Structural Resolution of the Initiation of Actomyosin's Force Generating Stroke

Yang Zhenhui

for the Degree of Doctor of Philosophy

Supervisor: András Málnási-Csizmadia PhD.

Structural Biochemistry Doctoral Program,

Doctoral School in Biology

Program Leader: Prof. László Gráf DSc.

Head of the School: Prof. Anna Erdei DSc.

Eötvös Loránd University
Budapest, Hungary
2012

Acknowledgement

First of all, I would like to thank my entire family for their love and support and for believing in me until the end. I am especially grateful to my wife, Judy for always trusting and supporting me, for listening to my complaints and worries.

I would also like to thank Professor László Gráf and László Nyitray for operating such a successful Biochemistry Department, for providing many overseas students with opportunities to study, and for making my stay at ELTE unforgettable.

Mere words will probably not be enough to convey my gratitude to my supervisor, Professor András Málnási-Csizmadia. He has been extremely supportive of my studies with his great scientific insights, and inspiration. He also provides me his opening academic ideas, exploring on the scientific issues, and optimistic spirit. Also, I would like to thank him for his companionship during my stay in Hungary. I would also like to thank Professor Mihály Kovács. He was always able to shed light on the various problems encountered in my work, and I benefited greatly from his shared knowledge of the technical literature. I would also like to thank Dr. Feng Jiannan for helping me with the molecular modeling theory at the beginning of my studies in Hungary.

I would like to thank Dr. Balázs Jelinek and Dr. Zoltán Simon for taking care of most of my administrative matters, especially my stay in Hungary, and for always finding a few minutes to listen to my complaints with patience. I must express my indebtedness to my colleague Boglárka Várkuti, Anna Rauscher and Dr. Bálint Kintses, with whom I have discussed almost all aspects of this work. They patiently answered my questions pertaining to science and have contributed to my understanding of myosin motors with their valuable advice and inspiring insight.

Furthermore, I would like to thank all of the members of the lab: Dr. Máté Gyimesi, Mrs. Ágnes Peragovics, László Végner, Miklós Képíró, István Lőrincz, Mrs. Margit Smeller Vígh, Gábor Szegvári, Mrs. Ilona Szász Ozoróczyné, Dr. Vitalii Stadnyk, Gergely Zahoránszky, and Dr. Chrystyna Mayor for all of their help and advice and for the discussions regarding my research during my time in Hungary.

Summary:

In this thesis, we present *in silico* simulations in one of the major conformational changes of myosin called recovery step and structural modeling of actomyosin in the pre-power stroke state which is experimentally difficult to approach because of its low proportions in the enzymatic cycle.

Different isoforms of myosin play important roles in the muscle contraction and series of biological functions. The myosin motor function is exerted in a complex with actin by consuming ATP. The effective work of lever movement, ATPase cycle, the actin binding and detachment cycle must be harmonized. In ATP form of myosin, lever is moved from the so called "down" position to the "up" position in actin detached state (recovery step); while after ATP hydrolysis, the actin rebinds to the "up" lever state of myosin to move the lever to the "down" position on the basis of effective power stroke. Actin binding region of myosin is a spacious surface, which is divided by a deep cleft. The strength of actin binding is regulated by the movement of this cleft as myosin weakly and strongly binds to actin in the open- and closed-cleft states, respectively. The initiating key state of actomyosin is myosin just weakly attached to actin in its up-lever state. The effective pathways of the power stroke are not accessible due to the short lifetime of their intermediates.

One of the major questions was whether actin binding induces structural rearrangements in myosin. Thus, we modeled the up-lever/open-cleft actomyosin state by protein-protein docking and relaxed in a long molecular dynamics simulation. In order to discover the communication pathways in actin induced conformational rearrangement of myosin, we also modeled some myosin mutants in the same weak actin-binding state, compared to their relaxed structures, and analyzed their motional correlations to those of the wild type. Furthermore, we were also interested in whether the power stroke of myosin runs in the same conformational reaction coordinates as the reverse of the recovery step. Thus, we also simulated the conformational pathway of the recovery step.

