博士論文

論文題目　Actin Stability and Dynamics Investigated by Molecular Dynamics Simulations

（アクチン安定性及び動態解析の分子動力学シミュレーション）

氏　名　若井　信彦

Abstract

In this study, I investigated the protein stability and dynamics of globular (G-) and filamentous (F-) actins. Actin is a protein responsible for numerous cellular functions. One of the main components of muscle fiber is actin, which plays many important roles in both muscle and non-muscle cells. I performed molecular dynamics simulations to analyze actin properties and mechanisms.

In the study of protein stability, the pressure tolerance of monomeric actin from the deep-sea fish *Coryphaenoides armatus* and *C. yaquinae* was compared to that of non-deep-sea fish *C. acrolepis*, carp, and rabbit/human/chicken actins using the simulations at 0.1 and 60 MPa. The amino acid sequences of actins are highly conserved across a variety of species. The actins from *C. armatus* and *C. yaquinae* have the specific substitutions Q137K/V54A and Q137K/L67P, respectively, relative to *C. acrolepis*, and are pressure tolerant to depths of at least 6000 m. At high pressure, I observed significant changes in the salt bridge patterns in deep-sea fish G-actins, and these changes are expected to stabilize ATP binding and subdomain arrangement. Salt bridges between ATP and K137, formed in deep-sea fish actins, are expected to stabilize ATP binding even at high pressure. At high pressure, deep-sea fish actins also formed a greater total number of salt bridges than non-deep-sea fish actins owing to the formation of inter-helix/strand and inter-subdomain salt bridges. Free energy analysis suggests that deep-sea fish actins are stabilized to a greater degree by the conformational energy decrease associated with pressure effect.

In the study of protein dynamics, I investigated the conformational transition of actin which is related to the elongation mechanism of F-actin. I proposed the method which combines the parallel cascade selection molecular dynamics (PaCS-MD) and Markov state model (MSM). With these methods, free energy changes upon conformational change were calculated without any bias during MD simulations. The dihedral angle defined by four subdomains in actin (propeller angle) is known to be different between ATP-bound G-actin and ADP-bound F-actin. PaCS-MD enhanced sampling and successfully generated wide range of propeller angles compared to the conventional MD simulations. By the free energy analysis along the propeller angle, ATP-bound actin was confirmed to be stable in a twisted

structure whereas ADP-bound actin was stable in a flat structure. Comparison among G-actin and actin pentamer was conducted to investigate the relation of the propeller angle difference to the ATP hydrolysis and filament elongation. The free energy difference in flat and twisted conformation for ATP-bound actin is larger than that for ADP-bound actin. This observation might imply that conformational transition between flat- and twisted-form could occur in ADP-bound actin more easily than in ATP-bound actin because the conformational transition from twisted- to flat-form would be allowed only when ATP is hydrolyzed into ADP.

Contents

Journal publication

Nobuhiko Wakai, Kazuhiro Takemura, Takami Morita, Akio Kitao. 2014. Mechanism of Deep-sea Fish α-Actin Pressure Tolerance Investigated by Molecular Dynamics Simulations. *PLoS ONE.* 9(1): e85852

Chapter 1

1 General introduction

1-1 Purpose of this thesis

It is well-known that many kinds of proteins work for life. The DNA and amino acid sequences are one of the most important data for proteins and 3D structures also determine their functions. Many biomolecular structures, especially protein, have been determined in molecular level; however, these functions are not fully understood yet. Furthermore, it is difficult to elucidate their mechanisms perfectly. The protein stability and dynamics are associated with the functions such as the motility, robustness, and enzymatic activity. Thus, it is important to clarify these mechanisms. In this study, I chose actin to analyze protein properties because it is partly composed of muscle fiber and has characteristic functions in stability and dynamics. The muscle-related diseases are one of the fatal diseases in human. For example, the muscular dystrophy is the genetic disease and spoils the muscle fibers (1). Additionally, the amino acid sequences of actins are highly conserved and only few amino acid substitutions can affect the actin property and functions. One of the effects of substitutions is actin pressure tolerance. Although pressure is a basic physical quantity, the mechanism of pressure tolerance is still unclear. The purpose of this thesis is to elucidate the stability and dynamics of proteins, especially actin. In protein stability study, I examined the pressure tolerance of deep-sea fish actins compared to the non-deep-sea fish actins. I investigated the mechanism of pressure tolerance induced by two amino acid substitutions in deep-sea fish actins using MD simulations. I performed also MD simulations to investigate the conformational changes between globular (G-) and filamentous (F-) actins.

1-2 Actin

1-2-1 *Actin in muscle*

Actin is a protein responsible for numerous cellular functions. For example, actin serves as a component of muscle fibers in muscle cells, supports cell structure, and is involved in cellular motility as a component of microfilaments even in non-muscle cells. Muscle is composed of distinct filaments such as actin and myosin. In sarcomeres, actin and myosin compose thin and thick filaments, respectively, and their lengths are almost constant except in the process of construction. Some actin binding proteins stabilize the filamentous structure without depolymerization. The muscle contraction process is dominantly conducted by these two kinds of filaments, namely actin and myosin filaments. Actin hydrolyzes ATP into ADP through the elongation of F-actin. In addition to actin hydrolysis, myosin heads move along F-actin as guide rail in sarcomeres using ATP hydrolysis (2). Actin performs its function in various cells and binds many proteins whereas myosin is localized in muscle. Actin is used in muscle and various cells, and one of the most abundant proteins in vertebrates.

1-2-2 *Actin isoforms*

The amino acid sequences of actins are known to be highly conserved across a variety of species. Actins are mainly categorized into three isoforms as α-, β-, and γ-actins. Vertebrate actin can be farther classified into six isoforms: α-skeletal actin, α-vascular actin, α-cardiac actin, β-non-muscle actin, γ-smooth actin, and γ-non-muscle actin (3). These six isoforms have different N-terminal amino acid sequences and they are acetylated during terminal processing. Class II α-actins basically consist of 375 amino acid residues, and class I β- and γ-actins are one amino acid shorter than class II α-actins at N-terminal regions (Table 1). Class I and II have the identical or highly conserved fragments whereas class III has different N-terminal regions and their domains (3). The class III actin was identified as actin-like proteins such as *Saccharoyces cerevisiae* Act2 (3). The difference between class III and class I or II actins was mainly three or eleven amino acid insertions at the residues 42 and 320, respectively (4). In addition to these insertions, the lengths of the

N-terminal regions were also different. Class III actins were most recently classified among the three actin classes. Although the amino acid sequences of Class III actins are highly conserved, interestingly, about 20 amino acid residues of the N-terminal regions show the divergence (5). The N-terminal region completely belongs to the subdomain 1 and this region may be associated with the distinct functions of these actins. The enzyme nitric oxide synthase type 3 produces nitric oxide molecules, which are one of the important paracrine mediators of vascular and platelet functions. The ternary complex among nitric oxide synthase type 3, globular β-actin, and heat shock protein 30 increase the activities of nitric oxide synthase type 3 and cyclic guanosine-3',5'-monophosphate (6). This synthesis is critically regulated by globular β-actin. Therefore, both G- and F-actins play important roles for various functions even in non-muscle cells.

A main component of muscle fiber is α-actin. This α-actin forms stable filaments and connects myosin filaments in sarcomeres whereas unstable actin filaments are polymerized and depolymerized as motor proteins, especially in filopodia and lobopodia (7). In the filopodia and lobopodia, cell movement is caused by thrusting the bacterial membrane with elongation of tangled F-actins. Most of α-actins are related to dynamics and stability of the cell. α-actin belongs to Class II and is classified into three types, namely skeletal, cardiac, and vascular (Table 1). Only nine amino acid residues in the N-terminal region have some variations. This N-terminal difference determines distinct properties of these actins. The N-terminal regions of Class I and II actins contain identical residues in residues 7-9, 11-15, and 18. In Class II actins, residues 6, 10, and 16 are also conserved. Many of the N-terminal regions in Class I and II shares the same charged residues; however, Class III actins are more different in the residues and length of the N-terminal region than others.

Table 1. Amino acid sequence comparison of the actin N-terminal regions (3, 5).

	Residue																	
Type	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Class I																		
β-non-muscle actin		D	D	D	I	A	A	L	V	V	D	N	G	S	G	M	C	K
γ-non-muscle actin		E	E	E	I	A	A	L	V	I	D	N	G	S	G	M	C	K
γ-smooth muscle actin		E	E	E	T	T	A	L	V	C	D	N	G	S	G	L	C	K
Class II																		
α-skeletal actin	D	E	D	E	T	T	A	L	V	C	D	N	G	S	G	L	V	K
α-cardiac actin	D	D	E	E	T	T	A	L	V	C	D	N	G	S	G	L	V	K
α-vascular actin	E	E	E	D	S	T	A	L	V	C	D	N	G	S	G	L	C	K

Shaded residues are identical between Class I and Class II. The residue numbers are based on the Class II actins.

1-2-3 *Actin amino acid sequence*

Class II α-skeletal actin in muscle filament consists of 375 amino acid residues. The amino acid sequence of the α-skeletal actin is highly conserved. Vertebrate actins of rabbit, chicken, and human have the same sequences. Therefore, mutation in actin amino acid sequence is considered to be not permitted in many cases. Related to the high conservation of the amino acid sequence, various actin binding proteins distinguish actins using different interaction surfaces. Actin mutants were rarely purified in experiment even in the case of single mutation, probably due to the fatal effects of the mutations on the cellular functions.

The whole amino acid sequence of the actin is shown in Table 2. Blue residues show identical or conserved fragments of the N-terminal region in actin isoforms. Purple residues represent the DNase I binding loop (residues 42-55). This region is very flexible, and it contacts with the neighboring actin protomers in F-actin, and has high-affinity with DNase I (8). The red residue shows an only modified histidine at residue 73. In all known actins except that of the ameba *Naegleria gruberi*, histidine 73 is known to be post-translationally modified to 3-methylhistidine (3-MeH) (9). This 3-MeH is protonated histidine with the methyl group located at the position 3 in the imidazole ring (Figure 1). pK_a of the imidazole ring in 3-MeH is higher than that in a standard histidine and the imidazole ring in 3-MeH can bind to negative atoms. Because the $N^{\delta 1}$ atom in 3-MeH is suggested to form a hydrogen bond with the carbonyl group of G158 (which forms a hydrogen bond with ATP), the hydrogen bond network of 3-MeH is probably associated with the ATP hydrolysis (9). In experiment, 3-MeH is used for detection of muscular collapse in blood. Orange residues show hinge regions (residues 141-142 and 336-337), which are considered to be an axis of the rotation in domain motions. Green residues show so-called V-stretch region (residues 227-237) located on the inner surface of F-actin. The V-stretch contains a serine rich helix with four continuous serine residues.

Table 2. Whole amino acid sequence of rabbit α-skeletal actin (10).

Residues and amino acids				
1	11	21	31	41
DEDETTALVC	DNGSGLVKAG	FAGDDAPRAV	FPSIVGRPRH	QGVMVGMGQK
Conserved domain				DNase I
51	61	71	81	91
DSTVGDEAQS	KRGILTLKYP	IEHGIITNWD	DMEKIWHHTF	YNELRVAPEE
binding loop		3-MeH		
101	111	121	131	141
HPTLLTEAPL	NPKANREKMT	QIMFETFNVP	AMYVAIQAVL	SLYASGRTTG
				Hinge
151	161	171	181	191
IVLDSGDGVT	HNVPIYEGYA	LPHAIMRLDL	AGRDLTDYLM	KILTERGYSF
201	211	221	231	241
VTTAEREIVR	DIKEKLCYVA	LDFENEMATA	ASSSSLEKSY	ELPDGQVITI
		V-stretch		
251	261	271	281	291
GNERFRCPET	LFQPSFIGME	SAGIHETTYQ	SIMKCDIDIR	KDLYANNVMS
301	311	321	331	341
GGTTMYPGIA	DRMQKEITAL	APSTMKIKII	APPERKYSVW	IGGSILASLS
			Hinge	
351	361	371		
TFQQMWITKQ	EYDEAGPSIV	HRKCF		

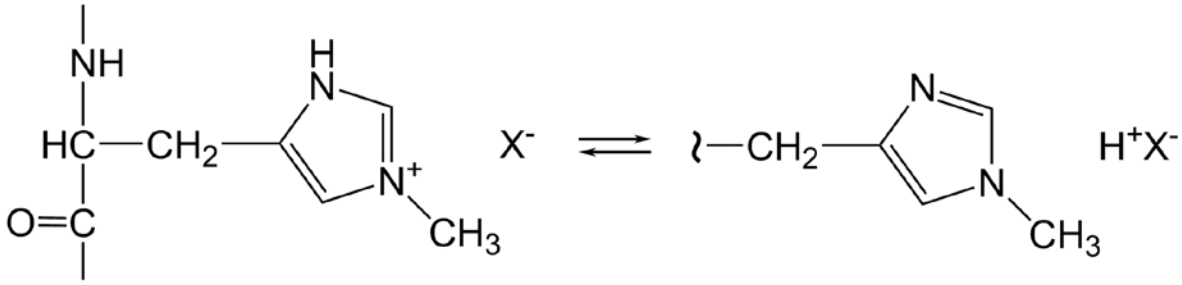

Figure 1. Structure of 3-methylhistidine.

Actin amino acid sequences are highly conserved even between rabbit and yeast (0.867). Among the α-actins of rabbit, yeast, carp, and sea fish *Coryphaenoides*, and the β-actin of fruit fly, the minimum homology rate was 0.859 between yeast and carp (Table 3). Although evolutionary distances between rabbit, yeast, and fruit fly are relatively far, the rates of homology were significantly high. This high rate (over 0.85) is suggested that the high conservation of the amino acid sequence is required for actin functions.

Table 3. Rates of homology in actin amino acid sequences (10-12).

	Yeast	Fruit fly	Carp	Ac1 [a]	Ac2 [a]	Arm [b]	Yaq [b]
Rabbit	0.867	0.928	0.987	0.981	0.984	0.979	0.979
Yeast	-	0.885	0.859	0.861	0.864	0.859	0.859
Fruit fly [c]	-	-	0.925	0.931	0.933	0.928	0.928
Carp	-	-	-	0.995	0.992	0.987	0.992
Ac1	-	-	-	-	0.997	0.992	0.992
Ac2	-	-	-	-	-	0.995	0.995
Arm	-	-	-	-	-	-	0.995

[a]Non-deep-sea and [b]deep-sea fish belong to *Coryphaenoides*. See the detailed definition in Chapter 2-2-1. [c]Fruit fly β-actin.

1-2-4 *Structure of actin*

The first X-ray crystal structure of G-actin was reported in 1990 for rabbit skeletal actin (8) Since then over 70 G-actin structures determined by the X-ray crystallography and solution nuclear magnetic resonance (NMR) have been registered in the Protein Data Bank (PDB). Structure models of F-actin were formerly deduced from the structures of only G-actin; however, a few filament structures were recently determined using cryo electron microscopy (13-16). Most of the actin crystal structures are determined for rabbit actin, and the structures of human, yeast, and fruit fly actins have been also determined. A divalent cation (typically Mg^{2+} or Ca^{2+}) and a nucleotide (ATP or ADP) bind to the center of G-actin at the nucleotide binding site (Figure 2). In general, the divalent cation in the enzymatic pocket is Mg^{2+} in physiological conditions (17). The divalent cation is coordinated with oxygen atoms of the phosphate groups at the tail of the nucleotide and with about four water molecules. The α-actin monomer consists of two major domains separated by the nucleotide binding site at the center of actin. These two major domains can be classified into four subdomains. The domain composed of subdomains 1 and 2 is arranged as the outer surface of F-actin, and another domain consisting of subdomains 3 and 4 forms the core of the F-actin (8) (Table 4). Subdomain 2 consists primarily of loops and is predicted to be very flexible. Most of the X-ray crystal structures do not provide the atomic coordinates of the flexible region.

Table 4. Relationship of α-actin subdomains and residues.

Subdomain	Residue
1	1-32, 70-144, and 338-375
2	33-69
3	145-180, 270-337
4	181-269

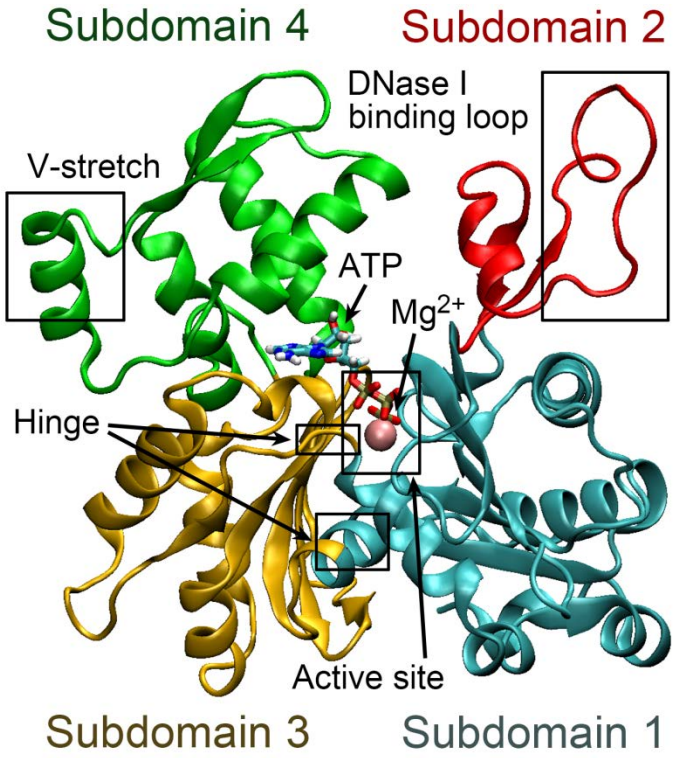

Figure 2. Subdomain arrangement of actin. Subdomains 1, 2, 3, and 4 are shown in cyan, red, yellow, and green, respectively. The pink sphere represents Mg^{2+} at the active site.

1-2-5 *Low-affinity ion binding sites*

Actin has one high-affinity ion binding site at the active site. Some X-ray crystal structures have low-affinity ion binding sites. In PDB, 17 α-actin PDB files have low-affinity ion binding sites (PDB files as of June 8, 2013). Among them, the differences between 2Q0R and 2Q0U and between 2ASM and 2ASP were minor. Low-affinity ion binding sites can be classified into nine positions (S1-9). S3, S4, and S7 are located at the protomer interface whereas the other positions are situated in intra protomer sites (Table 5). Almost all low-affinity ion binding sites are composed of acidic amino acids or backbone oxygen atoms. Therefore, these sites can bind to cations such as Mg^{2+}. The low-affinity ion binding sites would bind between protomers in F-actin and it probably contributes to the F-actin stabilization. Among the 17 PDBs, only 1J6Z has ADP at the active site as well as five

low-affinity ion binding sites. In general, G-actin is stabilized when it binds to ATP. Actin bound to ATP has 1-3 low-affinity ion binding sites. S1 and S6 are the most frequently found binding sites among the nine sites (Table 6).

Table 5. Positions of the low-affinity ion binding sites.

Label	Inter/intra protomer	Residue:Atom name
S1	Intra	V30:O and N
S2	Intra	E125:O, N128:OD1
S3	Inter	G268:O, E270:OE1 and OE2
S4	Inter	E205:OE2
S5	Intra	D222:OD1 and OD2, E224:OE1
S6	Intra	Q263:OE1, S265:OG
S7	Inter	D286:OD1 and OD2, ASP288:OD2
S8	Intra	S323:O
S9	Intra	Q354:O, W356:O, E361:OE1 and OE2

The atom names are corresponded to the PDB file format.

Table 6. Low-affinity ion binding sites in PDB files.

PDB entry	Nucleotide	S1	S2	S3	S4	S5	S6	S7	S8	S9
3M6G (18)	ATP	1			1					1
2VYP (19)	ATP							1		1
2Q1N (20)	ATP						1			
2Q0R/2Q0U [a] (21)	ATP						1	1	1	
2HMP (22)	ATP	1	1							
2FXU (23)	ATP						1	1		
2ASM/2ASP [a] (24)	ATP						1			
2A5X (25)	ATP						1			
1YXQ (26)	ATP				1					
1P8Z (27)	ATP					1				
1NWK (28)	ATP			1						
1MDU (29)	ATP								1	
1LCU (30)	ATP	1								
3U9D (31)	ATP	1								
1J6Z (32)	ADP	1		1	1	1				1
Total		5	1	2	3	2	5	3	2	3

[a] Two PDB files were generated using the same structure except for minor changes.

1-2-6 *Propeller angle rotation*

During elongation of the filament, ATP is hydrolyzed in F-actin and inorganic phosphate (Pi) is released. This activation is a key for the polymerization and dynamics of actins. G-actin basically binds to ATP and a divalent cation as the ligands whereas F-actin mainly binds to ADP and a divalent cation except for the terminal of the filament or minor protomers which bind to ATP. Three or four actins in the end regions are expected to bind to ATP and the actins bound to ATP have different conformation to hydrolyze ATP. Although the ATP hydrolysis and Pi release were observed in experiment, the mechanism of the conformational changes and the ATP hydrolysis have not been fully understood yet. Initial elongation process is suggested to begin when about four actins are aggregated as a

nucleus of filament assisted by actin related protein 2/3 (Arp2/3). Actin can be elongated without the Arp2/3 in the initial process; however, the terminal region is easy to depolymerize due to the unstable conformation. Arp2/3 binds only to one side of the filament and the rate of elongation is different between the two terminal regions.

The rate of the polymerization at a so-called “barbed end” is 5-fold faster than that of the other end, “pointed end”. G-actin takes a U-shape form composed of the four subdomains. The hinge regions connect the two major domains. A preceding study showed the difference of the conformation between G- and F-actin using X-ray crystal structure and cryo electron microscopy structure (16). The two planes defined by two near-axial pairs of these domains are known to be relatively twisted (propeller angle ~20°) in G-actin whereas those in F-actin are relatively flat (~5°) (Figure 3). Therefore, the propeller angle twists when subdomains 2 and 4 contact the barbed end of F-actin. The main difference was seen in these propeller angles. The flat conformation of F-actin is stabilized by the interaction of the neighboring actin protomers whereas the twisted structure is appropriate conformation for G-actin isolated in solution.

