baek, X. Liu, F. Ferron, S. Shu, E.D. Korn, R. Dominguez, Modulation of actin structure and function by phosphorylation of Tyr-53 and profilin binding, Proc. Natl. Acad. Sci. United States Am. 105 (2008) 11748–11753, https://doi.org/10.1073/pnas.0805852105.

[12] J.-F. Gaucher, C. Maugé, D. Didry, B. Guichard, L. Renault, M.-F. Carlier, Interactions of isolated C-terminal fragments of neural Wiskott–Aldrich syndrome protein (N-WASP) with actin and Arp2/3 complex, J. Biol. Chem. 287 (2012) 34646–34659, https://doi.org/10.1074/jbc.M112.394361.

[13] T. Oda, Y. Maéda, Multiple conformations of F-actin, Struct. 18 (2010) 761–767, https://doi.org/10.1016/j.str.2010.05.009.

[14] W. Kabsch, H.G. Mannherz, D. Suck, E.F. Pai, K.C. Holmes, Atomic structure of the actin:DNase I complex, Nature. 347 (1990) 37–44, https://doi.org/10.1038/347037a0.

[15] A. Narita, T. Oda, Y. Maéda, Structural basis for the slow dynamics of the actin filament pointed end, EMBO J. 30 (2011) 1230–1237, https://doi.org/10.1038/emboj.2011.48.

[16] R. Koike, M. Ota, A. Kidera, Hierarchical description and extensive classification of protein structural changes by motion tree, J. Mol. Biol. 426 (2014) 752–762, https://doi.org/10.1016/j.jmb.2013.10.034.

[17] W. Kabsch, A solution for the best rotation to relate two sets of vectors, Acta Cryst. A32 (1976) 922, https://doi.org/10.1107/S0567739476001873.

[18] W. Kabsch, A discussion of the solution for the best rotation to relate two sets of vectors, Acta Cryst. A34 (1978) 827, https://doi.org/10.1107/S0567739478001680.

[19] R Core Team, The R project for statistical computing, 2017.

[20] N. Zumel, Practical Data Science With R, Manning Publications, 2014.

[21] T. Galili, Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering, Bioinforma. 31 (2015) 3718–3720, https://doi.org/10.1093/bioinformatics/btv428.

[22] P. Langfelder, B. Zhang, with contibutions from S. Horath, dynamicTreeCut: Method for Detection of Cluster in Hierarchical Clustering Dendrogram, R package version 1.63-1. 2016.

[23] C. Henning, Fpc: Flexible Procedures for Clustering, R package version 2.1-11.1. 2018.

[24] S. Mouilleron, S. Guettler, C.A. Langer, R. Treisman, N.Q. McDonald, Molecular basis for G-actin binding to RPEL motifs from the serum response factor coactivator MAL, EMBO J. 27 (2008) 3198–3208, https://doi.org/10.1038/emboj.2008.235.

[25] E. Paunola, P.K. Mattila, P. Lappalainen, WH2 domain: a small, versatile adapter for actin monomers, FEBS Lett. 513 (2002) 92–97https://www.ncbi.nlm.nih.gov/pubmed/11911886.

[26] S.A. Rizvi, V. Tereshko, A.A. Kossiakoff, S.A. Kozmin, Structure of bistramide A–actin complex at a 1.35 angstroms resolution, J. Am. Chem. Soc. 128 (2006) 3882–3883, https://doi.org/10.1021/ja058319c.

[27] K. Hirata, S. Muraoka, K. Suenaga, T. Kuroda, K. Kato, H. Tanaka, M. Yamamoto, M. Takata, K. Yamada, H. Kigoshi, Structure basis for antitumor effect of aplyronine a, J. Mol. Biol. 356 (2006) 945–954, https://doi.org/10.1016/j.jmb.2005.12.031.

[28] F. Ferron, G. Rebowski, S.H. Lee, R. Dominguez, Structural basis for the recruitment of profilin–actin complexes during filament elongation by Ena/VASP, EMBO J. 26 (2007) 4597–4606, https://doi.org/10.1038/sj.emboj.7601874.

[29] J.H. Pereira, C. Petchprayoon, A.C. Hoepker, N.W. Moriarty, S.J. Fink, G. Cecere, I. Paterson, P.D. Adams, G. Marriott, Structural and biochemical studies of actin in complex with synthetic macrolide tail analogues, ChemMedChem. 9 (2014) 2286–2293, https://doi.org/10.1002/cmdc.201402150.