Recently, Fisher *et al.* modeled the recovery step by the minimum energy path (MEP)

method and distinguished two phases in this reaction. They indicated a so-called seesaw movement of the myosin relay region, which is a long helical element of myosin and mechanically connects the nucleotide-binding region to the converter/lever region. They supposed that the seesaw tilt directly results in the swing of the myosin lever. Based on our results, a three-phase model of the recovery step was deduced on the basis of a free energy profile and the post-recovery state in a lower free energy is more preferred than the pre-recovery state. Formation of a hydrogen bond cluster accelerates structural transformation to overcome the activation energy barrier in the first phase. We found that the first phase of the recovery step is equivalent with the seesaw-like tilt described by Fischer, while the two others are related to bending and dynamic rearrangement of the relay region.

Recently, Holmes *et al.* modeled the strong actin-binding state by the molecular dynamics flexible fitting (MDFF) method. We modeled both the weak and the strong actin-binding states applying the protein-protein docking and molecular dynamic simulations. We revealed a novel actin-binding site in myosin named activation loop. The conserved positively charged tip (Arg520) of the activation loop interacts with four negatively charged residues in the N-terminus of actin in various binding patterns in the weak and the strong actin-binding states. Three specific myosin conformational changes induced by actin binding were observed in the weak actin-binding state: 1. the partial closure of the actin-binding cleft, 2. the further up rotation of the lever arm and 3. further closure of SW-II loop. The R520Q mutation in myosin prevented these conformational changes. The motions of the activation loop are correlated with four functional regions of myosin in the weak actin-binding state. Two communicational pathways were speculated between the actin binding regions and the myosin nucleotide binding site, which might be a reasonable mechanism of the actin-induced myosin conformational changes at the initial stage of the power stroke.

Contents

CHAPTER 1: Introduction

1.1 The molecular motor - myosin superfamily

Molecular motors are biological molecular machines that play a significant role in the essential movement in various living organisms. They use external energy sources by operating as molecular motors, in order to produce a directed motion and convert chemical energy into mechanical work [1]. This special function of molecular motors is involved in a large number of cellular tasks. It is able to generate cell movement, ion gradient, molecular transport across membranes, organelle transport and protein folding and unfolding processes [2]. However, the general mechanism they used to convert chemical energy into mechanical work is simple and ingenious. Myosins, kinesins, and dyneins are three types of cytoplasmic molecular motor families, which are involved in motions along filaments: kinesins and dyneins translocate along microtubules, whereas myosins move along actin filaments and are responsible for muscle contraction and other cellular movements. In the past 50 years and recently, one of the most extensively studied molecular motors is the myosin. It drives contraction through nucleotide binding and hydrolysis, which causes conformational changes in the globular motor domain and this conformational change, is translated into movements with the aid of accessory functional motifs [3].

Myosins belong to a large superfamily of proteins that share a common feature with other motor proteins, to hydrolyze ATP, to interact with actin and to generate the force [4]. However, not all the members of the myosin superfamily possess such biochemical characteristics, and some of them have lost one or more of these features. Both muscle myosin (conventional myosin) and unconventional myosin are typically constructed from three functional subdomains. The foremost important subdomain is called the motor domain, whose core sequence is highly conserved in all myosin classes. This motor domain contains three functional motifs: a converter region, the nucleotide-binding pocket and the actin-binding interface. This converter region is the motion and force generation linkage between the core motor and the following lever