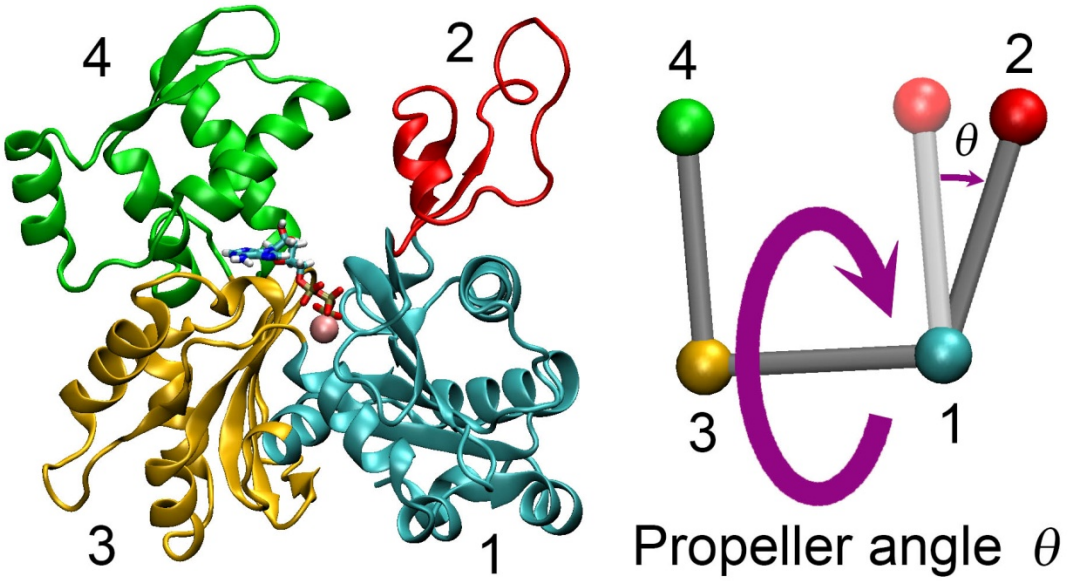

Figure 3. The definition of the propeller angle. Numbers denote the subdomain number. Subdomains 1, 2, 3, and 4 are shown in cyan, red, yellow, and green, respectively. The pink sphere represents Mg^{2+} at the active site.

1-2-7 *Elongation of actin filament*

G-actin polymerizes into F-actin for various cellular functions. The elongation rate controls these functions. F-actin has two different ends due to the asymmetry of G-actin structure. The polymerization rate of the pointed end is slower than that of the barbed end. In general, depolymerization mainly occurs at the pointed end whereas the filament grows at the barbed end. In actin elongation, ATP-bound actin is mainly used as a protomer for the F-actin elongation because ADP-bound actin has a significantly low binding affinity.

The rate constant of the F-actin elongation has been measured by experiment (33). The measurements showed that the dissociation rate constant of the pointed end was about 5-fold larger than that of the barbed end (Table 7). This difference was derived by the association rate constant of ATP-bound actin at the barbed end, i.e., ATP-bound actin has notably high-affinity at the barbed end of F-actin. The minimum association rate constant in

Table 7 was shown for ADP-bound actin at the pointed end, which is only ~1% of ATP-bound actin in the barbed end. In addition to the association rate constant, the dissociation rate constant of ADP-bound actin at the barbed end was largest in the four conditions. The barbed end has high-affinity for ATP-bound actin whereas ADP-bound actin can dissociate more easily. This difference might be related to the conformational changes triggered by the nucleotide change. Actin binding interface changes depending on if G-actin binds to the pointed or barbed end. When G-actin binds to the barbed end, subdomains 2 and 4 of G-actin primarily bind to the barbed end protomers. In contrast, subdomains 1 and 3 of G-actin bind to the pointed end protomers. Subdomain 2 has a flexible loop (DNase I binding loop) and this loop can bind between the two protomers of the barbed end like a bridge. A variety of actin binding proteins regulate the length and stability of F-actin (34). Actin molecule also interacts with many actin-binding proteins, such as profilin and cofilin (35, 36), which accelerates and decelerates the rate of polymerization, respectively. Most of the actin-binding proteins are associated with nucleation, formation, or stabilization of F-actin.

Table 7. F-actin rate constants in solution (33).

	Barbed end		Pointed end	
Constant value	ATP	ADP	ATP	ADP
Association rate constant ($\mu M^{-1}s^{-1}$)	11.6	3.8	1.3	0.16
Dissociation rate constant (s^{-1})	1.4	7.2	0.8	0.27
Dissociation constant (μM)	0.12	1.9	0.62	1.7

Solution condition is 50 mM KCl, 1 mM $MgCl_2$, 1 mM EGTA, pH 7.0.

A preceding study proposed a model of elongation and ATP hydrolysis (14). In this model, conformational changes are considered to be triggered at the end region when G-actin bound to ATP contacts the barbed end of F-actin. The contacted protomer changes its conformation to the flat form. Then, H161 moves and releases the nucleophilic water. The nucleophilic water attacks the γ-phosphate of ATP to hydrolyze to ADP and Pi. It takes about two seconds for this hydrolysis whereas the Pi releases takes about two minutes. This delayed release would be associated with the conformational changes in the terminal region. Pi is released through the back door of the actin. Therefore, ATP is quickly hydrolyzed and Pi is slowly released during the elongation of F-actin in this model.

When the polymerization rate at barbed end equals to that of depolymerization at pointed end, the length of F-actin is maintained. This process is called treadmilling. The rate of treadmilling of pure muscle actin is 0.1 subunits/s or 1.5 μm/h, which is two orders of magnitudes slower than that of F-actin accelerated by actin binding proteins (37). More rapid elongation rate is required for the motility such as cellular movement by the pseudopodia. The rate of elongation to push the membrane forward is 10 μm/min, which is equivalent to the barbed-end elongation rate, 100 subunits/s; however, the treadmilling rate was 200-fold slower when measured *in vitro* using pure F-actin at steady state (37). The elongation speed of the lamellipodium of goldfish epithelial keratocytes is typically ~5-15 μm. The membrane protrusion or bacterium propulsion velocities are 1-25 μm/min (38).

1-3 Molecular dynamics simulations

1-3-1 *Aims of the analysis of molecular dynamics*

Biomolecules such as protein, nucleotide, and lipid have complicated functions and dynamics. These biomolecules dominate biological activities. To evaluate the biological features at molecular level, many biomolecular structures have been investigated using the X-ray crystallography, NMR, and cryo electron microscopy. In general, these methods provide the stable state conformation. Computational methods are recently improved and accelerated by remarkable development of computer hardware and software. One of the oldest computers is the Atanasoff Berry Computer made in Iowa State University around 1939-1942 by Dr. John Vincent Atanasoff and his graduate student, Clifford Berry. The calculation of this computer was conducted by vacuum tubes and its throughput was extremely slow. Nowadays, the processors, central processing unit and graphics processing unit, are highly integrated into a small chip using the transistors. The transmission speeds are amazingly faster among processers, memories, and storage than computers in early days. Additionally, the parallel computation in intra-processer and inter-computers improves the performance. Since the throughput of the modern computers is fast enough to explore many computational problems, many computational methods have been developed. One of the computational methods for molecular science is molecular dynamics (MD) simulation, which deals with not only inorganic substance but also the complex biomolecules. MD simulation is a powerful tool to investigate dynamics and stability of various proteins. However, the computational cost of MD simulations is still relatively high.

In MD simulation, all atom (or atom group) positions are sequentially calculated by integrating the equation of motion. Since MD simulation generates the time evolution of atomic positions, we can observe dynamics of biomolecules and even with surrounding solvent molecules, typically water molecules. From another point of view, we can also treat the trajectories as an ensemble. This ensemble average can be derived as the time average of the physical quantities such as the energy, entropy, temperature, pressure, and volume based on the Ergodic hypothesis.

1-3-2 *The basis of the molecular dynamics simulations*

In general, the MD simulation is performed through the Newton's equation of motions in the Cartesian coordinate space as follows,

$$m_i \frac{d\mathbf{r}_i}{dt^2} = \mathbf{F}(\mathbf{r}_i) = -\nabla U(\mathbf{r}_i) \tag{1}$$

where m, $\mathbf{r}$, t, and U represent the mass of atom, Cartesian coordinate, time, and potential function, respectively. Subscript i represents the i-th atom. To solve this equation, the potential energy function has been developed for biomolecular systems, such as CHARMM (39-41), AMBER (42-49), and OPLS/AA (50) force fields. The CHARMM potential energy function which I employed for the research is defined as follows (51),

$$\begin{aligned} U(\mathbf{r}_i) = & \sum_{bond} k_{bond}(r_{ij} - r_0)^2 \\ & + \sum_{angle} \left\{ k_{angle}(\theta - \theta_0)^2 + k_{Urey-Bradley}(r_{ik} - r_{ub})^2 \right\} \\ & + \sum_{dihedral} \begin{cases} k_{dihedral}(1 + \cos(n\phi + \phi_0)) \; if \; n > 0 \\ k_{dihedral}(\phi - \phi_0)^2 \; if \; n = 0 \end{cases} \\ & + \sum_{Lennard-Jones} (-E_{\min}) \left\{ \left(\frac{R_{\min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{\min}}{r_{ij}} \right)^{6} \right\} \\ & + \sum_{electrostatic} \varepsilon_{14} \frac{C q_i q_j}{\varepsilon_0 r_{ij}} \end{aligned} \tag{2}$$

where k, θ, ϕ, ϕ_0, and n represent the force constant, bond angle, dihedral angle, phase shift angle, and periodicity of the dihedral angle, respectively. Subscript 0 in r_0 and θ_0 represents the equilibrium values. r_{ij} and r_{ik}, represent the distance between i- and j-th and between i- and k-th atoms, respectively. The second term in the angle potential is the Urey-Bradley term. This energy term is calculated between i- and k-th atoms (atom numbers are 1, 2, … , i, j, k, …). The first three energy terms in equation (2) are related to the covalent bonds formed by consecutive two to four atoms. These energies are defined using the harmonic potentials except for the dihedral angle term in $n > 0$. The forth energy term represents the nonbonded energy associated with the repulsive and London dispersion interactions between a pair of atoms. This term is called Lennard-Jones potential. $-E_{\min}$ and $R_{\min}$ represent the minimum of the potential term, namely the well-depth and the distance of the minimum energy, respectively. The last energy term represents the electrostatic potential

between *i*- and *j*-th atoms. ε_0, ε_{14}, *C*, and *q* represent the dielectric constant, scaling factor, Coulomb's constant, and charge of the atom. The force constants in equation (2) are defined using data generated by experiments and quantum mechanics.

If three or more atoms exist in the system, it is impossible to solve the equation of motion analytically. The numerical integration is used for the MD simulations. Several numerical integration methods have been developed such as Verlet method (52), velocity-Verlet method (53), reversible reference system propagator algorithms (54), and predictor-corrector method (55, 56). In the velocity-Verlet method which I employed of the MD simulation, the following two equations are solved as,

$$\begin{aligned} \mathbf{v}_i(t) &= \mathbf{v}_i(t-\Delta t) + \frac{\mathbf{F}_i(t)+\mathbf{F}_i(t-\Delta t)}{2m_i}\Delta t + O(\Delta t^3) \\ \mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t^2 + O(\Delta t^4) \end{aligned} \tag{3}$$

where **v** and Δt represent the velocity and time step, respectively. This method sequentially calculates the atom positions and velocities. For the numerical integration, the initial atom positions and velocities are required to start the MD simulations. In many cases, the Cartesian coordinate such as PDB files and the Maxwell distribution are used for the initial atom coordinates and velocities, respectively. The time step in protein systems is typically used in 0.5-2.0 fs.

In the MD simulations, the simulated system maintains some physical values to be constant to reproduce a target ensemble. For example, the temperature and pressure of the system are kept constant for the isothermal-isobaric ensemble. To maintain the temperature, several thermostat methods have been developed, for example, Langevin dynamics (57), Nosé-Hoover thermostat (58), Nosé-Hoover chain thermostat (59), Berendsen thermostat (60), and Gaussian isokinetic method (61). In addition to the thermostat, the barostat methods are used to maintain the pressure, for example, Langevin piston (62) and Parrinello-Rahman method (63). Additionally, Anderson method can control both temperature and pressure in the MD simulations (64). The periodic boundary condition is employed to avoid the effect of the boundary of the system. In this case, the particle mesh Ewald method procedure (65) is one of the typical methods to reduce the calculation cost of the electrostatic interactions. Additionally, the hydrogen bonds are kept rigid to reduce the time step because hydrogen atoms are involved with the fastest fluctuation in the system. A representative method to achieve this is the SHAKE algorithm (66). Since biomolecules

are typically immersed in solution, many water molecules are included in the simulated system. The SETTLE algorithm (67) is an efficient method particularly designed for water molecule to reduce the computational cost through the rigid of the internal geometry of water.

1-3-3 *Applications*

The first MD simulation in the biomolecules was performed for bovine pancreatic trypsin inhibitor *in vacuo* in 1977 for 9.2 ps (68). Recently, an MD simulation of a protein in solution (17,758 atoms) was also conducted over 1 ms in 2010 (69). Since this simulation was using the special-purpose computers for the MD simulations, the supercomputers for general use can perform shorter simulation time. At the moment, the simulation in ~1 ms is thought to be the maximum limit of long time MD simulation in biomolecules.

The MD simulations have been used to examine the G- and F-actin structures in preceding studies. The results of G-actin and actin-related protein 3 simulations showed that the nucleotide-mediated conformational changes were associated with the nucleotide cleft open/close motions (70). The nucleotide binding loop, P1 loop, which interacted with the end of phosphate tail in ATP, was also examined in this work. A coarse-grained model of F-actin model was employed to investigate the polymerization and depolymerization process and treadmilling (71, 72). The elongation process was simulated using many actins bound to ATP or ADP; however, individual actin structures ware ignored for simplicity. In contrast, all-atom MD simulations of F-actin results showed the difference of the number of hydrogen bonds in F- and G-actin models (73). These F-actin models were conducted with the periodic boundary condition and the length of the filament was effectively infinity. Since the F-actin systems did not have the end of filaments, the terminal regions of F-actin were not observed in this model.

1-3-4 *Molecular dynamics software packages*

The MD simulations performed in this thesis were performed with NAMD program package, which is one of the major programs for the MD simulations (51). The NAMD can

be efficiently carried out in parallel computers, and it can treat various biomolecules and biosystems. The temperature and pressure of the simulated system can be also controlled. The constraint or fix methods are employed if needed. Other major program packages are, for example, AMBER (74) and GROMACS (75). These program packages are similar to the NAMD program package. They are also used for various biomolecules and biosystems.

Chapter 2

2 Mechanism of deep-sea fish α-actin pressure tolerance

2-1 Introduction

In this chapter, I examined the stability of proteins using MD simulations. In particular, I analyzed the pressure tolerance in deep-sea fish actins. Deep-sea fish have the adaption to live in the abyss. Deep-sea fish can be found at depths down to ~6000 m, where the pressure reaches ~60 MPa. *Coryphaenoides*, a marine fish known as rattail or grenadier, inhabits a wide depth range in the Pacific Ocean, making it a suitable species for studying pressure tolerance. The *Coryphaenoides* species *C. acrolepis*, *C. armatus*, and *C. yaquinae* inhabit depths of approximately 180-2000, 2700-5000, and 4000-6400 m, respectively. *C. armatus*, and *C. yaquinae* are deep-sea fish. This depth of ~6000 m is perfectly dark and the temperature is low in 1-4 ºC (76). The environment is harsh to survive for many kinds of marine species. Deep-sea fish is a relatively large creature among living things in the abyss. Since *Coryphaenoides* species inhabit wide depths, deep-sea fish is suggested to obtain the adaptation through the course of evolution. I focused on α-actins in the one of the muscle components. The amino acid sequences of α-actins from these three species are known to be highly conserved, and actins of terrestrial animals or shallow-water fish species also have similar sequences (12, 77). The rates of actin polymerization and, ATP and Ca^{2+} dissociation in non-deep-sea fish are significantly affected by pressures but these rates in deep-sea fish are unaffected by pressures at least up to 60 MPa, even though the actin sequences differ by only a few residues from those of other species (Table 8), especially two residue substitutions between deep-sea and non-deep-sea fish actins. α-skeletal actin is composed of the 375 amino acid residues and the N-terminal residue is generally modified by the acetylation (Table 2).

Table 8. Sequence features of the various actins examined in this study (12).

Species and type	MD label	Residue									
		2	3	54	67	137	155	165	278	299	358
Rabbit / Chicken	Rab	E	D	V	L	Q	S	I	T	M	T
C. acrolepis actin 1 [a]	Ac1W, Ac1Q [c]	D	E	V	L	Q	A	V	A	L	S
C. acrolepis actin 2a [a]	Ac2	D	E	V	L	Q	S	V	A	L	S
C. armatus actin 2b [b]	Arm	D	E	**A**	L	**K**	S	V	A	L	S
C. yaquinae actin 2b [b]	Yaq	D	E	V	**P**	**K**	S	V	A	L	S
Carp	-	D	D	V	L	Q	A	V	A	L	T

[a]Non-deep-sea fish actin. [b]Deep-sea fish actin (12). [c]See methods section in chapter 2 for the differences. Unlisted amino acid residues are identical.

Table 8 shows that several amino acid substitutions have occurred among species. Deep-sea fish actins have a lysine as residue 137 (K137) near the active site in the center of actin (Figure 4), where other species have a glutamine (Q137). Residue Q137 is expected to affect the hydrolysis reaction and is believed to be one of the key residues in the polymerization process based on the evidence suggesting that it plays an important role in controlling water molecules that behave as nucleophiles and attack the γ-phosphate of ATP (78). During hydrolysis process of ATP, it is essential to cleavage the covalent bond using nucleophiles such as water molecules. The attack on ATP by water molecules can significantly impact the rate of polymerization (79) and it starts the hydrolysis of ATP. In addition to the Q137K substitution, deep-sea fish actins have either L67P or V54A substitutions, both of which are distant from the active site and are located on the near protein surface bound to neighboring F-actin protomers (Figure 4), suggesting that they affect the pressure tolerance of F-actin polymerization. These two substitutions, residues 54 and 67, locate near the neighboring protomers but they do not contact directly in F-actin (16). Therefore, specific Q137K is predicted to play an essential role in the pressure tolerance of deep-sea fish α-actin because the position is near the active site of hydrolysis and the two deep-sea fish commonly have the residue.

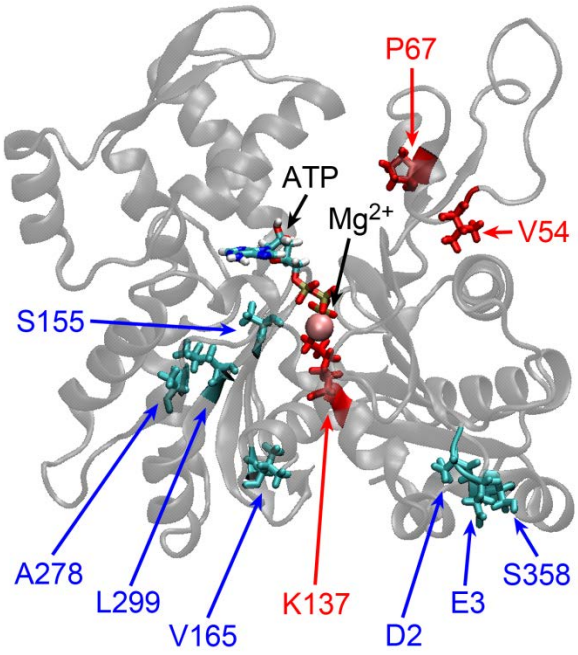

Figure 4. Positions of substituted residues in *C. yaquinae* actin as compared to rabbit/chicken actin. The residues shown in red and cyan in the licorice model represent the specific substitutions in deep-sea fish actins and those of terrestrial animals and shallow-water fish species, respectively.

Protein denaturation, conformational changes, and loss of enzymatic activity are observed in experiments at high pressure, (12, 80-82). The ligand dissociation rates of hydrolases and dehydrogenases were shown to increase at high pressure in experiments (12, 83, 84). Fluctuation of hen egg white lysozyme in picosecond time range was also shown to be affected by pressure and temperature (85). The pressure effect on actin was first reported in 1966 (81). This preceding study showed that pressure-induced denaturation of rabbit G-actins begins to occur at 250 MPa and completes at 400 MPa (81). Mariana Trench is well-known as the deepest part of the ocean in the world. Although the depth is ~11,000 m and the pressure reaches ~110 MPa, the pressure of the actin denaturation is larger than the pressure at the Mariana Trench. In addition, pressure prevents G-actin from assembling due to denaturation or conformational changes (12). High pressure has been shown to induce significant changes in actins purified from

terrestrial animals or shallow-water fish species, as evidenced by decreased actin volume change upon polymerization, a decrease in DNase I inhibition, an increase in the critical concentration, and increases in ligand dissociation rates (12). Therefore, high pressure affects proteins in many properties and the most effects are negative for cellular functions such as polymerization.

MD simulation is one of powerful tools for analyzing the biomolecules such as proteins and also investigating the effects of pressure. Several studies have used MD simulation to examine conformational changes, denaturation, water penetration, and volume changes in proteins under a wide range of pressures up to tens of thousands MPa (86-89). In the case of ubiquitin, water penetration is induced at ~300 MPa, and at higher pressures denaturation was clearly observed (86). Collapse of the secondary structure and an increase in the radius of gyration at high pressure were studied using the water-insertion method, which adds the water molecule at the internal of the secondary structures (89). Changes in protein structure induced by high pressure can also be investigated by calculating the solvent-accessible surface area (SASA) and volume through both experiments and simulations. NMR analyses and volume calculations based on atomic coordinates such as PDB files showed that high pressure compresses protein volume by only 1-3% (89). Since most globular proteins form highly packed structures in the native state, the volume change of this magnitude is relatively small. Most high-pressure simulations involve pressures well over 100 MPa, which is sufficient to induce denaturation of many proteins even highly packed proteins; however, relatively few studies have addressed the effects of pressures below 100 MPa as in the abyss on protein structure.

In this chapter, the effects of pressure up to 60 MPa on the structure and dynamics of G-actin from two deep-sea fish (*C. armatus* and *C. yaquinae* actin 2b), two actins from non-deep-sea fish (*C. acrolepis* actin 1 and actin 2a), and rabbit/chicken actins were investigated using MD simulations at 0.1 and 60 MPa except for the initial equilibrium simulations. The results of free energy analysis show that deep-sea fish actins at high pressure are stabilized in the conformational energy decrease of actin without significant change of the solute entropy. I also show that only two amino acid differences are sufficient to induce significant changes in the pattern of salt bridging, which is suggested to play a significant role in stabilization of ATP binding and subdomain arrangement at high pressure. Possible differences in ATP hydrolysis mechanisms, especially the substituted residue and nucleophilic water molecule, are also discussed.

2-2 Methods

2-2-1 *Structure modeling of actins*

To analyze the effect of amino acid substitutions *in vivo*, it is necessary to obtain actin mutants; however, it is well-known that most mutant actins produced by mutagenesis cannot be expressed (79). Therefore, most of experiments in the previous studies were carried out using actins purified from muscle fiber. Instead of determining actin mutant structures, I conducted molecular modeling of the various actins based on the rabbit one from X-ray crystal structures *in silico*. Because the amino acid sequence of actin is highly conserved and the difference was small, this modeling was relatively straightforward. The actin molecules studied in this chapter are listed in Table 8. First, the atomic coordinates of rabbit skeletal muscle α-actin, including crystal water molecules, ATP, and Ca^{2+} were adopted from a high-resolution crystal structure (PDB entry 1WUA, resolution 1.45 Å (10)). Since residues 1-4 of the N-terminus, residues 42-50 of the flexible DNase I binding loop, and residues 372-375 of the C-terminus were missing in this structure, lower resolution structures from PDB entries 1ATN (resolution 2.80 Å) (8) and 1ESV (resolution 2.00 Å) (90) were used for residues 1-4, 39-53, and 372-375, respectively. Hydrogen-atom coordinates were complemented using the visual molecular dynamics (VMD) software package (91). The modeled rabbit actin coordinates were also used as the template to model other actin structures. Actin 2b of the deep-sea fish species *Coryphaenoides armatus* and *C. yaquinae* are referred to as Arm and Yaq, respectively. Actin 1 and 2a of the non-deep-sea fish species *C. acrolepis* are labelled as Ac1 and Ac2, respectively. *C. acrolepis* actins contain 60% Ac2 and 40% Ac1. In contrast, actins of *C. armatus* consist of 20% Ac2 and 80% Arm, and those of *C. yaquinae* consist of 19% Ac2 and 81% Yaq (12). Additionally, the sequence of rabbit actin (labelled as Rab) is identical to those of human and chicken actins.