[30] S.M. Margarit, W. Davidson, L. Frego, C.E. Stebbins, A steric antagonism of actin polymerization by a salmonella virulence protein, Struct. 14 (2006) 1219–1229, https://doi.org/10.1016/j.str.2006.05.022.

[31] J.N. Rao, Y. Madasu, R. Dominguez, Mechanism of actin filament pointed-end capping by tropomodulin, Sci. 345 (2014) 463–467, https://doi.org/10.1126/science.1256159.

[32] S. Hayward, A. Kitao, H.J. Berendsen, Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme, Proteins. 27 (1997) 425–437https://www.ncbi.nlm.nih.gov/pubmed/9094744.

[33] S. Hayward, H.J. Berendsen, Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme, Proteins. 30 (1998) 144–154https://www.ncbi.nlm.nih.gov/pubmed/9489922.

[34] B. Xue, C. Leyrat, J.M. Grimes, R.C. Robinson, Structural basis of thymosin-β4/profilin exchange leading to actin filament polymerization, Proc. Natl. Acad. Sci. United States Am. 111 (2014) E4596–E4605, https://doi.org/10.1073/pnas.1412271111.

[35] E. Ballweber, E. Hannappel, T. Huff, H.G. Mannherz, Mapping the binding site of thymosin beta4 on actin by competition with G-actin binding proteins indicates negative

co-operativity between binding sites located on opposite subdomains of actin, Biochem. J. 327 (1997) 787–793 , Pt 3 https://www.ncbi.nlm.nih.gov/pubmed/9581557.

[36] E. Ballweber, K. Giehl, E. Hannappel, T. Huff, B.M. Jockusch, H.G. Mannherz, Plant profilin induces actin polymerization from actin : beta-thymosin complexes and competes directly with beta-thymosins and with negative co-operativity with DNase I for binding to actin, FEBS Lett. 425 (1998) 251–255https://www.ncbi.nlm.nih.gov/pubmed/9559659.

[37] T. Kotila, K. Kogan, G. Enkavi, S. Guo, I. Vattulainen, B.L. Goode, P. Lappalainen, Structural basis of actin monomer recharging by cyclase-associated protein, Nat. Commun. 9 (2018), 1892. https://doi.org/10.1038/s41467-018-04231-7.

[38] P.K. Mattila, O. Quintero-Monzon, J. Kugler, J.B. Moseley, S. C. Almo, P. Lappalainen, B.L. Goode, A high-affinity interaction with ADP–actin monomers underlies the mechanism and in vivo function of Srv2/cyclase-associated protein, Mol. Biol. Cell 15 (2004) 5158–5171, https://doi.org/10.1091/mbc.e04-06-0444.

[39] X. Liu, S. Shu, M.-S.S. Hong, B. Yu, E.D. Korn, Mutation of actin Tyr-53 alters the conformations of the DNase I-binding loop and the nucleotide-binding cleft, J. Biol. Chem. 285 (2010) 9729–9739, https://doi.org/10.1074/jbc.M109.073452.

[40] E.E. Grintsevich, P. Ge, M.R. Sawaya, H.G. Yesilyurt, J. R. Terman, Z.H. Zhou, E. Reisler, Catastrophic disassembly of actin filaments via Mical-mediated oxidation, Nat. Commun. 8 (2017), 2183. https://doi.org/10.1038/s41467-017-02357-8.

[41] Scilab Enterprises, Scilab: Free and Open Source software for numerical computation (Windows 10, Version 6.0) [Software] Available from: http://www.scilab.org, www.scilab.org, 2012.

[42] M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, GROMACS User Manual Version 2016.1, www.gromacs.org, 2016.

[43] A.D. Mackerell, M. Feig, C.L. Brooks, Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations, J. Comput. Chem. 25 (2004) 1400–1415, https://doi.org/10.1002/jcc.20065.

[44] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815, https://doi.org/10.1006/jmbi.1993.1626.

[45] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8592, https://doi.org/10.1063/1.470117.

[46] B. Hess, P-LINCS: a parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput. 4 (2008) 116–122, https://doi.org/10.1021/ct700200b.

[47] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007), 014101. https://doi.org/10.1063/1.2408420.

[48] M. Parrinello, A. Rahman, Polymorphic transitions in single crystal: a new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190, https://doi.org/10.1063/1.328693.