arm region [5-7]. The myosin motor domain, which contains a common nucleotide-binding pocket like a microtubule-based kinesin motor, is responsible for the conformational changes and force generation [8]. Actin-binding interfaces of myosin are principally composed of the upper and lower 50 kDa domains, which are separated by a deep cleft. This cleft with many conserved residues contributes a negotiation between the nucleotide-binding pocket and the actin-binding interfaces [3]. The second subdomain is the neck region or lever arm that is composed of a long helix. The length of this helix characteristically depends on the number of IQ motifs, which share a consensus helical sequence (IQxxxRGxxxR) requisite to either light chains or calmodulin [9, 10]. The number of IQ motifs presented from the neck region in different myosins can vary from zero to six except for one special case: the Toxoplasma myosin A of class XIV lost this neck region [9]. The last subdomain called the tail region is extremely variable in sequence length, domain composition and organization. The tail region contains a helical coiled-coil domain, which is believed to be an impetus for the dimerization of two motor domains, whereas one motor domain is formed with only one helical tail, but not in a coiled-coil conformation [10, 11]. Its biological or mechanochemical roles have not been completely ascertained for all of its conformational variety.

The newly reconstructed tree of eukaryotic life is based on 2,269 myosin motor domains from 328 different organisms. Phylogenetic analysis groups myosins into 35 distinct classes based on the motor domain, but sixteen kinds of myosin classes have not been proposed previously. Some customary relationships of major taxa and preliminary classifications are confirmed by the resultant phylogenetic tree [12]. The Clustal-W package was used to perform a distance matrix analysis [13]. Myosin I as the first unconventional myosin was famous for its singular head. This myosin is responsible for cytoskeletal reorganization and organelle translocation, but it is not able to be self-associated into bipolar filaments [14, 15]. Myosin II can be divided into four groups: skeletal/cardiac muscle myosins; vertebrate smooth muscle/nonmuscle myosins; *Dictyostelium/Acanthamoeba*-type myosin and yeast-type myosins. The common feature of myosin II is self-assemble to form a

variety of filament structures based on their helical coiled-coil tails, which is essential to exhibit their functions [11]. Myosin III differs from all other myosins by its N terminal kinase domain, which can serve as a link between the signaling complex and cytoskeleton for photo-transduction [16]. Myosin IV is predicted to have a single motor domain, one IQ motif and a tail with a myosin tail homology domain (myTH4) [17]. Myosin V has been identified in most of the eukaryotic cells and examined except those of plants. Myosin Va is a processive motor protein with a high duty ratio and a large step size. These properties make it ideally suitable for taking part in polarized cell growth, membrane trafficking and specific transport pathways [18]. Myosin VI has a single IQ motif, a tail region and a unique globular head domain. There is a unique insertion between the neck region and the motor domain, called the "reverse gear" that enables myosin VI to move to the minus end of actin filaments (in the opposite direction, unlike all other myosins) [19]. Myosin VII has two talin-binding (FERM) domains and an SH3-domain [11]. Myosin classes VIII, XI and XIII have been detected in plant cells [20]. Myosin IXs have been identified in a wide variety of tissues and cell types are supposed to be involved in intracellular signaling pathways, which are yet to be ascertained [21]. Myosin XIV contains a motor domain, no classic IQ motifs and a variable-length tail, which exhibit a unique substrate dependent gliding motility, which is essential for host cell invasion [22]. Little is known about the functions of myosin X, XII, and other myosin classes.

1.2 Myosin II structural features

Members of the myosin II class have been considered as conventional myosins in the past few decades [11]. This class can be proteolytically cleaved into several functional domains (see Figure 1). The first cleavage site is located at the junction of myosin head and tail. The break-up of this point will separate myosin into a tail fragment of a coiled-coil dimer conformation and a soluble fragment named S1. Meanwhile, the soluble fragment S1 still keeps two light chains and binds to the actin and nucleotides. The second cleavage site situated in the tail region can divide myosin

into two fragments named heavy meromyosin (HMM) and light meromyosin (LMM) [23].