Except for the Q137K substitution, the other fish actin substitutions relative to the rabbit/chicken actin sequence are located on the protein surface distant from the active site and distant from each other, making the modeling of these side chains straightforward. Because Q137K is located in the active site and is expected to be important for pressure tolerance, I was careful in my modeling approach. All 81 rotamers (3^4 rotations) were

considered as possible side-chain structures of K137. After energy minimization, I compared the structures and chose the structure for which the Ca^{2+}–N^{ζ} distance was the longest (5.55 Å). The Ca^{2+} in the enzymatic pocket bound to a nucleotide was exchanged with Mg^{2+} to reproduce physiological conditions (17). The N-terminus was modified with acetyl-aspartate, and 3-MeH was adopted for the residue 73 (Figure 1). All five G-actin models examined (i.e., Yaq, Arm, Ac1, Ac2, and Rab) were solvated in a periodic boundary box with 50 mM KCl and a solvation water layer of at least 10 Å.

After a 3000-step energy minimization using the conjugate gradient method, MD simulations were first performed by keeping non-modeled protein atoms, K137, ATP, and Mg^{2+} fixed at the initial position for 1 ns and then restraining the same atoms with harmonic positional restraints for 1 ns. The force constraints were started from 1 kcal/mol/$Å^2$ and gradually decreased by 0.1 kcal/mol/$Å^2$ every 0.1 ns. After the restrained MD simulation, I began 32 MD simulations for both the Ac1 and Rab systems with distinct initial velocities and carried out the simulations independently at 0.1 MPa for 0.1 ns without restraints. I obtained two distinct Mg^{2+} coordination patterns in Ac1: 1) coordination by four water molecules and two γ-oxygen atoms of ATP in 26 cases (Ac1W), and 2) coordination by three water molecules, two γ-oxygen atoms of ATP, and a Q137 side-chain oxygen atom in six cases (Ac1Q) (Figure 5). I examined cation coordination patterns for 73 globular α-actin structures in the PDB and found that 42 X-ray crystal structures had coordination patterns similar to that of Ac1W, whereas Ac1Q-type coordination was formed only in six relatively low-resolution structures with the following PDB ID resolutions: 1ATN (8), 2.80 Å; 1IJJ (30), 2.85 Å; 1LCU (30), 3.50 Å; 1H1V (92), 3.00 Å; 1RFQ (93), 3.00 Å; 1Y64 (94), 3.05 Å. The initial side-chain dihedral angles of residue 137, which had dihedral angles χ_1 to χ_3 from the C^{α} to the C^{δ} of the end of the side chain in Ac1W, were $\chi_1 = -61.7°$, $\chi_2 = -176.5°$, and $\chi_3 = 177.3°$. The dihedral angles for Ac1Q were $\chi_1 = -58.1°$, $\chi_2 = 166.9°$, and $\chi_3 = -168.8°$. Therefore, the initial difference between Ac1W and Ac1Q was primarily the χ_2 and χ_3 dihedral angles of residue 137. For each pattern, one representative structure was selected and investigated with longer MD simulations (Table 8). In all 32 Rab simulations, Q137 was not coordinated to Mg^{2+}.

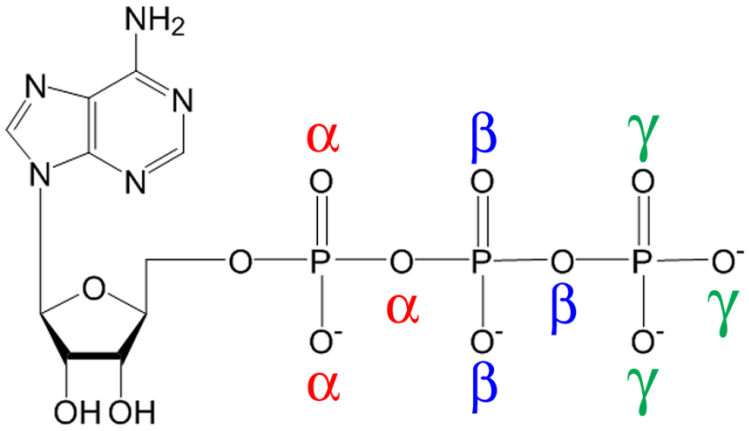

Figure 5. Chemical formula of ATP. Oxygen atoms in the phosphate tail of ATP are distinguished by α, β, and γ.

2-2-2 *Molecular dynamics simulations*

A total of six different models (i.e., Ac1W, Ac1Q, Ac2, Rab, Arm, and Yaq) were simulated using MD at pressures of 0.1 and 60 MPa. The high-pressure MD simulation was conducted after gradually raising the pressure from 0.1 to 0.2 MPa in 30 ps and then to 60 MPa in 0.2-MPa increments per 0.03 ns. MD simulations were performed with the CHARMM22 force field (39-41) and SPC/E water model (95) using the NAMD software package (51). The parameter files were modified to accommodate 3-MeH (96). SPC/E model was employed because its translational diffusion constant and the rotational correlation time are the closest to the experiment values among TIP3P, TIP4P, SPC, and SPC/E (97). The 3-MeH parameters and topology were generated in the CHARMM22 files using a doubly protonated histidine and an N-methylamide C-terminus patch (98). Electrostatic potentials were calculated using the particle mesh Ewald procedure (65), and van der Waals interactions were computed using a 12-Å cutoff and a smooth switching function. The simulations were conducted with periodic boundary conditions in an isobaric-isothermal ensemble, with the exception of the initial 2-ns simulations, in which a canonical ensemble (constant *NVT*) was employed. Constant temperature was maintained using Langevin dynamics (57) for non-hydrogen atoms with a damping coefficient of 5 ps^{-1}, whereas constant pressure was maintained using a Langevin piston (62) with an oscillation

period of 100 fs and a decay time of 50 fs. The bond between each hydrogen and the atom to which it is bonded in the solute is constrained using the SHAKE algorithm (66), and the internal geometry of water molecules was kept rigid using the SETTLE algorithm (67). To reproduce experimental conditions, MD simulations were carried out at 277 K and either 0.1 or 60 MPa. This temperature was selected to reproduce the condition in the experiment (12). Deep-sea temperatures are 1-4 ºC (76). The effects of possible difference between simulation and real temperatures (< 3 ºC) are expected to be very small. The MD time step of 2 fs was used for the simulations, which were performed for 100 ns (50-ns equilibration and 50-ns sampling). The coordinates and energy data were stored every 0.5 and 0.1 ps, respectively.

2-2-3 *Analysis of physical properties*

To examine the effects of high pressure on G-actin, I calculated the excluded volume (V_{ex}), SASA, and the isothermal compressibility (κ_T). Both V_{ex} and SASA were calculated using the CAVE software package (99). SASA is defined by the track of the probe center as the probe rolls around the whole surface of the protein, and the space inside the track of the probe is defined as V_{ex}. The probe radius was 1.4 Å, and the van der Waals radius was used for each protein atom. These van der Waals radii were 2.0 Å for the sp^3 carbon, sp^3 nitrogen, and sulfur with a hydrogen, 1.7 Å for sp^2 carbon without hydrogen and sp^2 nitrogen with a hydrogen, 1.85 Å for sp^2 carbon with hydrogen(s) and sulfur without hydrogen, 1.8 Å for sp^2 nitrogen with hydrogens and 1.4 Å for oxygen (100).

Protein compressibility is a property that is associated with the structure of a protein. The packing density of a protein is non-uniform due to the presence of small cavities. Because compressibility is significantly affected by internal cavities and hydration, it is an effective measure of protein structure property. Compressibility is also related to toughness and fluctuation. The adiabatic compressibility of various proteins has been measured experimentally using the velocity of sound in solution (101), and from these measurements the isothermal compressibility can be estimated. I calculated the isothermal compressibility, κ_T, which is defined as follows,

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = \frac{\left\langle V^2 \right\rangle - \left\langle V \right\rangle^2}{k_B T \left\langle V \right\rangle} \tag{4}$$

where V, p, k_B, and T represent the system volume, pressure, the Boltzmann constant, and the absolute temperature, respectively. The angle bracket denotes the average over last 50 ns of the simulation.

To understand the stability of actin at high pressure, free energy shift ΔG caused by pressure change from 0.1 to 60 MPa was estimated using the following thermodynamic cycle,

$$\begin{array}{ccc} \text{State N } \textit{in vacuo} & \xrightarrow{\Delta\mu_{0.1\text{MPa}}} & \text{State N in solution (0.1 MPa)} \\ \Delta E_{\text{conf}} - T\Delta S \quad \downarrow & & \downarrow \Delta G \\ \text{State H } \textit{in vacuo} & \xrightarrow{\Delta\mu_{60\text{MPa}}} & \text{State H in solution (60 MPa)} \end{array} \tag{5}$$

where states N and H represent stable states of actin at normal (0.1 MPa) and high (60 MPa) pressure, respectively. In this conceptual cycle, we consider that solvation steps do not alter actin structures in states N and H. $\Delta\mu_{0.1\text{MPa}}$ and $\Delta\mu_{60\text{MPa}}$ indicate solvation free energies, i.e., transfer free energies from vacuum to solution at 0.1 and 60 MPa, respectively. ΔE_{conf} and ΔS are change of conformational energy and solute entropy from states N to H. From the cycle shown by (2), ΔG can be calculated as,

$$\begin{aligned} \Delta G &= \Delta E_{\text{conf}} - T\Delta S + \left(\Delta\mu_{60\text{MPa}} - \Delta\mu_{0.1\text{MPa}}\right) \\ &= \Delta E_{\text{conf}} - T\Delta S + \Delta\Delta\mu \end{aligned}. \tag{6}$$

ΔE_{conf} was calculated as change of the average conformational energy of solute (actin) from states N to H using the NAMD software package (51). As shown in the cycle (2), the solvation steps are defined not to alter actin structures. Therefore, ΔE_{conf} *in vacuo* is calculated using snapshots of MD simulations in solution. $\Delta\mu$ was estimated by the method proposed in the reference (102) and is divided into two contributions,

$$\Delta\mu = \Delta\mu_{\text{polar}} + \Delta\mu_{\text{nonpolar}}, \tag{7}$$

where $\Delta\mu_{\text{polar}}$ and $\Delta\mu_{\text{nonpolar}}$ are the polar and nonpolar solvation free energy, respectively. $\Delta\mu_{\text{polar}}$ was estimated by calculating the Poisson dielectric continuum model using DelPhi

software package (103) with water dielectric constants estimated using the Harris and Alder *g*-factor (104) at the simulated temperature and pressure. $\Delta\mu_{\text{nonpolar}}$ can be approximately decomposed into the contributions from three components (105),

$$\Delta\mu_{\text{nonpolar}} \approx pV + \gamma A + \Delta\mu^{\text{vdW}}, \tag{8}$$

where γ, A, and $\Delta\mu^{\text{vdW}}$ are the surface tension, the surface area, and the free energy of van der Waals attraction. From the values given in (106), γ at 277 K was set to 0.1091 kcal/mol/Å^2. I ignored the $\Delta\mu^{\text{vdW}}$ term because its contribution to yield improvements was less than 0.1% (102). Solute entropy, *S*, is calculated by the total sum of entropy in the translational (S^{trans}), rotational (S^{rot}), and internal motion (S^{int}) (107).

$$S = S^{\text{trans}} + S^{\text{rot}} + S^{\text{int}}, \tag{9}$$

which are defined by,

$$S^{\text{trans}} = k_{\text{B}} \ln\left\{\frac{(2\pi M k_{\text{B}} T)^{3/2}}{h^3} e^{5/2} V\right\}, \tag{10}$$

$$S^{\text{rot}} = k_{\text{B}} \ln\left\{\frac{(2\pi k_{\text{B}} T)^{3/2} (I_x I_y I_z)^{1/2}}{h^3} e^{3/2} 8\pi^2\right\}, \tag{11}$$

$$S^{\text{int}} = \sum_i k_{\text{B}} \left\{\frac{1}{\alpha_i (\exp \alpha_i^{-1} - 1)} - \ln(1 - \exp(-\alpha_i^{-1}))\right\}, \qquad \alpha_i = 2\pi k_{\text{B}} T / h\omega_i \tag{12}$$

where *h*, *M*, *V*, and ω_i are the Planck constant, mass of protein, volume in L/mol, and angular frequency of normal mode, respectively. I_x, I_y, and I_z denote protein principal moments of inertia. S^{int} is deduced from the covariance matrix of coordinates, similar to so-called configurational entropy (108). The ω_i is calculated as the effective frequency of principal mode (109) obtained by the principal component analysis of simulations using mass-weighted all-atom coordinates. Equation (12) of S^{int} includes the kinetic term whereas the configurational entropy has only configurational integral of potential term. Assuming that states N and H are invariant in the process of solvation in the thermodynamics cycle, I calculated the solute entropy from the MD trajectories in solution. Last 50-ns simulation was divided into five 10-ns simulations. I independently calculated the entropy using these

10-ns simulations and then obtained the average entropies of actins.

All molecular images shown in figures of this thesis were generated using VMD software package (91).

2-3 Results and Discussion

2-3-1 *Comparison of structure and fluctuation*

First, the average structures resulting from the last 50-ns trajectories of 12 MD simulations were compared based on the mutual root mean-square deviation (RMSD) of the backbone heavy atoms. The maximum RMSD in all 66 combinations was 2.5 Å between Arm and Rab at 0.1 MPa, whereas the minimum RMSD was 1.5 Å between Arm and Ac1Q at 0.1 MPa. The average RMSD was 2.0 ± 0.3 Å. For deep-sea fish actins, the RMSDs between high- and low-pressure structures were 1.5 and 2.3 Å for Yaq and Arm, respectively. For non-deep-sea fish actins, the RMSDs between high- and low-pressure structures were in the range 2.0-2.2 Å. There were no significant differences between the different actins with respect to the magnitude of the average conformational change induced by high pressure.

The “propeller angle” defined by the relative rotation between subdomains (Figure 3) was shown to change upon polymerization to F-actin (14). I calculated the propeller angle as the torsion angle defined by the centers of four subdomains (Table 9). Variations of the propeller angle among the species and shift of the average propeller angle caused by high pressure were not largely different from the standard deviations.

Table 9. Propeller angle defined by the actin subdomains.

	Propeller angle (°)		
Label	0.1 MPa	60 MPa	Δ
Rab	16.9 ± 1.9	17.8 ± 3.9	0.8 ± 4.4
Ac1W	20.2 ± 1.9	22.7 ± 2.6	2.5 ± 3.2
Ac1Q	18.0 ± 1.7	21.1 ± 3.3	3.1 ± 3.7
Ac2	20.7 ± 2.0	21.1 ± 2.7	0.4 ± 3.4
Arm	**17.6 ± 1.8**	**21.3 ± 2.0**	**3.7 ± 2.7**
Yaq	**17.5 ± 2.2**	**14.8 ± 2.5**	**−2.6 ± 3.4**

Δ = $(\text{Propeller angle})_{60\text{MPa}} - (\text{Propeller angle})_{0.1\text{MPa}}$. The value after "±" indicates standard deviation.

The root mean-square fluctuation (RMSF) per residues from the average structure using actin backbone heavy atoms at 60 MPa is shown in Figure 6. The most flexible region is the DNase I binding loop (residues 42-55) of subdomain 2 (Figure 2). This region had a notably high RMSF compared with all other regions. The V-stretch region (residues 227-237) corresponded to the second highest peak except for the N-terminal region peak. The V-stretch consists of one α-helix (residues 222-233) of subdomain 4 exposed in G-actin, which also make no contact with adjacent protomers in F-actin. RMSF of the substituted residues in deep-sea fish actins (residues V54, P67, and K137) were very small (RMSF < 1 Å). Figure 6 also shows the secondary structure profiles of the actins at 60 MPa. Simulations of the DNase I binding loop returned a variety of secondary structures because this region is very flexible. The position of the α-helices and β-strands in the deep-sea fish actins corresponded to those of the non-deep-sea fish actins except for some end regions of the α-helix and β-strand. Although high pressure affects actin function, the secondary structures were well-maintained even under high pressure.

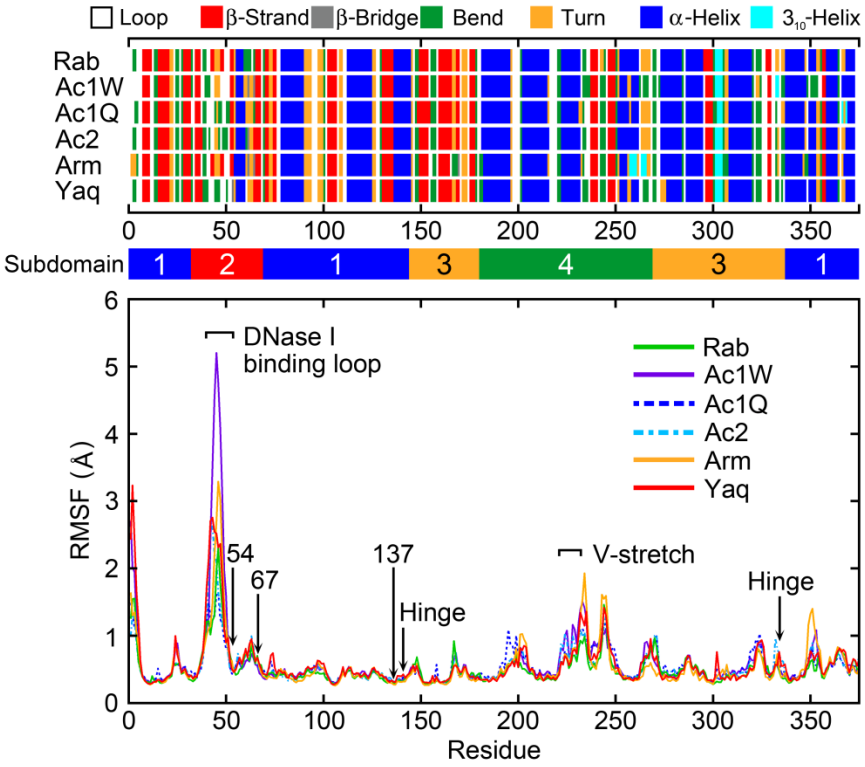

Figure 6. The RMSF was calculated by best-fitting the backbone heavy atoms of each snapshot to the average structure. Secondary structure and subdomain assignments are also shown.

2-3-2 *Analysis of Vex, SASA, and* κ_T

As representative quantities to examine the effects of pressure, V_{ex}, SASA, and κ_T were calculated and are shown in Table 10. No notable differences were observed between deep-sea and non-deep-sea fish actins with respect to these parameters. Although V_{ex} tended to decrease slightly at higher pressure, the differences were comparable to the standard deviations, as were the differences in the values for SASA. Denaturation associated with high pressure generally induces a decrease in V_{ex} and an increase in SASA due to protein unfolding; however, the pressures examined in this work were much lower than that necessary for denaturation of actin, which begins to occur at 250 MPa (81).

Table 10. Effect of high pressure on excluded volume (v_{ex}) and solvent accessible surface area (SASA).

	V_{ex}			SASA		
Label	0.1 MPa	60 MPa	Δ	0.1 MPa	60 MPa	Δ
Rab	7.39 ± 0.02	7.37 ± 0.02	−0.02	1.82 ± 0.02	1.84 ± 0.02	0.02
Ac1W	7.40 ± 0.02	7.41 ± 0.03	0.01	1.84 ± 0.02	1.87 ± 0.02	0.03
Ac1Q	7.40 ± 0.02	7.39 ± 0.02	−0.01	1.84 ± 0.02	1.86 ± 0.03	0.02
Ac2	7.41 ± 0.03	7.36 ± 0.02	−0.05	1.84 ± 0.02	1.83 ± 0.02	−0.01
Arm	**7.39 ± 0.02**	**7.36 ± 0.03**	**−0.03**	**1.83 ± 0.02**	**1.84 ± 0.02**	**0.01**
Yaq	**7.42 ± 0.02**	**7.40 ± 0.02**	**−0.02**	**1.85 ± 0.02**	**1.87 ± 0.02**	**0.02**

Units: V_{ex} (10^4 Å^3), SASA (10^4 Å^2), κ_T, (GPa^{-1}). $\Delta = X_{60MPa} - X_{0.1MPa}$ where $X = V_{ex}$, SASA, or κ_T. The value after "±" indicates standard deviation.

A positive globular protein adiabatic compressibility suggests a more compact structure at high pressure. Ultrasonic measurements indicate that the adiabatic compressibilities of filamentous proteins including F-actin and myosin are negative, which confirms that this property is related to the hydration of protomer surfaces (110). In this work, the κ_T of the various actins examined was in the range 0.13-0.15 GPa^{-1}; no clear systematic differences were observed. The reported κ_T values as determined from sound velocity measurements for 25 globular proteins whose molecular weights ranged from 12,400-232,000 were in the range 0.0192-0.150 GPa^{-1} (111). The results also indicated that larger proteins are more compressible. The molecular weight of actin is ~41,800. The reported κ_T values of comparable seven proteins with a molecular weight between 30,000 and 70,000 is in the range 0.0932-0.150 GPa^{-1} except for peroxidase whose κ_T is exceptionally small (0.0670 GPa^{-1}) (Table 12). Therefore, the κ_T of actin is comparable to that of proteins of similar size. Of note, the κ_T for pure water determined from sound velocity measurements at different pressures is 49.175 Mbar^{-1} (0.49175 GPa^{-1}) at 278 K and 0.1 MPa and 41.912 Mbar^{-1} (0.41912 GPa^{-1}) at 278 K and 60 MPa (112).

Table 11. Effect of high pressure on isothermal compressibility (κ_T).

Label	0.1 MPa	60 MPa	Δ
Rab	0.14 ± 0.02	0.13 ± 0.02	−0.01
Ac1W	0.14 ± 0.02	0.15 ± 0.04	0.01
Ac1Q	0.14 ± 0.02	0.14 ± 0.01	−0.01
Ac2	0.15 ± 0.04	0.14 ± 0.02	−0.01
Arm	**0.13 ± 0.02**	**0.15 ± 0.04**	**0.02**
Yaq	**0.13 ± 0.02**	**0.15 ± 0.03**	**0.02**

Units: κ_T, (GPa^{-1}). $\Delta = \kappa_{T\ 60MPa} - \kappa_{T\ 0.1MPa}$. The value after "±" indicates standard deviation.

Table 12. Isothermal compressibility (κ_T) and molecular weight in globular proteins (111).

Protein	κ_T (GPa^{-1})	Molecular weight
Carbonic anhydrase	0.1050	30,000
Pepsin	0.1270	35,500
Insulin	0.1340	36,000
Peroxidase	0.0670	40,000
α-actin	-	**41,800**
α-amylase	0.0932	45,500
Ovalbumin	0.121	46,000
Bovine serum albumin	0.146	68,000
Hemoglobin	0.150	68,000

2-3-3 *Free energy analysis*

Energy shifts caused by high pressure (Table 13) were examined by the method described in Methods section. Details of each energy term are shown in Table 14 to Table 16. Free energy differences between 60 and 0.1 MPa (ΔG) were all positive, indicating that actin at 60 MPa is less stable compared to 0.1 MPa. ΔG values of Arm and Yaq were the lowest and second lowest, respectively, and were significantly lower than the others. This is consistent to the fact that Arm and Yaq are stable at high pressure. ΔE_{conf} were significantly negative for Arm and Yaq, which might be the primary factor of the stabilization of Arm and Yaq. Contributions of electrostatic interactions in ΔE_{conf} were −158 ± 69 (Arm) and −167 ± 99 kcal/mol (Yaq), which were dominant term in ΔE_{conf}. These results indicate that deep-sea fish actins at high pressure are stabilized by the conformational energy decrease.

Table 13. Energy differences of actins between 60 and 0.1 MPa.

Label	ΔE_{conf}	$\Delta\Delta\mu$	$T\Delta S$	ΔG	$\Delta\Delta G$
Rab	−59 ± 53	520 ± 32	18 ± 18	444 ± 57	−135
Ac1W	11 ± 103	532 ± 14	16 ± 26	527 ± 122	−51
Ac1Q	16 ± 85	575 ± 43	13 ± 15	579 ± 65	0
Ac2	−52 ± 127	512 ± 29	19 ± 24	441 ± 112	−138
Arm	**−147 ± 67**	**510 ± 22**	**29 ± 20**	**334 ± 69**	**−244**
Yaq	**−153 ± 92**	**535 ± 25**	**30 ± 11**	**352 ± 77**	**−226**

Unit: kcal/mol. $\Delta X = X_{60MPa} - X_{0.1MPa}$ where $X = E_{conf}$, $\Delta\mu$, TS, or G. $\Delta G = \Delta E_{conf} + \Delta\Delta\mu - T\Delta S$. $\Delta\Delta G$ is the difference from ΔG of Ac1Q.

Table 14. Effect of high pressure on solute energy.

	E_{conf} (kcal/mol)		
Label	0.1 MPa	60 MPa	Δ
Rab	−3735 ± 41	−3794 ± 33	−59 ± 53
Ac1W	−3632 ± 68	−3621 ± 77	11 ± 103
Ac1Q	−3705 ± 52	−3688 ± 68	16 ± 85
Ac2	−3612 ± 123	−3664 ± 30	−52 ± 127
Arm	**−3578 ± 44**	**−3725 ± 50**	**−147 ± 67**
Yaq	**−3543 ± 82**	**−3696 ± 41**	**−153 ± 92**

$\Delta = (E_{conf})_{60MPa} - (E_{conf})_{0.1MPa}$. The value after "±" indicates standard deviation.

Table 15. Effect of high pressure on solute entropy.

	0.1 MPa			
Label	TS^{trans}	TS^{rot}	TS^{int}	TS
Rab	126 ± 0	104 ± 0	4432 ± 11	4662 ± 11
Ac1W	126 ± 0	104 ± 0	4413 ± 19	4643 ± 19
Ac1Q	126 ± 0	104 ± 0	4400 ± 12	4630 ± 12
Ac2	126 ± 0	104 ± 0	4382 ± 8	4612 ± 8
Arm	**126 ± 0**	**104 ± 0**	**4405 ± 14**	**4635 ± 14**
Yaq	**126 ± 0**	**104 ± 0**	**4393 ± 10**	**4623 ± 10**

	60 MPa				
Label	TS^{trans}	TS^{rot}	TS^{int}	TS	ΔTS
Rab	126 ± 0	104 ± 0	4450 ± 15	4680 ± 15	18 ± 18
Ac1W	126 ± 0	104 ± 0	4429 ± 18	4659 ± 18	16 ± 26
Ac1Q	126 ± 0	104 ± 0	4413 ± 8	4642 ± 8	13 ± 15
Ac2	126 ± 0	104 ± 0	4402 ± 23	4631 ± 23	19 ± 24
Arm	**126 ± 0**	**104 ± 0**	**4434 ± 14**	**4664 ± 14**	**29 ± 20**
Yaq	**126 ± 0**	**104 ± 0**	**4423 ± 4**	**4653 ± 4**	**30 ± 11**

Unit: kcal/mol. $\Delta TS = TS_{60MPa} - TS_{0.1MPa}$. The value after "±" indicates standard deviation.

Table 16. Effect of high pressure on actin solvation energy.

	0.1 MPa			60 MPa		
Label	$\Delta\mu_{polar}$	$\Delta\mu_{nonpolar}$	$\Delta\mu$	$\Delta\mu_{polar}$	$\Delta\mu_{nonpolar}$	$\Delta\mu$
Rab	−4638 ± 10	1952 ± 10	−2686 ± 13	−4771 ± 14	2606 ± 17	−2165 ± 29
Ac1W	−4663 ± 15	1971 ± 10	−2691 ± 10	−4801 ± 16	2642 ± 22	−2159 ± 10
Ac1Q	−4666 ± 15	1971 ± 11	−2696 ± 11	−4757 ± 23	2636 ± 23	−2121 ± 43
Ac2	−4681 ± 35	1978 ± 10	−2703 ± 27	−4789 ± 6	2598 ± 14	−2191 ± 10
Arm	**−4657 ± 8**	**1962 ± 11**	**−2695 ± 18**	**−4793 ± 14**	**2608 ± 13**	**−2185 ± 13**
Yaq	**−4670 ± 24**	**1990 ± 15**	**−2680 ± 22**	**−4791 ± 7**	**2647 ± 14**	**−2144 ± 12**

Label	$\Delta\Delta\mu_{polar}$	$\Delta\Delta\mu_{nonpolar}$	$\Delta\Delta\mu$
Rab	−133 ± 17	653 ± 20	520 ± 32
Ac1W	−139 ± 22	671 ± 25	532 ± 14
Ac1Q	−91 ± 28	666 ± 25	575 ± 44
Ac2	−107 ± 35	619 ± 17	512 ± 29
Arm	**−136 ± 16**	**646 ± 17**	**510 ± 22**
Yaq	**−121 ± 25**	**657 ± 20**	**535 ± 25**

Unit: kcal/mol. $\Delta\Delta X = X_{60MPa} - X_{0.1MPa}$, $X = \Delta\mu_{polar}$, $\Delta\mu_{nonpolar}$, or $\Delta\mu$. The value after "±" indicates standard deviation.

2-3-4 *Hydrogen bond and salt bridge analyses*

The results of the free energy analysis suggest that intra-solute interaction in deep-sea fish actins is a key to understand high pressure tolerance. To examine this, I conducted hydrogen bond and salt bridge analyses. The number of hydrogen bonds in actin monomers and between actin monomers and water was shown in Table 17. Within actin, variations in the hydrogen-bond numbers among the species and shift of the average hydrogen-bond numbers caused by high pressure were not largely different from the standard deviations. The number of hydrogen bonds between actin and water molecules showed more variations among the species. This quantity can be related to solvation free energy, rather than conformational energy. The number of hydrogen bonds increased at high pressure in all the

cases. There was no clear correlation between the number of hydrogen bonds and pressure tolerance.

Table 17. The number of hydrogen bonds in actin and between actin and water.

	The number of hydrogen bonds within actin			The number of hydrogen bonds between actin and water		
Label	0.1 MPa	60 MPa	Δ	0.1 MPa	60 MPa	Δ
Rab	273 ± 6	272 ± 6	−1 ± 8	874 ± 24	937 ± 16	63 ± 29
Ac1W	273 ± 6	265 ± 6	−8 ± 8	927 ± 16	953 ± 21	26 ± 26
Ac1Q	273 ± 6	269 ± 6	−4 ± 8	913 ± 23	939 ± 15	27 ± 27
Ac2	271 ± 6	272 ± 6	2 ± 8	995 ± 99	1011 ± 95	16 ± 137
Arm	**273 ± 7**	**274 ± 6**	**1 ± 9**	**922 ± 15**	**955 ± 22**	**33 ± 26**
Yaq	**268 ± 6**	**265 ± 6**	**−3 ± 8**	**926 ± 17**	**957 ± 15**	**31 ± 23**

$\Delta = (\text{Hydrogen bond})_{60\text{MPa}} - (\text{Hydrogen bond})_{0.1\text{MPa}}$. The value after "±" indicates standard deviation.

Table 18 shows the number of salt bridges between ATP and surrounding residues. A salt bridge was considered to be formed if the distance between oxygen and nitrogen atoms of charged groups was less than or equal to 3.2 Å (91). It should be noted that multiple salt bridges can be formed in a pair of residues with this definition. Deep-sea fish actins have K137 at the active site, whereas non-deep-sea fish actins have Q137. The Q137K substitution in deep-sea fish actin changes the charge at this position from neutral to positive. Deep-sea fish actins formed a salt bridge between the γ-oxygen atoms of ATP (Figure 5) and the side chain of K137 at 0.1 and 60 MPa. Deep-sea fish actins formed more salt bridges between ATP and the surrounding residues than non-deep-sea fish actins. Ac2 had the least number of salt bridges at 60 MPa. Although deep-sea fish actins include some proportion of Ac2 (20% in *C. armatus* and 19% in *C. yaquinae*), Ac2 is found predominantly in non-deep-sea fish actins (60% in *C. acrolepis*). The ligand dissociation

rate constants of non-deep-sea fish actins were shown to increase notably at high pressure, whereas those of deep-sea fish actins are less affected (12). The salt bridge between ATP and K137 is expected to stabilize ATP binding at high pressure, thus enhancing the protein's pressure tolerance. Residue K137 in deep-sea fish actins is located near the hinge region (residues 141-142 or 336-337) of the propeller motion. All of the subdomains also make contact with the active site. Furthermore, transformation of G- to F-actin is associated with the hydrolysis of ATP. Therefore, a conformational change in the ATP γ-phosphate bound to K137 is expected to play a large role in affecting the change in actin structure occurring upon filament formation. The Q137K substitution might trigger a propagation of the changes in protein conformation and salt bridge pattern.

Table 18. Number of salt bridges formed between ATP and surrounding residues.

Label	K18-O_α	K18-O_β	K137-O_γ	Total
		0.1 MPa		
Rab	1.0 ± 0.0	1.5 ± 0.6	-	2.5 ± 0.6
Ac1W	0.8 ± 0.4	1.4 ± 0.6	-	2.2 ± 0.7
Ac1Q	0.7 ± 0.6	1.1 ± 0.6	-	1.8 ± 0.8
Ac2	1.0 ± 0.1	1.4 ± 0.6	-	2.4 ± 0.6
Arm	**1.0 ± 0.2**	**0.6 ± 0.6**	**1.1 ± 0.4**	**2.7 ± 0.8**
Yaq	**1.0 ± 0.1**	**1.0 ± 0.3**	**1.7 ± 0.5**	**3.7 ± 0.5**
		60 MPa		
Rab	1.0 ± 0.0	1.1 ± 0.7	-	2.1 ± 0.7
Ac1W	1.0 ± 0.0	1.4 ± 0.5	-	2.4 ± 0.5
Ac1Q	1.0 ± 0.0	1.3 ± 0.5	-	2.3 ± 0.5
Ac2	0.3 ± 0.5	1.0 ± 0.1	-	1.3 ± 0.5
Arm	**1.0 ± 0.0**	**0.9 ± 0.3**	**1.4 ± 0.5**	**3.3 ± 0.6**
Yaq	**1.0 ± 0.2**	**0.8 ± 0.6**	**1.0 ± 0.3**	**2.8 ± 0.7**

K18-O_α, K18-O_β, and K137-O_γ represent salt bridges between K18 and α-oxygen, between K18 and β-oxygen, and between K137 and γ-oxygen of ATP, respectively. The value after "±" indicates standard deviation.

Table 19 shows the number of salt bridges formed between secondary structures and subdomains. I first focus on differences in the number of salt bridges among different species. Deep-sea fish actins tended to form more salt bridges than the actins of other species. With the exception of Yaq and Arm, the rank order with respect to the total number of salt bridges formed corresponded to the experimentally determined rank order with respect to pressure tolerance (i.e., Yaq was the most pressure tolerant, followed by Arm, Rab, Ac2, and Ac1). The differences of the total number of salt bridges between Yaq and Arm at 0.1 and 60 MPa were within the range of standard deviations. These data suggest that the number of salt bridges formed is closely related to the degree of pressure tolerance.

Deep-sea fish actins also formed more inter-helix/strand salt bridges than non-deep-sea fish actins. The number of intra-helix/strand salt bridges formed by the various actins was comparable, except for Rab, which tended to form more salt bridges within the helix and strand at high pressure (Table 19). Table 8 shows the positions of residues involved in the formation inter-helix/strand salt bridges.

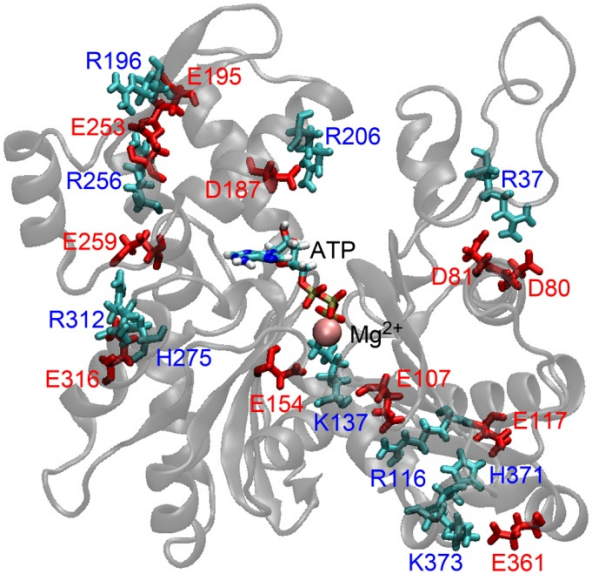

Figure 7. Salt bridge between secondary structures in Yaq at 60 MPa. The residues that form salt bridges with a formation rate of more than 0.5 are shown in Yaq at 60 MPa. Red and blue represent acidic and basic amino acids, respectively.

These residues are expected to stabilize the arrangements between secondary structures at multiple sites in deep-sea fish actins. The number of inter- and intra-subdomain salt bridges is also shown in Table 19. It is clear that deep-sea fish actins form more inter-subdomain salt bridges than do actins from other species, and this might play a role in stabilizing the subdomain arrangement. It should be noted that no inter-subdomain salt bridges were found between subdomains 2 and 4 (Figure 8). Therefore these inter-subdomain salt bridges do not interfere with propeller angle rotation. Changes in the number of salt bridges induced

by high pressure were also seen in Table 19. The average total number of salt bridges did not largely change in Rab, Arm and Yaq (From −0.2 to +0.6 changes whereas the standard deviations are 2.6-3.1) that are relatively high pressure tolerant, but larger increase in the salt bridge number (From +3.7 to +6.0 increases whereas the standard deviations are 2.3-3.8) was observed in Ac1W, Ac1Q and Ac2 that are less tolerant to pressure. The results indicate that only small change in terms of salt bridges was caused by high pressure in Rab, Arm and Yaq, showing the robustness of these actins.

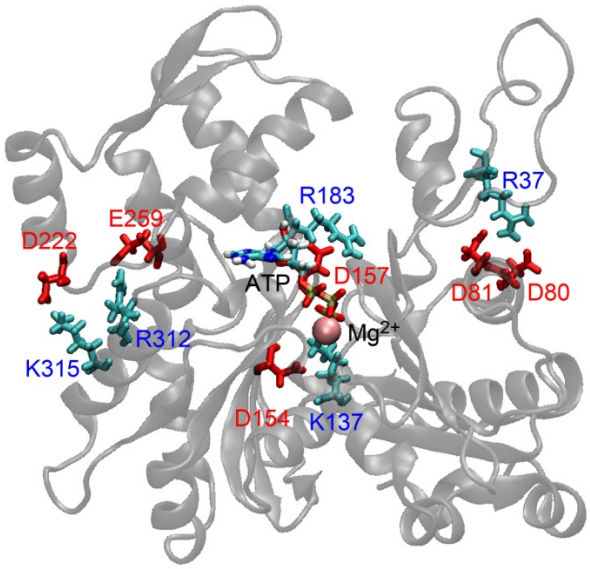

Figure 8. Salt bridge between subdomains in Yaq at 60 MPa. The residues that form salt bridges with a formation rate of more than 0.5 are shown in Yaq at 60 MPa. Red and blue represent acidic and basic amino acids, respectively.

Table 19. Number of salt bridges formed between secondary structures and subdomains.

	Secondary structure				ATP	Subdomain		
Label	Inter helix/ strand [a]	Helix/strand and loop [b]	Loop and loop [c]	Intra helix/ strand [d]	ATP and residues [e]	Inter subdomain [f]	Intra subdomain [g]	Total [h]
0.1 MPa								
Rab	16.4 ± 1.5	9.9 ± 2.1	0.0 ± 0.0	11.3 ± 1.7	2.5 ± 0.6	10.2 ± 1.4	27.4 ± 2.6	40.2 ± 3.1
Ac1W	14.7 ± 1.5	6.9 ± 1.5	0.2 ± 0.4	6.8 ± 1.3	2.2 ± 0.7	8.0 ± 0.9	20.5 ± 2.0	30.7 ± 2.3
Ac1Q	11.9 ± 2.1	8.4 ± 1.5	0.4 ± 0.5	6.7 ± 1.4	1.8 ± 0.8	7.4 ± 1.3	20.0 ± 2.4	29.2 ± 2.8
Ac2	12.8 ± 1.8	8.5 ± 2.1	0.7 ± 0.9	8.0 ± 1.7	2.4 ± 0.6	7.0 ± 1.4	23.0 ± 3.2	32.4 ± 3.8
Arm	**20.5 ± 1.6**	**9.2 ± 1.4**	**1.0 ± 1.2**	**8.9 ± 1.4**	**2.7 ± 0.8**	**10.7 ± 1.3**	**28.9 ± 2.4**	**42.4 ± 2.6**
Yaq	**19.1 ± 1.8**	**9.0 ± 2.6**	**0.1 ± 0.3**	**9.5 ± 1.7**	**3.7 ± 0.5**	**11.1 ± 1.4**	**26.6 ± 3.5**	**41.4 ± 4.4**
60 MPa								
Rab	15.9 ± 1.5	11.2 ± 1.9	0.0 ± 0.1	11.6 ± 2.2	2.1 ± 0.7	9.2 ± 1.5	29.4 ± 2.8	40.8 ± 2.9
Ac1W	15.7 ± 1.4	9.9 ± 2.2	0.0 ± 0.0	6.4 ± 1.2	2.4 ± 0.5	8.2 ± 1.4	23.8 ± 2.7	34.4 ± 3.2
Ac1Q	14.1 ± 1.4	8.9 ± 1.6	0.0 ± 0.2	7.9 ± 1.4	2.3 ± 0.5	7.9 ± 1.0	23.1 ± 2.6	34.2 ± 2.4
Ac2	14.7 ± 1.8	11.1 ± 1.5	2.3 ± 0.7	9.0 ± 1.6	1.3 ± 0.5	9.1 ± 1.3	28.0 ± 2.2	38.4 ± 2.6
Arm	**19.4 ± 1.4**	**11.1 ± 1.6**	**0.0 ± 0.0**	**8.3 ± 2.1**	**3.3 ± 0.6**	**11.5 ± 1.3**	**27.4 ± 2.5**	**42.2 ± 2.8**
Yaq	**18.8 ± 1.6**	**11.8 ± 1.8**	**0.2 ± 0.4**	**8.4 ± 1.7**	**2.8 ± 0.7**	**10.5 ± 1.2**	**28.7 ± 2.6**	**41.9 ± 2.6**

Salt bridges [a]between distinct helices or strands, [b]between a helix/strand and a loop, [c]between distinct loops, [d]within helix or strand. [e]Salt bridges between ATP and a residue. [f]Inter and [g]intra subdomain salt bridge. [h]The sum of "Secondary structure" + "ATP" or "Subdomains" + "ATP". The value after "±" indicates standard deviation.

2-3-5 *Possible ATP hydrolysis mechanism*

It was reported that the Q137A mutant actin polymerized four times faster than wild-type actin, but cleavage of the ATP γ-phosphate group occurred at only one-fourth of the rate of wild-type actin, indicating the residue 137 has a significant effect on these processes (79). I observed that the difference in the amino acid residue at residue 137 between deep-sea

fish and non-deep-sea fish actins also alters the orientation of the side chain. Actin hydrolyzes ATP mainly in F-actin elongation process and the enzyme activity is very weak in G-actin. The nucleophilic water attacks the γ-phosphate of ATP during the hydrolysis. The inactive nucleophilic water is probably held by the residue 137 and the water molecule bound to H161 (113) (Figure 9). Residue 137 is located in subdomain 1, whereas H161 is located in subdomain 3. It was suggested that H161 is moved by the conformational changes that occur when the adjacent actin protomer makes contact and ATP hydrolysis subsequently begins (14). In all of the MD simulations, I confirmed that ATP binds Mg^{2+} at the active site tightly, with a coordination number of 6.0 for all of the actins examined (Table 20).

Table 20. The coordination number of Mg^{2+} at the active site.

	Coordination number		
MD	0.1 MPa	60 MPa	Δ
Rab	6.0 ± 0.1	6.0 ± 0.2	0.0
Ac1W	6.0 ± 0.1	6.0 ± 0.1	0.0
Ac1Q	6.0 ± 0.1	6.0 ± 0.1	0.0
Ac2	6.0 ± 0.1	6.0 ± 0.1	0.0
Arm	**6.0 ± 0.1**	**6.0 ± 0.1**	**0.0**
Yaq	**6.0 ± 0.1**	**6.0 ± 0.1**	**0.0**

$\Delta = (\text{Coordination number})_{60MPa} - (\text{Coordination number})_{0.1MPa}$.

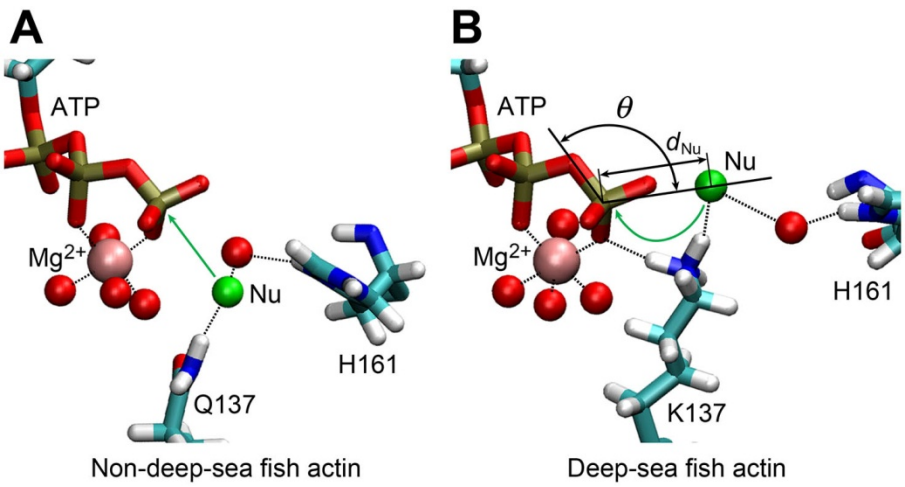

Figure 9. Arrangement of the water molecule expected to initiate nucleophilic attack on the γ-phosphate of ATP. The arrangement of non-deep-sea fish actins (A) and deep-sea fish actins (B). Green spheres show water molecules expected to be nucleophilic water for ATP hydrolysis. Red spheres indicate the water molecules coordinated to Mg^{2+} and those bridging the expected nucleophilic water and H161 with hydrogen bonds. Black dotted lines show typical hydrogen bonds formed during the MD simulation. Angle θ and distance d_{Nu} are defined by O^{β}-P^{γ}-O^{w} and P^{γ}-O^{w}, respectively, where O^{w} represents the oxygen of the expected nucleophilic water (see Figure 5 for the definition of the other atoms).