The heavy meromyosin (HMM), containing the head region and part of the coiled-coil fragment from the tail region, is soluble even at low ionic strength. This tiny portion of the tail region termed subfragment 2(S2), can assist in the formation of the two-headed conformation by dimerization. Moreover, the LMM still retains solubility properties like the original intact myosin II molecule. Numerous researches have been investigated for both S1 and HMM in their kinetic features, biophysical characteristics and structural changes [23, 24].
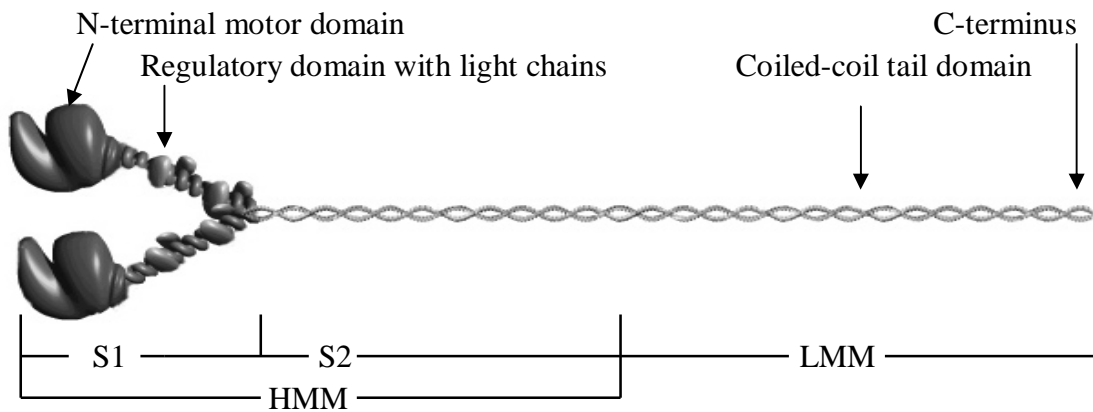
Figure 1: The homodimer myosin II molecule with indication to its proteolytic fragments. The head of the myosin (subfragment 1, S1) consists of the catalytically active N-terminal motor domain and a light chain binding to the regulatory domain. The long coiled-coil tail serves for dimerization and myosin filament formation. S1 and S2 create the heavy meromyosin (HMM), while the C-terminal fragment is the low meromyosin (LMM).

The X-ray crystallographic three dimensional structures of several myosin II molecules have been solved after proteolysis or by expression of recombinant molecules. They are including the chicken fast skeletal muscle myosin S1 [5], the S1 of scallop myosin [7], and the recombinant fragment of the smooth muscle myosin motor domain with the essential light chain (ELC) [6], and a motor domain of *Dictyostelium discoideum* myosin [25]. Structures with several different nucleotides or nucleotide analogs bound to the myosin active site have also been resolved. The myosin motor domain essentially consists of discrete regions connected by flexible