In yeast G-actin, the energy barrier of hydrolysis neglecting entropic contribution was 28.8 kcal/mol estimated using the minimum energy path of quantum chemical calculations (113). A QM/MM model of ATP hydrolysis without protein was carried out including the entropic contribution and the energy barrier was 33.4 kcal/mol (114). Since ATP hydrolysis in enzyme active sites is not investigated using QM/MM models with the entropic contribution, the detailed energy of the hydrolysis is unclear yet. I considered the effect of K137 binding to the γ-phosphate group. In general, a divalent cation assists the process of ATP hydrolysis. If K137 is bound to the O^{γ} in deep-sea fish actin during the hydrolysis as observed in the MD, more positive charge is coordinated to the O^{γ} and the energy barrier of the hydrolysis could be lowered in the intermediate state (Figure 10). In this case, the rate of hydrolysis can be accelerated in deep-sea fish actins; however, the rate of hydrolysis in

actins of deep-sea fish and non-deep-sea fish are comparable at low pressure (12). Therefore, Q137K substitution probably has also disadvantage for the hydrolysis.

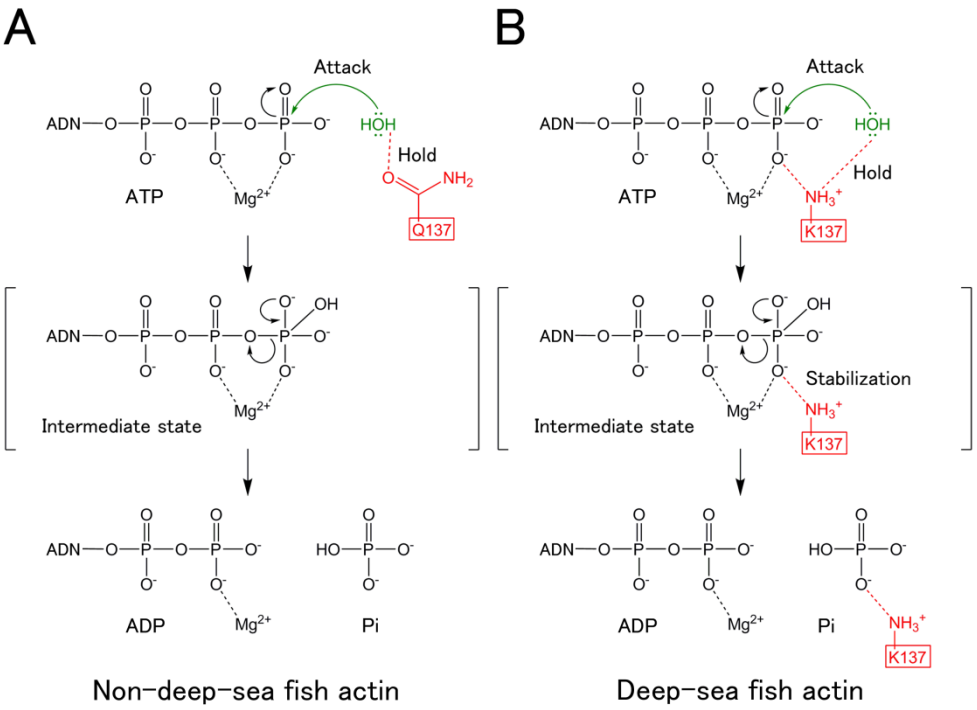

Figure 10. Possible mechanism of ATP hydrolysis in (A) non-deep-sea fish actins and (B) deep-sea fish actins. Green water molecules indicate the nucleophilic water attacking the γ-phosphate of ATP. Brackets indicate the intermediate state of ATP hydrolysis resulting from binding of the nucleophilic water. ADN denotes the adenosine.

I focused on the positions of water molecules in the active site. Figure 11 indicates the distribution of the expected nucleophilic water in the hydrolysis (see Figure 11 legend for the definition of expected nucleophilic water) shown by free energy scale as the function of the angle θ and distance d_{Nu}, which are defined by O^{β}-P^{γ}-O^{w} and P^{γ}-O^{w}, respectively, where O^{w} is the oxygen atom of the expected nucleophilic water molecule. The free energy minimum in deep-sea fish actins was around 150° at 0.1 MPa, whereas in non-deep-sea fish

actins the free energy minimum was around 170° (Figure 11A). Since the nucleophilic water linearly attacks the γ-phosphate of ATP ($\theta = \sim 180°$), non-deep-sea fish actins at 0.1 MPa maintained one water molecule at the favorable position for nucleophilic attack. This in-line arrangement was also observed in a preceding study of MD simulations (115). At 60 MPa, the θ angle had a free energy minimum at around 160-170° in non-deep-sea fish actins and around 140-150° in deep-sea fish actins (Figure 11B and Table 21). In deep-sea fish actins, residue K137 is directly bound to the γ-oxygen atoms of ATP and expected nucleophilic water (Figure 9). The P^{γ}-O^{w} distance (d_{Nu}) was 4.6 Å or less in each of the non-deep-sea fish actins at 60 MPa and 5.3-5.5 Å in the deep-sea fish actins (Table 21). Thus, the expected nucleophilic water in deep-sea fish actins is slightly shifted both in θ and d_{Nu} from the best in-line position. The stabilization of the γ-phosphate group in the intermediate state with more positive charge is suggested to be compensated with the effect of the less favorable position of the expected nucleophilic water.

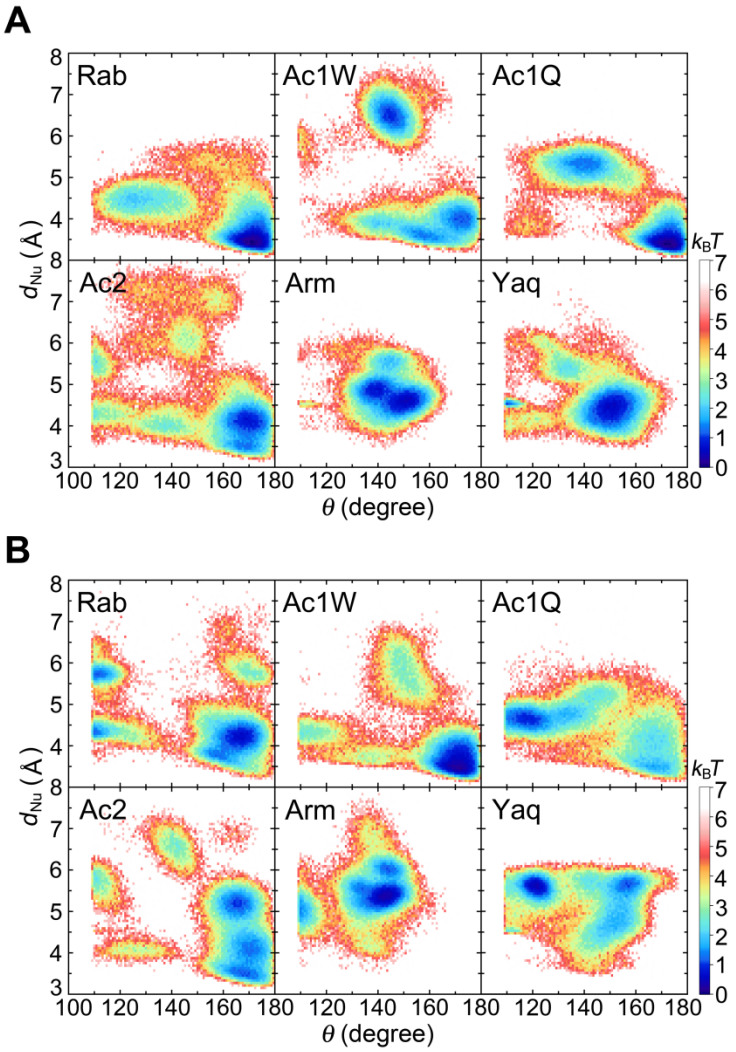

Figure 11. Spatial distribution of expected nucleophilic water. Distribution of expected nucleophilic water as a function of angle θ and distance d_{Nu}, as defined in the legend for Figure 9, converted as free energy scale at (A) 0.1 and (B) 60 MPa. A water molecule having the minimum d_{Nu} value and a θ greater than 109.3° was assigned as the expected nucleophilic water in each simulation snapshot. The free energy was shown as the relative value against the minimum free energy in k_BT.

Table 21. Coordination of the expected nucleophilic water to ATP.

	θ			d_{Nu}			d_{side}		
Label	0.1 MPa	60 MPa	Δ	0.1 MPa	60 MPa	Δ	0.1 MPa	60 MPa	Δ
Rab	159 ± 19	154 ± 21	−5 ± 28	4.0 ± 0.6	4.5 ± 0.7	0.5 ± 0.9	4.0 ± 0.6	3.9 ± 1.1	−0.1 ± 1.3
Ac1W	154 ± 14	162 ± 16	8 ± 21	4.9 ± 1.3	4.1 ± 0.8	−0.9 ± 1.5	3.3 ± 0.8	3.3 ± 0.7	0.0 ± 1.0
Ac1Q	158 ± 18	142 ± 21	−16 ± 27	4.3 ± 0.9	4.5 ± 0.5	0.2 ± 1.0	3.4 ± 0.5	3.4 ± 0.6	0.0 ± 0.8
Ac2	157 ± 19	161 ± 16	4 ± 24	4.5 ± 1.0	4.6 ± 0.9	0.1 ± 1.4	3.6 ± 1.2	3.3 ± 1.3	−0.3 ± 1.8
Arm	**145 ± 8**	**137 ± 11**	**−7 ± 13**	**4.8 ± 0.4**	**5.5 ± 0.5**	**0.7 ± 0.7**	**3.8 ± 0.6**	**4.7 ± 1.1**	**0.9 ± 1.2**
Yaq	**147 ± 12**	**139 ± 16**	**−7 ± 20**	**4.6 ± 0.5**	**5.3 ± 0.5**	**0.7 ± 0.7**	**3.4 ± 0.8**	**4.0 ± 0.6**	**0.6 ± 1.0**

Units: θ (°), d_{Nu} and d_{side} (Å). The value after "±" indicates standard deviation.

I also investigated the exchange of water molecules at the nucleophilic position. Table 22 shows the occupancy of the nucleophilic water molecules at 60 MPa during the last 50-ns MD simulations. The occupancy indicates the ratio of the duration of the specified water staying at the position versus the total sampling MD time, and the water molecules are rank-ordered by occupancy. A notably high value of 0.73 was obtained for rank 1 in Ac1W, meaning that one water molecule continuously remained at the nucleophilic position for 73% of the sampling time at 60 MPa. The other first rank values were 0.32-0.56. The values for deep-sea fish actins were relatively low (0.33-0.35).

Table 22. Occupancy of nucleophilic water molecules at 60 MPa.

	Rank				
MD	1	2	3	4	5
			0.1 MPa		
Rab	0.40	0.35	0.11	0.05	0.04
Ac1W	0.44	0.39	0.08	0.05	0.01
Ac1Q	0.51	0.22	0.19	0.05	0.02
Ac2	0.44	0.25	0.13	0.05	0.03
Arm	**0.86**	**0.03**	**0.03**	**0.03**	**0.03**
Yaq	**0.64**	**0.09**	**0.09**	**0.08**	**0.04**
			60 MPa		
Rab	0.56	0.09	0.09	0.07	0.06
Ac1W	0.73	0.15	0.08	0.03	0.00
Ac1Q	0.32	0.26	0.11	0.11	0.07
Ac2	0.37	0.28	0.12	0.08	0.02
Arm	**0.33**	**0.23**	**0.14**	**0.09**	**0.09**
Yaq	**0.35**	**0.30**	**0.08**	**0.05**	**0.05**

2-3-6 *Effect of V54A and L67P*

Deep-sea fish actins have either a V54A (Arm) or L67P (Yaq) substitution at the surface of subdomain 2. Because the DNase I binding loop of subdomain 2 binds to neighboring actin protomers in F-actin, these substitutions are also expected to contribute to pressure tolerance. Although these two residues are relatively close to protomer-protomer interface in F-actin, they do not directly interact with other protomers. The V54A and L67P substitutions do not significantly alter the polarity, but they do alter the hydrophobicity of the protein to a certain extent. The Kyte-Doolittle hydropathy indexes for valine, alanine, leucine, proline, isoleucine, and threonine are 4.2, 1.8, 3.8, −1.6, 4.5, and −0.7, respectively; that of isoleucine is the highest among the standard amino acids (116). Both V54 and L67 are highly hydrophobic residues in non-deep-sea fish actins,

whereas deep-sea fish actins have less hydrophobic residues at these positions (i.e., A at position 54 and P at position 67). Residues V54 and L67 in non-deep-sea fish actins are directed toward neighboring subdomains and can interact with I85 and T203, respectively (Figure 12).

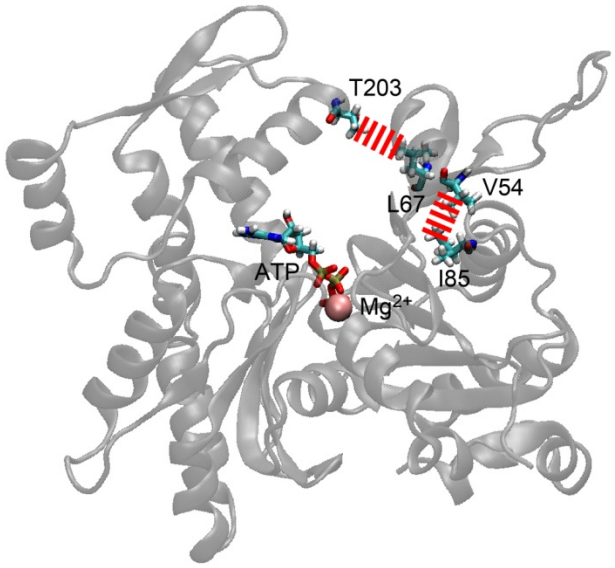

Figure 12. Hydrophobic interactions involving specific substituted residues in Ac1 actin. Red broken lines indicate the hydrophobic interaction. Residues 54 and 67 are different in the actins of deep-sea fish and other species.

Residue I85 is located in subdomain 1 near the boundary with subdomain 2. A strong hydrophobic interaction between the side chains of residues V54 and I85 was formed in all actins examined except Arm. The minimum distance between the side-chain carbon atoms ($d_{54\text{-}85}$) at 0.1 and 60 MPa was 3.9 and 3.8-3.9 Å, respectively, whereas the $d_{54\text{-}85}$ for A54 and I85 at 0.1 and 60 MPa in Arm was 4.4 and 4.3 Å, respectively (Table 23). Because alanine is smaller and less hydrophobic than valine, the $d_{54\text{-}85}$ in Arm was slightly longer than that in other actins. In spite of slightly weaker hydrophobic interaction, there was no change in this distance at 0.1 and 60 MPa in any of the actins examined.

The Arm DNase I binding loop had a tendency to form additional β-strand structure at residues 43-44 and 48-49 compared with actins other than Rab (Figure 6). The tendency might have some effect on the protomer-protomer interaction in F-actin.

Table 23. Minimum inter-residue distances at V54A and L67P.

	$d_{54\text{-}85}$			$d_{67\text{-}203}$		
Label	0.1 MPa	60 MPa	Δ	0.1 MPa	60 MPa	Δ
Rab	3.9 ± 0.2	3.9 ± 0.2	0.0 ± 0.3	4.5 ± 0.7	4.3 ± 0.9	−0.2 ± 1.1
Ac1W	3.9 ± 0.3	3.9 ± 0.3	0.0 ± 0.4	7.0 ± 0.6	7.0 ± 0.9	0.0 ± 1.1
Ac1Q	3.9 ± 0.3	3.9 ± 0.3	0.0 ± 0.4	7.5 ± 0.6	7.1 ± 0.8	−0.4 ± 1.0
Ac2	3.9 ± 0.3	3.9 ± 0.3	0.0 ± 0.4	4.0 ± 0.4	5.0 ± 1.5	1.0 ± 1.6
Arm	**4.4 ± 0.4**	**4.3 ± 0.3**	**−0.1 ± 0.5**	**8.1 ± 0.7**	**6.2 ± 1.0**	**−1.9 ± 1.2**
Yaq	**3.9 ± 0.3**	**3.8 ± 0.2**	**−0.1 ± 0.4**	**7.5 ± 0.4**	**6.8 ± 0.7**	**−0.7 ± 0.8**

Unit: Å. $\Delta = X_{60\text{MPa}} - X_{0.1\text{MPa}}$, $X = d_{54\text{-}85}$ or $d_{67\text{-}203}$. The value after "±" indicates standard deviation

The L67P substitution in Yaq is located at the surface of subdomain 2 near the DNase I binding loop and nucleotide binding cleft between subdomains 2 and 4 (Figure 4). T203 is the nearest residue to residue 67 in subdomain 4. The minimum distance between the side-chain carbon atoms of residues 67 and 203 ($d_{67\text{-}203}$) in Arm and Yaq at 60 MPa was 6.2 ± 1.0 and 6.8 ± 0.7 Å, respectively (Table 23). In non-deep-sea fish actins, $d_{67\text{-}203}$ had more variations among different actins at 60 MPa (4.3 [Rab]-7.1 Å [Ac1Q]).

In summary, the effect of V54A and L67P substitutions on pressure tolerance was not clear in the G-actin simulation. This effect should be examined with F-actin simulation.

2-3-7 *Comparison between two actin 1a models*

In this work, I considered two actin 1a structures, Ac1W and Ac1Q. In the crystal structures, the high-affinity cation binding site consists of a divalent cation, chelate water molecules, and some side-chain atoms. Of the 48 PDB structures having chelating groups at the α-actin active site (PDB files as of June 8, 2013), only six have Q137 as the chelating group. In this sense, the coordination state represented by Ac1W predominates that represented by Ac1Q among the PDB structures. A comparison of the MD simulation results showed that the two coordination states are closely related, both structurally and dynamically. Indeed, the physical properties and salt bridge patterns were almost the same, except for the salt bridge between K18 and the β-oxygen atom of ATP. This salt bridge would slightly attract the phosphate tail to the region included in Q137, also affecting the expected nucleophilic water position in Ac1. The minimum energy coordination of the nucleophilic water in Ac1W at 60 MPa was at θ = ~170° and d_{Nu} = ~3.0 Å, whereas at θ = ~120° and d_{Nu} = ~4.3 Å in Ac1Q. This difference indicates that Ac1W maintains the nucleophilic water at a more favorable in-line position. Since the Mg^{2+} in Ac1Q is maintained at an unfavorable position for attacking the γ-phosphate of ATP, direct coordination of Q137 to Mg^{2+} is probably disadvantageous for ATP hydrolysis. This is also consistent with the cases of Ac2 and Rab, which do not involve binding between Q137 and Mg^{2+} at the active site.

I also analyzed the dihedral angle of residue 137, which had dihedral angles χ_1 to χ_3. At 0.1 MPa, the χ_3 of Ac1Q bound to Mg^{2+} was significantly different than that of Ac1W. The peaks of the dihedral angles in Ac1Q differed at 0.1 and 60 MPa, whereas we did not observe a significant difference in the dihedral angles at 0.1 and 60 MPa in Ac1W (Figure 13). Additional peaks were found both in χ_1 (−44° and ±180°) and χ_3 (88°) at 60 MPa in Ac1Q. Presumably, at 60 MPa the conformation of residue 137 in Ac1Q is unstable, which is consistent with the fact that the crystal structures included in the bond between Q137 and Mg^{2+} are minor. ΔG of Ac1Q was larger than that of Ac1W and was the largest among the actins studied (Table 13). Therefore, Ac1W can be considered as a more plausible model for Ac1.

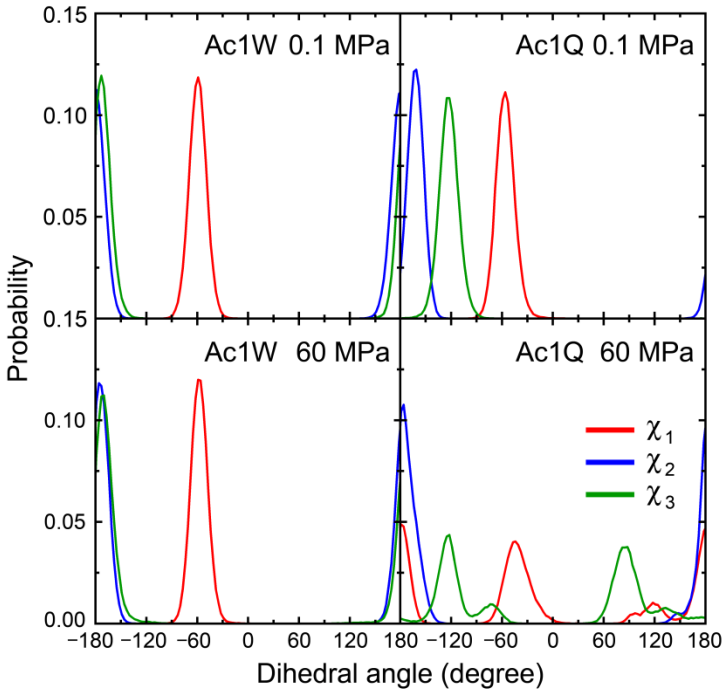

Figure 13. Probability distribution of dihedral angles of residue 137 in Ac1. The red, blue and green lines denote the χ_1, χ_2 and χ_3 dihedral angles of residue 137 in Ac1, respectively.

2-4 Conclusion

Deep-sea fish actins from *C. armatus* (Arm) and *C. yaquinae* (Yaq) have specific substitutions (Q137K and V54A [Arm] or I67P [Yaq]) not found in actins of terrestrial animals or species of shallow-water fish. Although the pressure of the deep-sea habitat is below the actin denaturation pressure, pressure has significant effects on polymerization and the dissociation rates of ATP and Ca^{2+} in non-deep-sea fish actins, whereas the actins of deep-sea fish are tolerant of pressures up to at least 60 MPa (12, 82). In this work, I investigated the effect of the amino acid substitutions on pressure tolerance using MD simulations. I found that high pressure causes only small changes in the excluded volume, SASA, isothermal compressibility, solvation energy, and solute entropy, of both deep-sea and non-deep-sea fish actins, but conformational energy of Arm and Yaq actins

significantly lowered at high pressure. Therefore, I conclude that deep-sea fish actins at high pressure are stabilized in the conformational energy decrease. Salt bridge pattern of Arm and Yaq showed notable differences compared to the others studied in this work. The salt bridges between ATP and K137, which were only formed in deep-sea fish actins, are expected to stabilize ATP binding even under high pressure. Deep-sea fish actins also formed a greater total number of salt bridges than non-deep-sea fish actins, owing to the formation of inter-helix/strand and inter-subdomain salt bridges at high pressure. Therefore, I conclude that two amino acid differences are sufficient to significantly stabilize ATP binding and subunit arrangement through the salt bridges.

Our results also suggest that ATP hydrolysis during elongation of F-actin differs slightly in deep-sea fish actins. Although direct binding of K137 to the γ-phosphate group stabilizes the binding of ATP to deep-sea fish actin, it also causes a slight shift of the water molecule responsible for the nucleophilic attack on the γ-phosphate group from the favorable in-line position. Since residue Q137 of non-deep-sea fish actins does not directly interact with ATP and maintains only the nucleophilic water, the intermediate state of ATP hydrolysis is supported primarily by Mg^{2+}. In deep-sea fish actins, residue K137 is directly bound to the γ-phosphate of ATP and may assist in ATP hydrolysis through electrostatic interaction. This interaction is expected to stabilize the intermediate state of ATP hydrolysis and lower the energy barrier of the reaction. Consequently, ATP is smoothly hydrolyzed in deep-sea fish actins when the nucleophilic water attacks the γ-phosphate of ATP. The effect of the less favorable position of the nucleophilic water is expected to be compensated for by stabilization of the γ-phosphate group, which lowers the energy barrier in the intermediate state. As a result, deep-sea fish and non-deep-sea fish actins have the same rate of hydrolysis at low pressure. Our results also suggest that H161 of subdomain 3, situated near the nucleophilic water, might affect conformational changes, which is consistent with preceding work (14).

The residue differences may also affect polymerization of G-actin into F-actin. Because the change in the propeller angle is related to subdomain rearrangement upon polymerization, residue K137, which is located near the hinge of the propeller motion in deep-sea fish actin, should also affect polymerization. The effect of the amino acid substitution on F-actin polymerization should be also investigated by MD simulations in the future.

Chapter 3

3 Conformational analysis for elongation of actin filament

3-1 Introduction

In this chapter, I examined the dynamics of proteins using MD simulations. In particular, I analyzed the conformational changes between flat and twisted forms of actin. These conformational changes are suggested to play an important role in the polymerization. Actin hydrolyzes ATP into ADP through the elongation of F-actin. In addition to actin hydrolysis, myosin head moves along F-actin as guide rails in sarcomeres using ATP hydrolysis. Actin can generate force for cellular motility without myosin in non-muscle cells. For example, when the rates of the elongation are equal between the pointed and barbed ends, the length of the actin filament does not change, the position of F-actin moves, which called treadmilling, as mentioned in General Introduction. G-actin takes a twisted conformation in regard to the propeller angle (~20°) whereas F-actin is relatively flat (~5°). The propeller angle difference is a main structure difference in monomer between G- and F-actin. In general, G-actin binds to ATP and F-actin binds to ADP. In the G-actin bound to ADP, the propeller angle in the crystal structure of uncomplexed actin (PDB entry 1J6Z, resolution 1.54 Å (32)) was 24.3° excluding the missing residues (residues 1-4 and 373-375). In addition to the uncomplexed structure, two ADP-actin structures complex with β-thymosin domain of drosophila ciboulot of two different length (residues 2-58, PDB entry 3U9Z, resolution 2.09 Å (117); residues 10-34, PDB entry 1SQK, resolution 2.50 Å (117)) were solved by X-ray crystallography. The propeller angle of the former actin was 19.5° excluding the missing residues (residues 1-5, 43-51, and 372-375) whereas the latter was 15.5° excluding the missing residues (residues 1-5, 43-49, and 372-375). Both structures are closer to the twisted form observed for G-actin rather than the flat form of F-actin.

I applied parallel cascade selection molecular dynamics (PaCS-MD) (118) to investigate the conformational changes of action monomers. In the original PaCS-MD, to generate conformational transition pathway between reaction and product states under the

condition that a set of reactant and product structures is known, the cycle of short multiple independent MD simulations and the selection of the structures close to the product structure for the next cycle are carried out iteratively until the simulated structures are sufficiently close to the product structure. The PaCS-MD approach can enhance the protein conformational changes efficiently.

The free energy profile according to the propeller angle was also calculated by Markov state model (MSM) (119-126). MSM is used widely to obtain kinetics information on transitions among conformational states such as protein folding processes (127-129), and the results from MSM can also be compared with experimental data, e.g., temperature-jump and electron spin resonance (127, 128, 130, 131). To use MSM, I need to partition the conformational space covered with MD trajectories into a discrete set of states appropriately, and compute a transition matrix describing the transition probabilities between the states from the simulation data. If the system was Markovian, the rates of interconversion between the states and the stationary probability distribution could be obtained by analyzing the transition matrix. Based on the free energy profile calculated from the stationary probability distribution, the kinetics of actin propeller angle change is discussed.

The elongation mechanism in the terminal regions of F-actin is still unclear. Many actin structures have been registered in PDB. Most of the monomeric and oligomer actin structures were solved as complexes with other protein molecules by the X-ray crystallography and NMR. The F-actin structures determined by the cryo electron microscopy and X-ray fiber diffraction were relatively low-resolution compared to G-actin. Since the images were generated by the averaging of the non-terminal region of F-actin, the structure of the terminal regions has not been solved by the cryo electron microscopy and X-ray fiber diffraction. Previous studies of G- and F-actin showed that stable structures of these actins and their properties; however, the elongation mechanism of F-actin is not clear at the molecular level. Therefore, I analyzed the ATP- and ADP-actins using MD simulations. Conformational and free energy changes along propeller angle rotation were investigated by PaCS-MD and MSM.

3-2 Methods

3-2-1 *Structure modeling of actins*

The actin molecules studied in this work are listed in Table 24. I performed structure modeling of rabbit G-actin bound to ATP (See chapter 2-2-1 for the detailed procedure). This model starting from the X-ray crystal structure is called a twisted-ATP. Actin protomer bound to ADP model built using F-actin structure from cryo electron microscopy (PDB entry 3MFP, resolution 6.60 Å) (16) is termed flat-ADP. I also constructed an actin pentamer model using the cryo electron microscopy. Hydrogen-atom coordinates were added using the VMD software package (91). The N-terminus was modified with acetyl-aspartate, and 3-MeH was adopted for residue 73. All actin models examined were solvated in a periodic boundary box with 139 mM K^+, 12 mM Na^+, and a water layer of at least 15 Å. The initial box size of all G-actin and pentamer models was 90 × 90 × 90 and 127 × 127 × 216 $Å^3$, respectively. All the simulation systems were neutralized by Cl^- ions.

3-2-2 *Molecular dynamics simulations*

The MD condition of this chapter was the same as that in chapter 2-2-2. After the 40,000-steps energy minimization using the conjugate gradient method was run, all actin models were gradually heated from 3 to 300 K by 3 K every 10 ps. After heating the systems, I performed MD simulations with the harmonic position constraints for the equilibration. The backbone heavy atoms, nucleotide heavy atoms, and Mg^{2+} were chosen for the constraint atoms. Since twisted-ATP was made using the X-ray crystal structures and suggested to be stable, the force constraint was started from 1 kcal/mol/$Å^2$ and gradually decreased by 0.01 kcal/mol/$Å^2$ every 10 ps. In contrast, the force constant of flat-ADP was started from 2 kcal/mol/$Å^2$ and slowly decreased by 0.01 kcal/mol/$Å^2$ every 100 ps. I performed MD simulations the twisted-ATP, flat-ADP and pentamer without the constraints for 100 ns (50-ns equilibration and 50-ns sampling).

Table 24. Overview of the simulation systems.

MD	Number of actin protomers	Nucleotide	Simulation box (Å^3)	System size (atoms)	Simulation time (ns)
Twisted-ATP	1	ATP	90 × 90 × 90	68,476	100
Flat-ADP	1	ADP	90 × 90× 90	68,473	100
Pentamer (detecting) [a]	5	ADP	127 × 127 × 216	336,613	2 or 10
Pentamer (sampling) [b]	5	ADP	127 × 127 × 216	336,574	100

[a]Actin pentamer for detecting the low-affinity ion binding sites. [b]Actin pentamer for equilibrium and sampling simulations.

3-2-3 *Detecting the low-affinity binding sites*

Almost all X-ray crystal structures of G-actin have a divalent cation at the active site (high-affinity cation binding site) whereas some X-ray crystal structures also have ions other than the active site (low-affinity ion binding site). I detected low-affinity ion binding site for simulations as follows. Four low-affinity ion binding sites (S3, S5, S7, and S9) were chosen for further analysis, these binding sites satisfied the condition that at least two different structure share the site; two acidic amino acid side-chain bound to the ion or the site was the inter protomers (Table 25). In addition to the four cation binding sites, I speculated about the low-affinity ion binding sites in the F-actin model. The expected low-affinity cation binding sites were selected if the distance between two acidic amino acid side-chain nitrogen atoms was 4-12 Å and there was no interrupted atom between the two oxygen atoms. Similarly, the expected low-affinity anion binding sites were selected if the distance between two basic amino acid side-chain nitrogen atoms was 4-12 Å and there was no interrupted atom between the two nitrogen atoms. The sites of S3, S5, S7, and S9 also satisfied with the criteria. As a result, I obtained three cations (S10-12) and three anions (S13-15) expected binding sites. To eliminate inappropriate low-affinity ion binding sites, I performed MD simulations using actin pentamer in solution using Mg^{2+} and Cl^- at the low-affinity ion binding sites. The E270 orientation of the side-chain at S3 was different between X-ray crystal structure and cryo electron microscopy models by ~170°. Therefore, the two different side-chain models were used for simulations. In the two models, one 10-ns

and four 2-ns simulations were run using constraints 10 to 0.01 kcal/mol/Å^2 gradually decreasing by 0.01 kcal/mol/Å^2 every 2 ps after 40,000-step minimization. After simulations, I judged whether the ion was still bound or not. Finally, I determined that S7 and S9-11 were chosen for simulations and the orientation of E270 side-chain was used in the cryo electron microscopy model (Table 26).

Table 25. Detailed low-affinity ion binding site positions.

Site	Ion	Inter/intra protomer	Residue:Atom name
S1	Cation	Intra	V30:O and N
S2	Cation	Intra	E125:O, N128:OD1
S3	Cation	Inter	G268:O, E270:OE1 and OE2
S4	Cation	Inter	E205:OE2
S5	Cation	Intra	D222:OD1 and OD2, E224:OE1
S6	Cation	Intra	Q263:OE1, S265:OG
S7	Cation	Inter	D286:OD1 and OD2, ASP288:OD2
S8	Cation	Intra	S323:O
S9	Cation	Intra	Q354:O, W356:O, E361:OE1 and OE2
S10	Cation	Inter	E49:OE1, E167:OE1 and OE2
S11	Cation	Inter	E241:OE1 and OE2, T324:O
S12	Cation	Inter	D292:OD2, S60:O, R62:O
S13	Anion	Inter	K113:HZ1, R256:HH11
S14	Anion	Inter	R177:HH11, K191:HZ3
S15	Anion	Inter	Q246:HE22, R290:HH12

The sites S1-9 are the same as Table 5 except for the ion column in General introduction. S10-15 are the expected low-affinity ion binding sites. The atom names are corresponded to the PDB file format.

Table 26. Detecting the low-affinity ion binding sites using cryo electron microscopy and X-ray crystal structure models for E270 side-chain.

No.	E270	Time (ns)	S3	S5	S7	S9	S10	S11	S12	S13	S14	S15
1	EM	10	0	0	1	1	1	1	0	0	0	0
2	EM	2	0	0	1	0	0	1	0	0	0	0
3	EM	2	0	0	1	0	0	1	0	0	0	0
4	EM	2	0	0	1	1	1	1	0	0	0	0
5	EM	2	0	0	1	1	1	1	0	0	0	0
6	X-ray	10	0	0	1	1	1	1	0	0	0	0
7	X-ray	2	0	0	1	1	1	1	0	0	0	0
8	X-ray	2	0	0	1	1	1	0	0	0	0	0
9	X-ray	2	0	0	1	1	1	1	0	0	0	0
10	X-ray	2	0	0	1	1	1	1	0	0	0	0

One and zero denote binding and unbinding between the ion and residue, respectively. EM and X-ray represent the side-chain E270 of the cryo electron microscopy and X-ray crystal structure model, respectively.

3-2-4 *Parallel cascade selection molecular dynamics*

PaCS-MD was proposed as a conformational sampling method to induce conformational change toward a target direction efficiently (118). In PaCS-MD, *M* independent MD simulations are performed for the each cycle of PaCS-MD. All the snapshots are ranked-ordered in terms of the propeller angle and top *M* snapshots are selected as the initial structures for the next cycle. In this study, I used $M = 10$ for PaCS-MD. In the beginning of the each cycle, initial velocities are randomly generated to reproduce the Maxwell-Boltzmann distribution. The cycle was terminated when the propeller angle was reached to the target angle or the number of the cycle limit. In this study, the number of the cycle limit was set to be 30. Each simulation was performed for 100 ps and the coordinate was stored every 1 ps, i.e., 100 snapshots stored except for the initial structure. I can calculate the free energy profile along the propeller angle using trajectories generated from PaCS-MD (see below).

3-2-5 *Markov state model*

I used the EMMA software for the MSM construction (132). The construction process used here had three steps as follows,

1. Cluster the snapshot data into states by regular spatial clustering with the minimum cutoff angle, 1.0°, which guarantees that cluster centers are chosen to be approximately equally separated in the conformational space with respect to the angle metric used, and then assign all conformations to the states.

2. Count transitions between all pairs of states in the data set. The result is saved in a count matrix $\mathbf{C}(\tau)$ which depends on a lag time τ. The elements of a count matrix, $c_{ij}(\tau)$ contain the number of transition between state i at time t and state j at time $t+\tau$, for all times t.

3. Estimate the transition probability matrix $\mathbf{T}(\tau)$ calculated from the count matrix $\mathbf{C}(\tau)$.

Eigenvalue-eigenvector pairs are calculated from a transition matrix $\mathbf{T}(\tau)$,

$$\mathbf{T}(\tau)\, p_i = \lambda_i(\tau)\, p_i \tag{13}$$

where p_i and λ_i represent the i-th left eigenvector and eigenvalue, respectively. The first eigenvector of $\mathbf{T}(\tau)$ with eigenvalue 1 corresponds to the stationary probability distribution. It has been showed that the eigenvalues of $\mathbf{T}(\tau)$ should follow an exponential decay in τ (119, 121, 123, 130, 133). The time scale of this decay called implied time scale is expressed as follows,

$$t_i(\tau) = -\frac{\tau}{\ln\left(\left|\lambda_i(\tau)\right|\right)}. \tag{14}$$

The implied time scales are experimentally observed as relaxation time scales (130), and also indicate whether or not the model is approximately Markovian (121, 133). If the model with a certain lag time is Markovian, the implied time scales should remain constant for any longer lag time. Therefore, I estimated the time scales as discussed at chapter 3-3-3. After obtaining the stationary probabilities π, which is the eigenvector of $\mathbf{T}(\tau)$ with eigenvalue 1,

the free energies, G_i , can be calculated as follows,

$$G_i = -k_B T \ln \pi_i + C \tag{15}$$

where k_B, T, and C represent the Boltzmann constant, absolute temperature, and a constant.

In the MSM construction for both twisted-ATP and flat-ADP, last 90-ps trajectories in PaCS-MD were employed. Totally, 270,000 frames (= 10-initial-structure × 30-cycle × 10-short-MD × 90-frame) were used in both cases. The transition matrices with decreasing and increasing propeller angles were 23 × 23 and 30 × 30, respectively.

3-3 Results and Discussion

3-3-1 *Conventional molecular dynamics*

I performed conventional molecular dynamics (CMD) simulations using twisted-ATP and flat-ADP. The RMSDs of backbone heavy atoms from initial structures in twisted-ATP and flat-ADP were 2.36 ± 0.18 and 3.05 ± 0.19 Å, respectively (Figure 14). Since twisted-ATP is modeled based on X-ray crystal structure of G-actin, it is expected to be more stable than flat-ADP. In both twisted-ATP and flat-ADP, the standard deviation of RMSD is ~0.19 Å through the simulations. Although flat-ADP was isolated from F-actin and simulated in monomer state, the nucleotide was stably bound to the actin. No significant changes were observed in terms of conformation of the nucleotide and position relative to the subdomains of actin. Twisted-ATP also bound to ATP at the active site appropriately. The initial propeller angle of twisted-ATP and flat-ADP were 13.1 and 5.2°, respectively. The G-actin was more twisted compared to F-actin structures. In last 50-ns simulations, the propeller angle of twisted-ATP and flat-ADP were 22.4 ± 1.9 and 7.9 ± 2.2°, respectively (Figure 15). Since these fluctuations of propeller angle were only ~2°, the conformational transition between flat- and twisted-form was not occurred in 100-ns simulations. Preceding study in simulations indicated that the propeller angle of twisted-ATP was ~20° (73). Current study also indicated the propeller angle of twisted-ATP corresponds to that observed the preceding study. The difference between these two average propeller angles was ~15°. This propeller angle difference is probably associated with the elongation of F-actin.

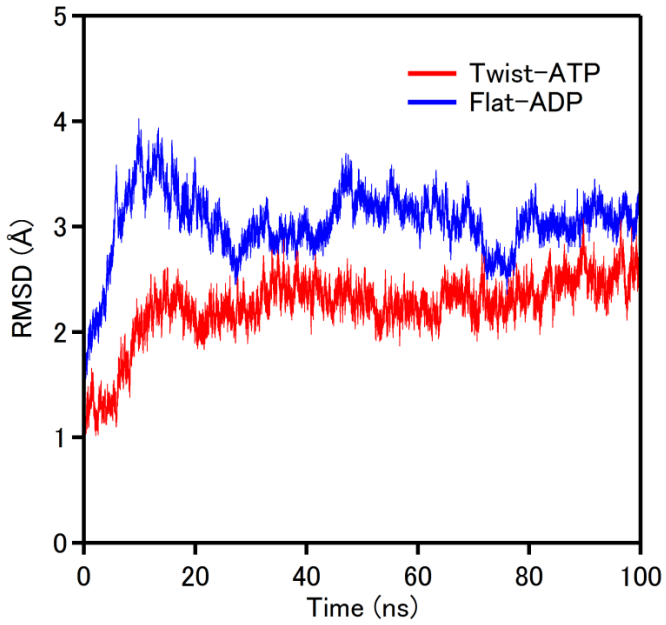

Figure 14. RMSD of backbone heavy atoms in CMD simulations of actins. Red and blue lines denote the twisted-ATP and flat-ADP, respectively.

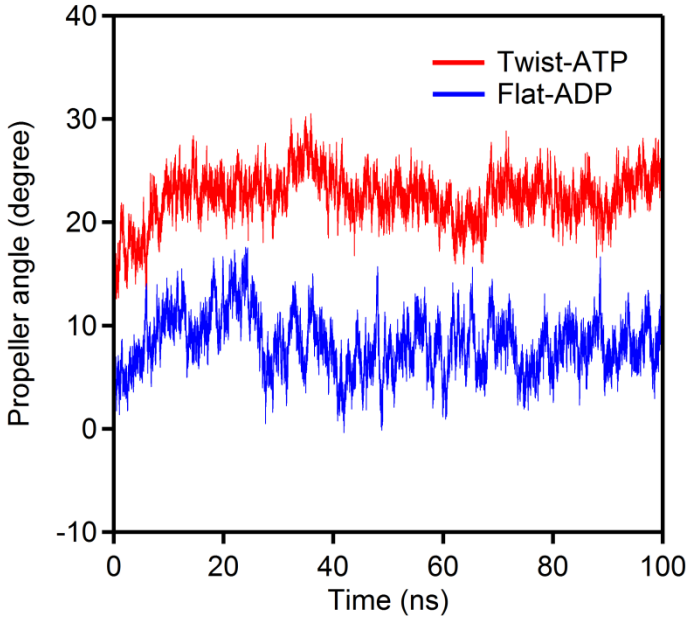

Figure 15. Propeller angles in CMD simulations of actins. Red and blue lines denote the twisted-ATP and flat-ADP, respectively.

3-3-2 *Parallel cascade selection molecular dynamics*

PaCS-MD was carried out using twisted-ATP or flat-ADP. The initial structure of twisted-ATP and flat-ADP were selected using the structures having maximum (28.8°) and minimum (0.92°) propeller angle in the last 50-ns CMD simulations, respectively. Two sets of 10 PaCS-MDs were performed for twisted-ATP and flat-ADP. The PaCS-MD cycles were terminated after 31 cycles were done in all cases. The propeller angles in PaCS-MD for twisted-ATP started with 28.8° and reached 2.1-9.6° (Figure 16). Similarly that for flat-ADP started with 0.92° and reached 26.3-39.2° (Figure 17). Therefore, PaCS-MD effectively rotated the propeller angle in both increasingly and decreasingly. Since 10 PaCS-MDs were run independently, the trajectories of the propeller angles obtained from PaCS-MD showed wide distributions in each cycle, which indicate that PaCS-MD sampled various transition passes. Twisted-ATP and flat-ADP were widely distributed in approximately 2 Å whereas the standard deviation of CMD simulations was ~0.2 Å. The RMSD of twisted-ATP reached 2.5-3.4 Å (Figure 18). Comparison to twisted-ATP, the RMSD of Flat-ADP reached greater in 3.6-5.1 Å (Figure 19). Therefore, PaCS-MD is suggested to generate sampling effectively in RMSD distributions. In CMD simulations, twisted-ATP and flat-ADP fluctuated in 15.9-28.8° and 0.92-16.7°, respectively. Therefore, PaCS-MD effectively rotated the propeller angle both increasingly and decreasingly including that actins in CMD simulations are not rotated. These rotations were generated by selection that I chose the highly-ranked structures. Some twisted-ATP and flat-ADP were trapped at small rotations. Since 10 PaCS-MDs were run independently, each simulation has variations of propeller angles through the initial to end transition. In addition to the propeller angel rotations, RMSD of backbone heavy atoms from initial structures were also analyzed. The RMSD of twisted-ATP reached 2.5-3.4 Å (Figure 18) whereas —that of flat-ADP reached 3.6-5.1 Å (Figure 19). These values are much larger than that observed in CMD simulations. Thus PaCS-MD efficiently sampled wide range of structures.

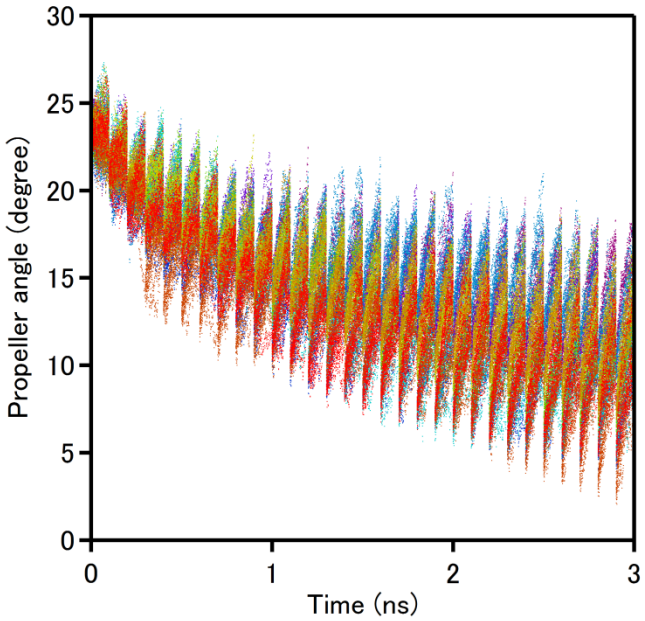

Figure 16. Propeller angle rotations of Twisted-ATP in 10 PaCS-MDs. The different colors of dots denote the individual PaCS-MDs.

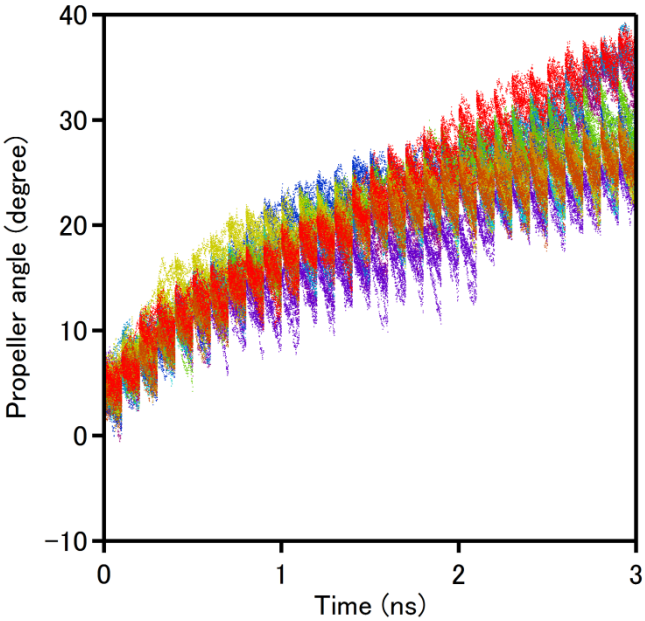

Figure 17. Propeller angle rotations of flat-ADP in 10 PaCS-MDs. The different colors of dots denote the individual PaCS-MDs.

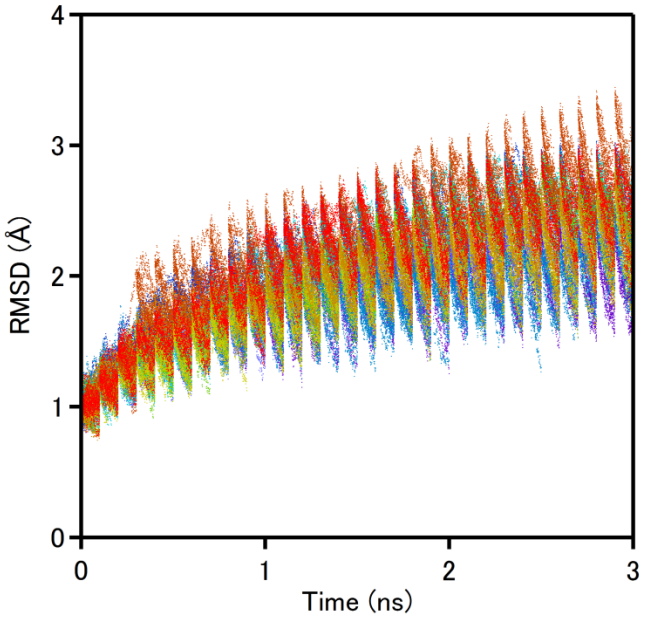

Figure 18. Twisted-ATP RMSD of backbone heavy atoms from the initial structure in 10 PaCS-MDs. The different colors of dots denote the individual PaCS-MDs.

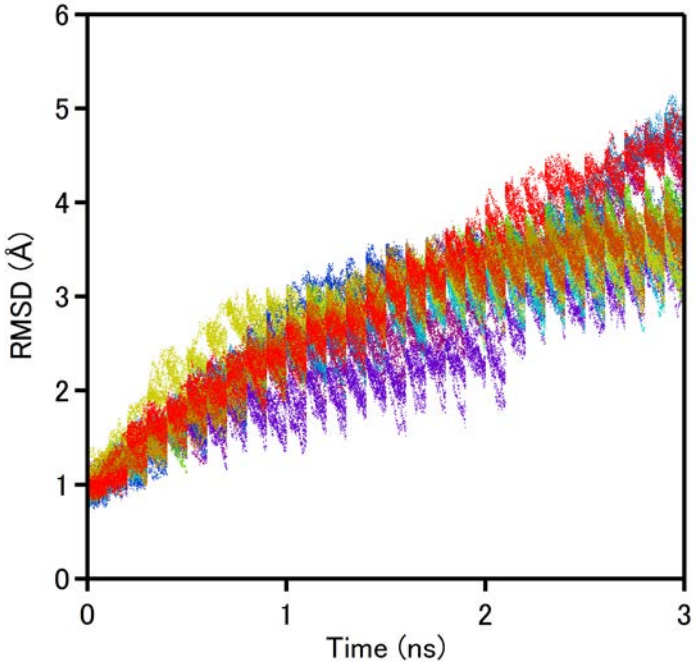

Figure 19. Flat-ADP RMSD of backbone heavy atoms from the initial structure in 10 PaCS-MDs. The different colors of dots denote the individual PaCS-MDs.

3-3-3 *Free energy analysis*

I conducted free energy analysis along the propeller angle rotations by MSM using multiple short trajectories generated from PaCS-MD. In this study 3000 short trajectories with 100 snapshots were generated. As shown in Figure 20, the temperature of each trajectory converged within a few ps. Thus trajectories after 10-ps in each cycle were used for analysis (90 snapshots from each cycle).

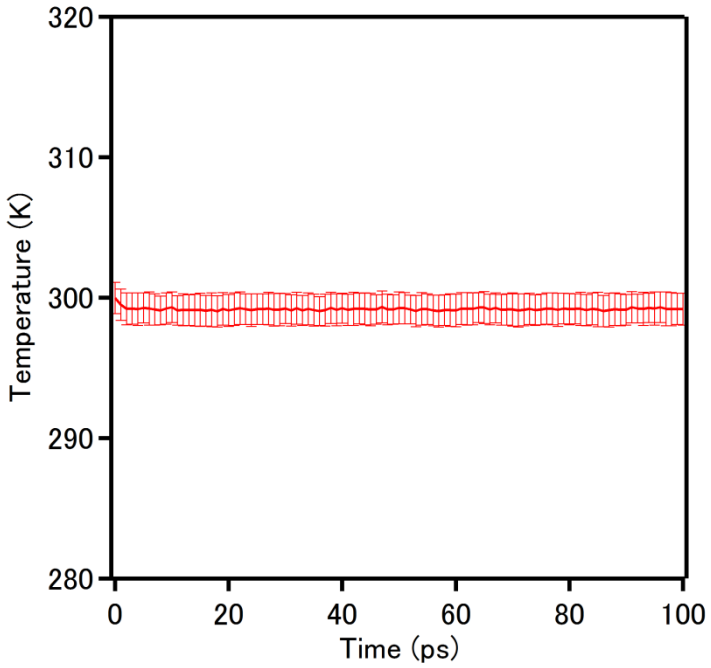

Figure 20. The convergence of temperature in PaCS-MD. Average temperatures were calculated using all PaCS-MD trajectories.

I calculated the implied time scale in twisted-ATP and flat-ADP to judge whether the model is approximately Markovian or not. Figure 21 showed the implied time scales as a function of the lag time. The three largest implied time scales were shown as examples. All the implied time scale well converged within 20 ps. Thus, I chose the lag time in 40 ps for

MSM construction. I calculated the transition matrix $\mathbf{T}(\tau)$ and obtained the stationary probabilities π. Finally, I obtained the free energy of the propeller angle rotations, ΔG, in twisted-ATP and flat-ADP using PaCS-MD trajectories. Finally I obtained the free energy profile along the propeller angle rotations, ΔG, in twisted-ATP and flat-ADP using PaCS-MD trajectories.

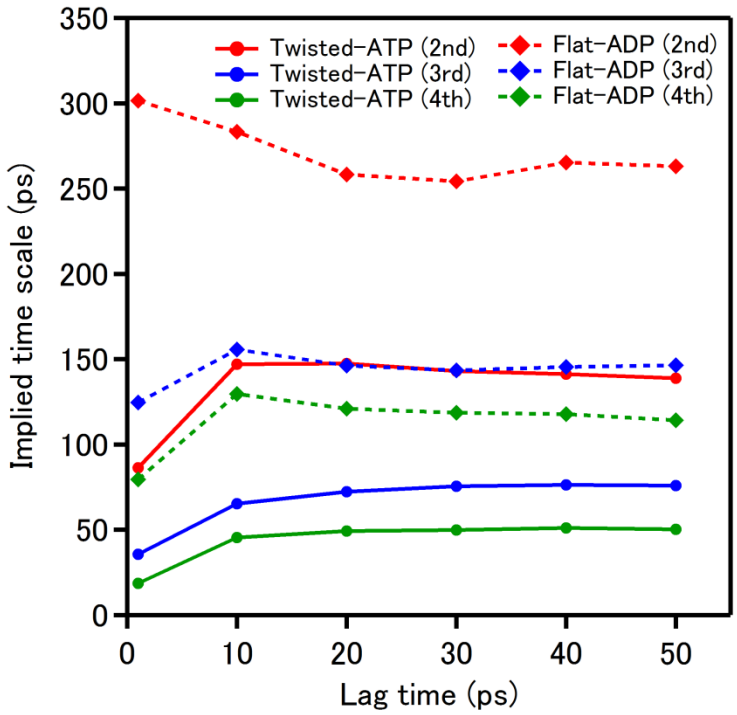

Figure 21. Implied time scale of twisted-ATP and flat-ADP. Red, blue, and green solid lines denote the implied time scale in 2nd, 3rd, and 4th eigenvalues, respectively. Similarly, red, blue, and green broken lines denote the implied time scale in 2nd, 3rd, and 4th eigenvalues, respectively.

The obtained free energy profiles along the propeller angle are shown in Figure 22. The free energy profiles obtained from CMD simulations using probability distributions of propeller angles are also shown. In PaCS-MD, the minimum of ΔG in twisted-ATP was

20-22° whereas the minimum in flat-ADP was 4°. These two energy minima were approximately corresponded to each energy minimum calculated by the CMD simulations of the probability distributions. The energies had single minimum and were similar curve between PaCS-MD and CMD simulations. These two energy minima roughly correspond to energy minima obtained from CMD simulations. The energy difference in twisted-ATP between the minimum and 4° in flat conformation was ~15 kcal/mol. This energy difference is suggested to prevent the conformational transition from twisted-ATP to flat-ADP in isolated actins. Additionally, the energy difference in flat-ADP between the minimum and 20° in twisted conformation was ~7 kcal/mol. There was also the energy barrier between flat and twisted conformation. Since the energy difference in flat-ADP was lower than that in twisted-ATP, flat-ADP would be rotatable compared to twisted-ATP. It is essential for flat-ADP to be rotated in propeller angle during the elongation. The energy difference in flat-ADP may be lower. Figure 22 showed the relative free energy of twisted-ATP and flat-ADP and the minima were adjusted to 0 kcal/mol. Therefore, I do not compare the difference between flat- and twisted-form. Actin protomers in F-actin was expected to be stabilized by the interaction among surrounded actin protomers. The conformational transition from twisted- to flat-form would be allowed only when ATP is hydrolyzed into ADP. Free energy analysis supports the fact that ATP-bound actin prefer twisted conformation whereas ADP-bound actin prefer flat conformation. The free energy differences in flat and twisted conformation for ATP- and ADP-bound actin are ~15 kcal/mol and ~7 kcal/mol, respectively. The observed energy difference in ADP-bound actin was smaller than that in ATP-bound actin. This observation might imply that conformational transition between flat- and twisted-form could occur in ADP-bound actin more easily than in ATP-bound actin because the conformational transition from twisted- to flat-form would be allowed only when ATP is hydrolyzed into ADP. The atom difference between twisted-ATP and flat-ADP was only the tail of nucleotide; however, the propeller angle rotation was critically difference. Although only the tail of nucleotide is different in ATP- and ADP-bound actin, the conformations of actins are critically different. The nucleotide at the center of actins is surrounded with all subdomains. The structural change between ATP and ADP may be sufficient to control the conformation.

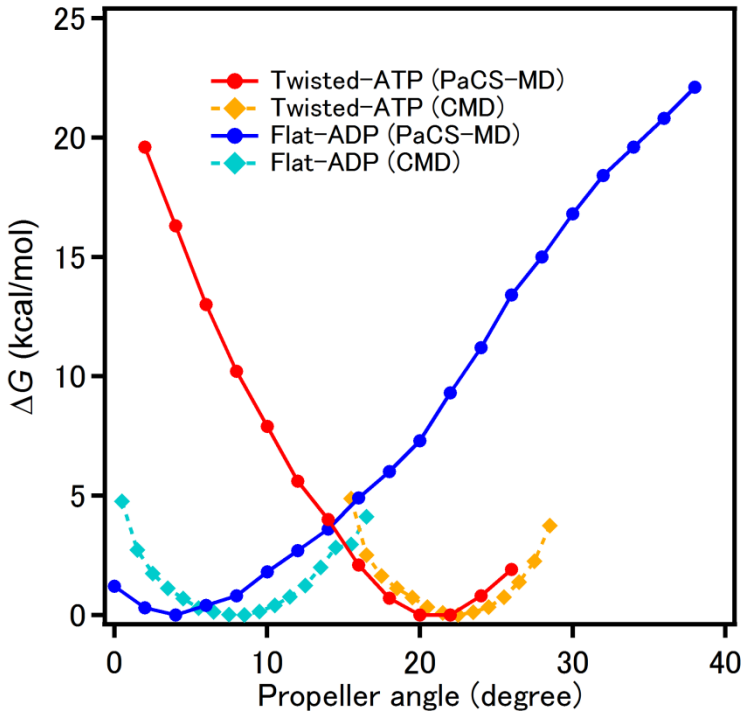

Figure 22. Free energy profiles along the propeller angle in twisted-ATP and flat-ADP. Red and blue solid lines denote the free energies in twisted-ATP and flat-ADP using PaCS-MD, respectively. Orange and cyan broken lines denote the free energies in twisted-ATP and flat-ADP using CMD simulations, respectively.

3-3-4 *Actin pentamer*

I performed MD simulation of actin pentamer as a model of F-actin. Table 27 shows the propeller angles of five actin protomers. Protomers are labeled A to E from the pointed end to the barbed end. In F-actin, actin protomers form a double helical structure and each actin protomer interacts with neighboring four protomers. Thus the actin pentamer can be considered as a minimal model of F-actin and only actin protomer at position C has four neighboring actin. The obtained propeller angles are larger than the 5.2°, which was found in the cryo electron microscopy structure. The propeller angle obtained from MD simulation of flat-ADP is also larger than 5.2°. Thus obtained large propeller angles may be originated from lack of interaction which should exist in F-actin. The propeller angle at

position C is smaller than those at position A, B, and D, which implies that interaction with neighboring actin protomers stabilizes the less twisted-form.

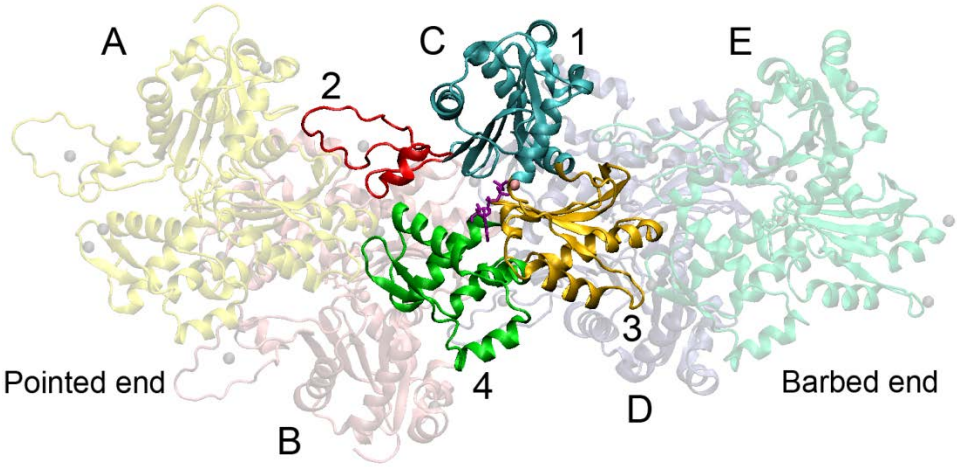

Figure 23. Structure of actin pentamer. The numbers and A to E denote the subdomain numbers and position labels, respectively. The pink sphere and purple licorice model represent Mg^{2+} and ADP, respectively.

Table 27. Propeller angles of actin pentamer.

Position label	Propeller angle (°)
A	9.6 ± 2.6
B	9.3 ± 2.1
C	7.6 ± 1.7
D	9.3 ± 1.4
E	4.5 ± 2.5

3-4 Conclusion

I analyzed conformational difference between G- and F-actin using CMD simulations of actin monomer and pentamer, and PaCS-MD of monomer. The analysis mainly focuses on the propeller angle which defined using four subdomains of actin. From the CMD simulation, I found that the stable structure of twisted-ATP is more twisted than that of flat-ADP. The propeller angle difference of twisted-ATP and flat-ADP is roughly 15°. The propeller angle of actin pentamer obtained from CMD simulation is similar with that found in the simulation of flat-ADP. The conformational transition from flat-form to twisted-form or twisted-form to flat-form was not observed in 100-ns CMD simulations. The combination of PaCS-MD and MSM was suggested to analyze conformational transition between flat- and twisted-form. PaCS-MD enhanced sampling and successfully generated wide range of propeller angles. Using multiple short simulations of PaCS-MD and MSM, free energy profiles along the propeller angle were calculated. The free energy analysis suggests that ATP-bound actin prefer the twisted conformation whereas ADP-bound actin prefers the flat conformation. The free energy differences in the flat and the twisted conformation for ATP-bound actin is larger than that for ADP-bound actin might imply that conformational transition between the flat- and the twisted-forms could occur in ADP-bound actin more easily than in ATP-bound actin because the conformational transition from the twisted- to the flat-form would be allowed only when ATP is hydrolyzed into ADP. Through these analyses I concluded that the procedure using PaCS-MD with MSM is powerful method and can be applied to various biomolecular phenomena.

Chapter 4

4. Conclusions and Perspective

Through this study, I mainly elucidated two important characteristics of actins. The first one is pressure tolerance of deep-sea fish which is related to the stability of actin. The second one is conformational difference between the flat- and the twisted-forms of actin which is related to the mechanism of elongation in F-actin.

I showed that deep-sea fish actin is more stable to high pressure. The substituted residues of Q137K/V54A and Q137K/L67P affected the actin salt bridge pattern and free energy. Deep-sea fish actins had the salt bridge between K137 and ATP, which is expected to prevent ligand dissociation at high pressure. Additionally, the energy differences between high and low pressure in deep-sea fish actins were lower than other species. Therefore, I elucidated that two substituted amino acid residues affect the stabilization to pressure tolerance in deep-sea fish actins. This effect is suggested to be induced by the subunit arrangement through the salt bridges.

I analyzed the conformational difference between twisted-ATP and flat-ADP. I did not observe the conformational transition between the flat- and the twisted-forms in 100-ns CMD simulations. In contrast, PaCS-MD enhanced sampling and successfully generated wide range of propeller angles. I proposed the method that combines PaCS-MD with MSM which enables us to conduct free energy analysis of various systems. The free energy analysis suggests that ATP-bound actin prefer twisted conformation whereas ADP-bound actin prefers flat conformation.

This work would contribute to the analysis of actin stability to high pressure and elongation mechanism of F-actin. In stability study in deep-sea fish actins, I investigated only G-actins for analysis of the pressure tolerance. The mechanism of the tolerance will be more clarified if MD simulations of F-actin from deep-sea fish are performed. Additionally, I will perform PaCS-MD and CMD simulations using twisted-ATP and flat-ADP in F-actin to investigate the detailed mechanism of elongation. The method which combines PaCS-MD and MSM proposed in this study would be useful to analyze biomolecular systems including F-actin.

Acknowledgements

I genuinely appreciate cooperation of Prof. Akio Kitao for this study. I am very grateful for his courteous supervision and generous guidance. Through the great discussions with him, I obtained the wonderful advices and strategies. Additionally, his encouragements helped me to improve my study. I would like to thank all Kitao laboratory members including past members. In particular, I would like to thank Dr. Kazuhiro Takemura daily support and great discussions. I also thank his technical support about the PC clusters and super computers. I am grateful for his advice about the programming and data analysis. I appreciate Dr. Yasutaka Nishihara for both the lavish discussions and technical support. I solved the trouble of the PC clusters and servers through his exploration. His wonderful advice inspires me. I thank Dr. Yu Yamamori about daily support and exciting discussions. I am grateful to secretary, Ms. Hidemi Shibayama and Ms. Chisato Iwasa, about a lot of daily supports. I also thank the past members. In particular, I wish to thank Dr. Shun Sakuraba for his sound advice. Additionally, I appreciate Dr. Ryuhei Harada about exciting discussions and MD techniques, especially PaCS-MD. I am thankful to Dr. Takami Morita for actin experimental data. I am much obliged to Prof. Keiich Namba about the cryo electron microscopy structure of F-actin and great discussions. I appreciate Dr. Takashi Fujii about the cryo electron microscopy structure of F-actin. Finally, I appreciate my family continuous supports very much.

References

1. Salmikangas, P., P. F. M. van der Ven, M. Lalowski, A. Taivainen, F. Zhao, H. Suila, R. Schröder, P. Lappalainen, D. O. Fürst, and O. Carpén. 2003. Myotilin, the limb-girdle muscular dystrophy 1A (LGMD1A) protein, cross-links actin filaments and controls sarcomere assembly. *Hum. Mol. Genet.* 12:189-203.
2. Masuda, T. 2008. A possible mechanism for determining the directionality of myosin molecular motors. *Biosystems* 93:172-180.
3. Herman, I. M. 1993. Actin isoforms. *Curr. Opin. Cell Biol.* 5:48-55.
4. Schwob, E., and R. P. Martin. 1992. New yeast actin-like gene required late in the cell cycle. *Nature* 355:179-182.
5. Vandekerckhove, J., and K. Weber. 1978. At least six different actins are expressed in a higher mammal: An analysis based on the amino acid sequence of the amino-terminal tryptic peptide. *J. Mol. Biol.* 126:783-802.
6. Ji, Y., G. Ferracci, A. Warley, M. Ward, K.-Y. Leung, S. Samsuddin, C. Lévêque, L. Queen, V. Reebye, P. Pal, E. Gkaliagkousi, M. Seager, and A. Ferro. 2007. β-Actin regulates platelet nitric oxide synthase 3 activity through interaction with heat shock protein 90. *Proc. Natl. Acad. Sci. USA.* 104:8839-8844.
7. Pollard, T. D., and G. G. Borisy. 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112:453-465.
8. Kabsch, W., H. G. Mannherz, D. Suck, E. F. Pai, and K. C. Holmes. 1990. Atomic-structure of the actin:DNase-I complex. *Nature* 347:37-44.
9. Kalhor, H. R., A. Niewmierzycka, K. F. Faull, X. Y. Yao, S. Grade, S. Clarke, and P. A. Rubenstein. 1999. A highly conserved 3-methylhistidine modification is absent in yeast actin. *Arch. Biochem. Biophys.* 370:105-111.
10. Hirata, K., S. Muraoka, K. Suenaga, T. Kuroda, K. Kato, H. Tanaka, M. Yamamoto, M. Takata, K. Yamada, and H. Kigoshi. 2006. Structure basis for antitumor effect of aplyronine A. *J. Mol. Biol.* 356:945-954.

11. Ducka, A. M., P. Joel, G. M. Popowicz, K. M. Trybus, M. Schleicher, A. A. Noegel, R. Huber, T. A. Holak, and T. Sitar. 2010. Structures of actin-bound Wiskott-Aldrich syndrome protein homology 2 (WH2) domains of Spire and the implication for filament nucleation. *Proc. Natl. Acad. Sci. USA.* 107:11757-11762.

12. Morita, T. 2003. Structure-based analysis of high pressure adaptation of α-actin. *J. Biol. Chem.* 278:28060-28066.

13. Oda, T., M. Iwasa, T. Aihara, Y. Maeda, and A. Narita. 2009. The nature of the globular- to fibrous-actin transition. *Nature* 457:441-445.

14. Murakami, K., T. Yasunaga, T. Q. P. Noguchi, Y. Gomibuchi, K. X. Ngo, T. Q. P. Uyeda, and T. Wakabayashi. 2010. Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release. *Cell* 143:275-287.

15. Galkin, V. E., A. Orlova, G. F. Schröder, and E. H. Egelman. 2010. Structural polymorphism in F-actin. *Nat. Struct. Mol. Biol.* 17:1318-1323.

16. Fujii, T., A. H. Iwane, T. Yanagida, and K. Namba. 2010. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* 467:724-729.

17. Carlier, M. F., D. Pantaloni, and E. D. Korn. 1986. Fluorescence measurements of the binding of cations to high-affinity and low-affinity sites on ATP-G-actin. *J. Biol. Chem.* 261:10778-10784.

18. Blain, J. C., Y.-F. Mok, J. Kubanek, and J. S. Allingham. 2010. Two molecules of lobophorolide cooperate to stabilize an actin dimer using both their "ring" and "tail" region. *Chem. Biol.* 17:802-807.

19. Hagelueken, G., S. C. Albrecht, H. Steinmetz, R. Jansen, D. W. Heinz, M. Kalesse, and W.-D. Schubert. 2009. The absolute configuration of rhizopodin and its inhibition of actin polymerization by dimerization. *Angewandte Chemie-Int. Edit.* 48:595-598.

20. Sawaya, M. R., D. S. Kudryashov, I. Pashkov, H. Adisetiyo, E. Reisler, and T. O. Yeates. 2008. Multiple crystal structures of actin dimers and their implications for interactions in the actin filament. *Acta Crystallogr. Sect. D-Biolog. Crystallogr.* 64:454-465.

21. Allingham, J. S., C. O. Miles, and E. Rayment. 2007. A structural basis for regulation of actin polymerization by pectenotoxins. *J. Mol. Biol.* 371:959-970.

22. Klenchin, V. A., S. Y. Khaitlina, and I. Rayment. 2006. Crystal structure of polymerization-competent actin. *J. Mol. Biol.* 362:140-150.

23. Rizvi, S. A., V. Tereshko, A. A. Kossiakoff, and S. A. Kozmin. 2006. Structure of bistramide α-actin complex at a 1.35 Å resolution. *J. Am. Chem. Soc.* 128:3882-3883.

24. Allingham, J. S., A. Zampella, M. V. D'Auria, and I. Rayment. 2005. Structures of microfilament destabilizing toxins bound to actin provide insight into toxin design and activity. *Proc. Natl. Acad. Sci. USA.* 102:14527-14532.

25. Kudryashov, D. S., M. R. Sawaya, H. Adisetiyo, T. Norcross, G. Hegyi, E. Reisler, and T. O. Yeates. 2005. The crystal structure of a cross-linked actin dimer suggests a detailed molecular interface in F-actin. *Proc. Natl. Acad. Sci. USA.* 102:13105-13110.

26. Klenchin, V. A., R. King, J. Tanaka, G. Marriott, and I. Rayment. 2005. Structural basis of swinholide a binding to actin. *Chem. Biol.* 12:287-291.

27. Irobi, E., L. D. Burtnick, D. Urosev, K. Narayan, and R. C. Robinson. 2003. From the first to the second domain of gelsolin: A common path on the surface of actin? *FEBS Lett.* 552:86-90.

28. Graceffa, P., and R. Dominguez. 2003. Crystal structure of monomeric actin in the ATP state - Structural basis of nucleotide-dependent actin dynamics. *J. Biol. Chem.* 278:34172-34180.

29. Dawson, J. F., E. P. Sablin, J. A. Spudich, and R. J. Fletterick. 2003. Structure of an F-actin trimer disrupted by gelsolin and implications for the mechanism of severing. *J. Biol. Chem.* 278:1229-1238.

30. Bubb, M. R., L. Govindasamy, E. G. Yarmola, S. M. Vorobiev, S. C. Almo, T. Somasundaram, M. S. Chapman, M. Agbandje-McKenna, and R. McKenna. 2002. Polylysine induces an antiparallel actin dimer that nucleates filament assembly. Crystal structure at 3.5-Å resolution. *J. Biol. Chem.* 277:20999-21006.

31. Didry, D., F.-X. Cantrelle, C. Husson, P. Roblin, A. M. E. Moorthy, J. Perez, C. Le Clainche, M. Hertzog, E. Guittet, M.-F. Carlier, C. van Heijenoort, and L. Renault. 2012. How a single residue in individual β-thymosin/WH2 domains controls their functions in actin assembly. *EMBO J.* 31:1000-1013.

32. Otterbein, L. R., P. Graceffa, and R. Dominguez. 2001. The crystal structure of uncomplexed actin in the ADP state. *Science* 293:708-711.

33. Pollard, T. D., L. Blanchoin, and R. D. Mullins. 2000. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Ann. Rev. Biophys. Biomol. Struct.* 29:545-576.

34. Fowler, V. M. 1996. Regulation of actin filament length in erythrocytes and striated muscle. *Curr. Opin. Cell Biol.* 8:86-96.

35. Schutt, C. E., J. C. Myslik, M. D. Rozycki, N. C. W. Goonesekere, and U. Lindberg. 1993. The structure of crystalline profilin–β-actin. *Nature* 365:810-816.

36. Arber, S., F. A. Barbayannis, H. Hanser, C. Schneider, C. A. Stanyon, O. Bernard, and P. Caroni. 1998. Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase. *Nature* 393:805-809.

37. Carlier, M. F., F. Ressad, and D. Pantaloni. 1999. Control of actin dynamics in cell motility - role of ADF/cofilin. *J. Biol. Chem.* 274:33827-33830.

38. Theriot, J. A., and T. J. Mitchison. 1991. Actin microfilament dynamics in locomoting cells. *Nature* 352:126-131.

39. MacKerell, A. D., N. Banavali, and N. Foloppe. 2001. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56:257-265.

40. Patel, S., and C. L. Brooks. 2004. CHARMM fluctuating charge force field for proteins: I parameterization and application to bulk organic liquid simulations. *J. Comput. Chem.* 25:1-15.

41. Patel, S., A. D. Mackerell, and C. L. Brooks. 2004. CHARMM fluctuating charge force field for proteins: II protein/solvent properties from molecular dynamics simulations using a nonadditive electrostatic model. *J. Comput. Chem.* 25:1504-1514.

42. Wang, Z. X., and Y. Duan. 2004. Solvation effects on alanine dipeptide: a MP2/cc-pVTZ//MP2/6-31G** study of (Φ, Ψ) energy maps and conformers in the gas phase, ether, and water. *J. Comput. Chem.* 25:1699-1716.

43. Cornell, W. D., P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. 1996. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules [*J. Am. Chem. Soc.* 1995, 117, 5179-5197]. *J. Am. Chem. Soc.* 118:2309-2310.

44. Kollman, P. A. 1996. Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. *Accoun. Chem. Res.* 29:461-469.

45. Wang, J. M., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049-1074.

46. Hornak, V., R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling. 2006. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins Struct. Funct. Bioinform.* 65:712-725.

47. Lindorff-Larsen, K., S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. shaw. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinform.* 78:1950-1958.

48. Nerenberg, P. S., C. So, A. Tripathy, and T. Head-Gordon. 2011. Evaluation and improvement of the AMBER ff99SB force field with an advanced water model. *Biophys. J.* 100:311-311.

49. Duan, Y., C. Wu, S. Chowdhury, M. C. Lee, G. M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. M. Wang, and P. Kollman. 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* 24:1999-2012.

50. Kaminski, G. A., R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* 105:6474-6487.

51. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781-1802.

52. Verlet, L. 1967. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* 159:98-103.

53. Swope, W. C., H. C. Andersen, P. H. Berens, and K. R. Wilson. 1982. A computer-simulation method for the calculation of equilibrium-constants for the formation of physical clusters of molecules - Application to small water clusters. *J. Chem. Phys.* 76:637-649.

54. Tuckerman, M., B. J. Berne, and G. J. Martyna. 1992. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* 97:1990-2001.

55. Gear, C. W. 1966. Numerical integration in ordinary differential equations operating under illiac 2 time-sharing system. *Commun. ACM* 9:475.

56. Gear, C. W. 1967. Numerical integration of ordinary differential equations. *Math. Comput.* 21:146.

57. Pastor, R. W., B. R. Brooks, and A. Szabo. 1988. An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol. Phys.* 65:1409-1419.

58. Nosé, S. 1991. Constant temperature molecular-dynamics methods. *Prog. Theor. Phys. Supp.* :1-46.

59. Martyna, G. J., M. L. Klein, and M. Tuckerman. 1992. Nosé-Hoover chains - The canonical ensemble via continuous dynamics. *J. Chem. Phys.* 97:2635-2643.

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

61. Mináry, P., M. E. Tuckerman, K. A. Pihakari, and G. J. Martyna. 2002. A new reciprocal space based treatment of long range interactions on surfaces. *J. Chem. Phys.* 116:5351-5362.

62. Feller, S. E., Y. H. Zhang, R. W. Pastor, and B. R. Brooks. 1995. Constant pressure molecular dynamics simulations: The Langevin piston method. *J. Chem. Phys.* 103:4613-4621.

63. Parrinello, M., and A. Rahman. 1980. Crystal structure and pair potentials: A molecular-dynamics study. *Phys. Rev. Lett.* 45:1196-1199.

64. Andersen, H. C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* 72:2384-2393.

65. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: An *N*·log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089-10092.

66. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327-341.

67. Miyamoto, S., and P. A. Kollman. 1992. SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13:952-962.

68. McCammon, J. A., B. R. Gelin, and M. Karplus. 1977. Dynamics of folded proteins. *Nature* 267:585-590.

69. Shaw, D. E., P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. 2010. Atomic-level characterization of the structural dynamics of proteins. *Science* 330:341-346.

70. Dalhaimer, P., T. D. Pollard, and B. J. Nolen. 2008. Nucleotide-mediated conformational changes of monomeric actin and Arp3 studied by molecular dynamics simulations. *J. Mol. Biol.* 376:166-183.

71. Jain, I., D. Lacoste, D. Panda, and R. Padinhateer. 2012. History-dependent depolymerization of actin filaments. *Biochemistry (Mosc).* 51:7580-7587.

72. Erlenkaemper, C., D. Johann, and K. Kruse. 2012. Impact of motor molecules on the dynamics of treadmilling filaments. *Phys. Rev. E* 86:051906.

73. Splettstoesser, T., K. C. Holmes, F. Noé, and J. C. Smith. 2011. Structural modeling and molecular dynamics simulation of the actin filament. *Proteins Struct. Funct. Bioinform.* 79:2033-2043.

74. Case, D. A., T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668-1688.

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

76. Hochachka, P. W., and G. N. Somero. 1984. Biochemical Adaptation. Princeton University Press, Princeton.

77. Morita, T. 2000. Amino acid sequences of α-skeletal muscle actin isoforms in two species of rattail fish, *Coryphaenoides acrolepis* and *Coryphaenoides cinereus*. *Fisher. Sci.* 66:1150-1157.

78. Vorobiev, S., B. Strokopytov, D. G. Drubin, C. Frieden, S. Ono, J. Condeelis, P. A. Rubenstein, and S. C. Almo. 2003. The structure of nonvertebrate actin: Implications for the ATP hydrolytic mechanism. *Proc. Natl. Acad. Sci. USA.* 100:5760-5765.

79. Iwasa, M., K. Maeda, A. Narita, Y. Maéda, and T. Oda. 2008. Dual roles of Gln^{137} of actin revealed by recombinant human cardiac muscle α-actin mutants. *J. Biol. Chem.* 283:21045-21053.

80. Ikeuchi, Y., A. Suzuki, T. Oota, K. Hagiwara, R. Tatsumi, T. Ito, and C. Balny. 2002. Fluorescence study of the high pressure-induced denaturation of skeletal muscle actin. *Eur. J. Biochem.* 269:364-371.

81. Ikkai, T., and T. Ooi. 1966. The effects of pressure on F-G transformation of actin. *Biochemistry (Mosc).* 5:1551-1560.

82. Morita, T. 2008. Comparative sequence analysis of myosin heavy chain proteins from congeneric shallow- and deep-living rattail fish (genus *Coryphaenoides*). *J. Exp. Biol.* 211:1362-1367.

83. Siebenaller, J. F., and G. N. Somero. 1979. Pressure-adaptive differences in the binding and catalytic properties of muscle-type (M_4) lactate dehydrogenases of shallow-living and deep-living marine fishes. *J. Comp. Phys. B* 129:295-300.

84. Siebenaller, J. F. 1984. Pressure-adaptive differences in NAD-dependent dehydrogenases of congeneric marine fishes living at different depths. *J. Comp. Phys. B* 154:443-448.

85. Meinhold, L., J. C. Smith, A. Kitao, and A. H. Zewail. 2007. Picosecond fluctuating protein energy landscape mapped by pressure-temperature molecular dynamics simulation. *Proc. Natl. Acad. Sci. USA.* 104:17261-17265.

86. Imai, T., and Y. Sugita. 2010. Dynamic correlation between pressure-induced protein structural transition and water penetration. *J. Phys. Chem. B* 114:2281-2286.

87. Collins, M. D., G. Hummer, M. L. Quillin, B. W. Matthews, and S. M. Gruner. 2005. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. *Proc. Natl. Acad. Sci. USA.* 102:16668-16671.

88. Grigera, J. R., and A. N. McCarthy. 2010. The behavior of the hydrophobic effect under pressure and protein denaturation. *Biophys. J.* 98:1626-1631.

89. Sarupria, S., T. Ghosh, A. E. García, and S. Garde. 2010. Studying pressure denaturation of a protein by molecular dynamics simulations. *Proteins Struct. Funct. Bioinform.* 78:1641-1651.

90. Morton, W. M., K. R. Ayscough, and P. J. McLaughlin. 2000. Latrunculin alters the actin-monomer subunit interface to prevent polymerization. *Nat. Cell Biol.* 2:376-378.

91. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: Visual molecular dynamics. *J. Mol. Graph.* 14:33-38.

92. Choe, H., L. D. Burtnick, M. Mejillano, H. L. Yin, R. C. Robinson, and S. Choe. 2002. The calcium activation of gelsolin: insights from the 3 Å structure of the G4-G6/actin complex. *J. Mol. Biol.* 324:691-702.

93. Reutzel, R., C. Yoshioka, L. Govindasamy, E. G. Yarmola, M. Agbandje-McKenna, M. R. Bubb, and R. McKenna. 2004. Actin crystal dynamics: structural implications for F-actin nucleation, polymerization, and branching mediated by the anti-parallel dimer. *J. Struct. Biol.* 146:291-301.

94. Otomo, T., D. R. Tomchick, C. Otomo, S. C. Panchal, M. Machius, and M. K. Rosen. 2005. Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain. *Nature* 433:488-494.

95. Berendsen, H. J. C., J. R. Grigera, and T. P. Straatsma. 1987. The missing term in effective pair potentials. *J. Phys. Chem.* 91:6269-6271.

96. Sept, D., and J. A. McCammon. 2001. Thermodynamics and kinetics of actin filament nucleation. *Biophys. J.* 81:667-674.

97. Takemura, K., and A. Kitao. 2012. Water model tuning for improved reproduction of rotational diffusion and NMR spectral density. *J. Phys. Chem. B* 116:6279-6287.

98. Wriggers, W., and K. Schulten. 1999. Investigating a back door mechanism of actin phosphate release by steered molecular dynamics. *Proteins Struct. Funct. Genet.* 35:262-273.

99. Buša, J., S. Hayryan, C. K. Hu, J. Skřivánek, and M. C. Wu. 2010. CAVE: A package for detection and quantitative analysis of internal cavities in a system of overlapping balls: Application to proteins. *Comput. Phys. Commun.* 181:2116-2125.

100. Rashin, A. A. 1984. Buried surface area, conformational entropy, and protein stability. *Biopolymers* 23:1605-1620.

101. Gekko, K., and H. Noguchi. 1979. Compressibility of globular proteins in water at 25 ºC. *J. Phys. Chem.* 83:2706-2714.

102. Lee, M. S., and M. A. Olson. 2013. Comparison of volume and surface area nonpolar solvation free energy terms for implicit solvent simulations. *J. Chem. Phys.* 139:044119.

103. Li, L., C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, L. Wang, N. Smith, M. Petukh, and E. Alexov. 2012. DelPhi: A comprehensive suite for DelPhi software and associated resources. *BMC Biophys.* 5:9.

104. Wagner, W., and A. Pruß. 2002. The IAPWS formation 1995 for the thermodynamic properties of ordinary water substance for general and scietific use. *J. Phys. Chem. Ref. Data* 31:387-535.

105. Wagoner, J. A., and N. A. Baker. 2006. Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proc. Natl. Acad. Sci. USA.* 103:8331-8336.

106. Vargaftik, N. B., B. N. Volkov, and L. D. Voljak. 1983. International tables of the surface tension of water. *J. Phys. Chem. Ref. Data* 12:817-820.

107. McQuarrie, D. A. 2000. Statistical Mechanics. University Science Books, Sausalito.

108. Karplus, M., and J. N. Kushick. 1981. Method for estimating the configurational entropy of macromolecules. *Macromolecules* 14:325-332.

109. Kitao, A., F. Hirata, and N. Gō. 1991. The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular-dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* 158:447-472.

110. Sarvazyan, A. P. 1979. Relaxational contributions to protein compressibility from ultrasonic data. *Biopolymers* 18:3015-3024.

111. Gekko, K., and Y. Hasegawa. 1986. Compressibility-structure relationship of globular proteins. *Biochemistry (Mosc).* 25:6563-6571.

112. Fine, R. A., and F. J. Millero. 1973. Compressibility of water as a function of temperature and pressure. *J. Chem. Phys.* 59:5529-5536.

113. Freedman, H., T. Laino, and A. Curioni. 2012. Reaction dynamics of ATP hydrolysis in actin determined by *ab Initio* molecular dynamics simulations. *J. Chem. Theory Comput.* 8:3373-3383.

114. Harrison, C. B., and K. Schulten. 2012. Quantum and classical dynamics simulations of ATP hydrolysis in solution. *J. Chem. Theory Comput.* 8:2328-2335.

115. Saunders, M. G., and G. A. Voth. 2011. Water molecules in the nucleotide binding cleft of actin: effects on subunit conformation and implications for ATP hydrolysis. *J. Mol. Biol.* 413:279-291.

116. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157:105-132.

117. Hertzog, M., C. van Heijenoort, D. Didry, M. Gaudier, J. Coutant, B. Gigant, G. Didelot, T. Préat, M. Knossow, E. Guittet, and M. F. Carlier. 2004. The β-thymosin/WH2 domain: Structural basis for the switch from inhibition to promotion of actin assembly. *Cell* 117:611-623.

118. Harada, R., and A. Kitao. 2013. Parallel cascade selection molecular dynamics (PaCS-MD) to generate conformational transition pathway. *J. Chem. Phys.* 139.

119. Prinz, J.-H., H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schüette, and F. Noé. 2011. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* 134:174105.

120. Pan, A. C., and B. Roux. 2008. Building Markov state models along pathways to determine free energies and rates of transitions. *J. Chem. Phys.* 129:064107.

121. Noé, F., I. Horenko, C. Schüette, and J. C. Smith. 2007. Hierarchical analysis of conformational dynamics in biomolecules: transition networks of metastable states. *J. Chem. Phys.* 126:155102.

122. Buchete, N.-V., and G. Hummer. 2008. Coarse master equations for peptide folding dynamics. *J. Phys. Chem. B* 112:6057-6069.

123. Chodera, J. D., N. Singhal, V. S. Pande, K. A. Dill, and W. C. Swope. 2007. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J. Chem. Phys.* 126:155101.

124. Bowman, G. R., K. A. Beauchamp, G. Boxer, and V. S. Pande. 2009. Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.* 131:124101.

125. Hubner, I. A., E. J. Deeds, and E. I. Shakhnovich. 2006. Understanding ensemble protein folding at atomic detail. *Proc. Natl. Acad. Sci. USA.* 103:17747-17752.

126. Schultheis, V., T. Hirschberger, H. Carstens, and P. Tavan. 2005. Extracting Markov models of peptide conformational dynamics from simulation data. *J. Chem. Theory Comput.* 1:515-526.

127. Noé, F., C. Schüette, E. Vanden-Eijnden, L. Reich, and T. R. Weikl. 2009. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. *Proc. Natl. Acad. Sci. USA.* 106:19011-19016.

128. Bowman, G. R., V. A. Voelz, and V. S. Pande. 2011. Atomistic folding simulations of the five-helix bundle protein $\lambda_{6\text{-}85}$. *J. Am. Chem. Soc.* 133:664-667.

129. Voelz, V. A., G. R. Bowman, K. Beauchamp, and V. S. Pande. 2010. Molecular simulation of *ab Initio* protein folding for a millisecond folder NTL9(1—39). *J. Am. Chem. Soc.* 132:1526-1528.

130. Noé, F., S. Doose, I. Daidone, M. Löellmann, M. Sauer, J. D. Chodera, and J. C. Smith. 2011. Dynamical fingerprints for probing individual relaxation processes in biomolecular dynamics with simulations and kinetic experiments. *Proc. Natl. Acad. Sci. USA.* 108:4822-4827.

131. Sezert, D., J. H. Freed, and B. Roux. 2008. Using Markov models to simulate electron spin resonance spectra from molecular dynamics trajectories. *J. Phys. Chem. B* 112:11014-11027.

132. Senne, M., B. Trendelkamp-Schroer, A. S. J. S. Mey, C. Schüette, and F. Noé. 2012. EMMA: A software package for Markov model building and analysis. *J. Chem. Theory Comput.* 8:2223-2238.

133. Swope, W. C., J. W. Pitera, and F. Suits. 2004. Describing protein folding kinetics by molecular dynamics simulations. 1. Theory. *J. Phys. Chem. B* 108:6571-6581.