linkers [6]. The N-terminal subdomain forms a SH3-like motif in myosin II that is not present in other myosin classes [25]. This N-terminal subdomain is functionally connected with the upper 50 kDa and lower 50 kDa subdomain. The name "50 kDa" is used for historical reasons. After two sensitive surface loops proteolytically cleaved, there will be a band on the sodium dodecyl sulfate polyacrylamide gel (SDS-PAGE) after electrophoresis with a molecular weight of 50 kDa [26]. The upper and lower 50 kDa domains are the main actin-binding interface in most myosin families, which is separated by a deep cleft with many conserved residues. This cleft closure is functionally coupled with conformational changes of the nucleotide-binding sites [23] (see Figure 2).
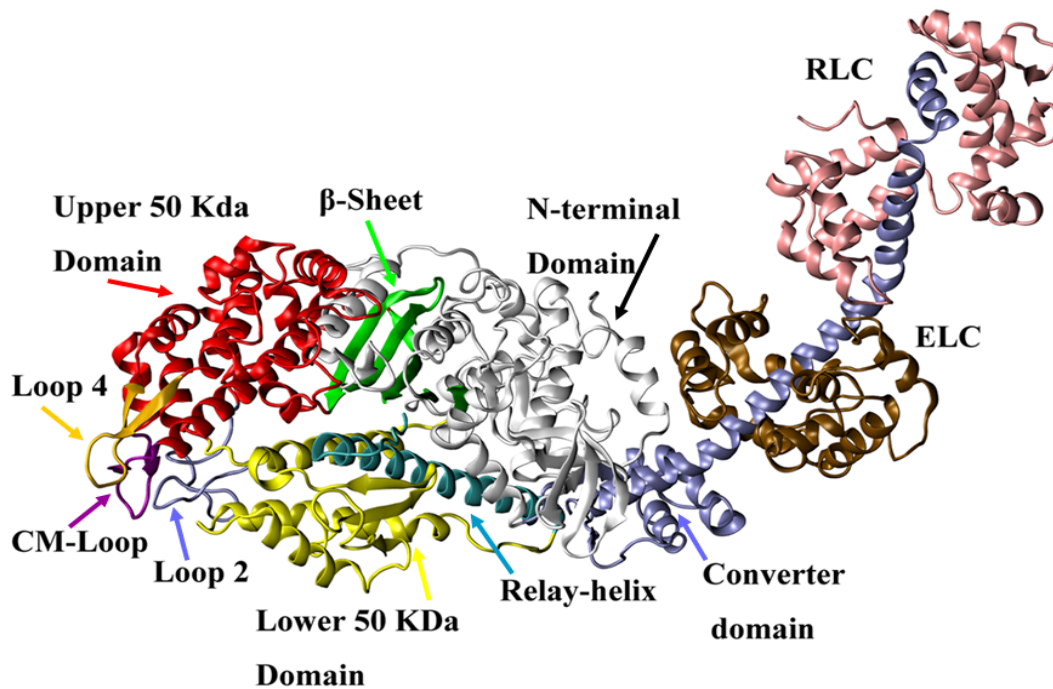
Figure 2: Three-dimensional structures of chicken fast skeletal muscle myosin S1. These color-codes for the functional subdomains of myosin S1 are represented by drawing method of cartoon. (Red: upper 50 kDa; Yellow: lower 50 kDa; Purple: cardiomyopathy loop (CM-loop); Light blue: loop 2; Orange: loop 4; Cyan: relay-helix; Ice blue: converter domain; Green: β-sheet; White: N-terminal domain; Brown: essential light chain (ELC); Pink: regulatory light chain (RLC)).

From the structural point of view, the nucleotide-binding pocket is located between the upper 50 kDa and the lower 50 kDa subdomains and includes three functional loops: the P-loop and the switch-1(SW-I) and switch-2(SW-II) loops. The P-loop enables the nucleotide binding to the myosin proteins regarded as an evolutionarily

conserved structure in the P-loop NTP-ases. The SW-I loop takes part in the structural changes accompanied by an actin-binding cleft opening and closure. The SW-II loop is involved in the force generation related lever arm movement. The conformations of these two loops (SW-I and SW-II) are very sensitive to different kinds of nucleotide binding or release, and conformational changes of the lever arm [23]. Besides these loops, there is a seven-stranded β-sheet surrounded by numerous loops and helices, which are one of the most typical protein structures for an evolutionary conserved platform of P-loop NTP-ases [27]. The last important subdomain is the converter domain and related lever arm. Two calmodulin-like "light-chains" bind to the helix-like lever arm and stabilize it. The x-ray crystallographic structure of the lever arm has been found in two conformations, which are apart from 60° and appear to be the two ends of the power stroke. The converter domain enables to amplify the myosin movement via the lever arm based on its rotations [23]. The SW-II loop, having a tight communication with the converter domain via the relay-helix region, can transmit the conformational changes from the nucleotide-binding site to the converter domain [23].

1.3 Actin structural features

Although actin has two forms, the G-actin (globular) and the F-actin (fibrous), the available crystal structures are only from the G-actin form. The crystallographic structure of G-actin was solved in complex with DNAase I, gelsolin and profillin [28, 29]. The actin structure consists of two similar domains (the DNA-binding domain and the myosin-binding domain) and each of them contains the associated-helices and a five-stranded β-sheet. The nucleotide (ATP or ADP) coupled with Mg^{2+} or Ca^{2+} is located between two of five-stranded β-sheets. Each of the domains also contains two subdomains. One is used to form the parts of the nucleotide-binding pocket, and the other is involved in actin-actin interactions [30] (see Figure 3).
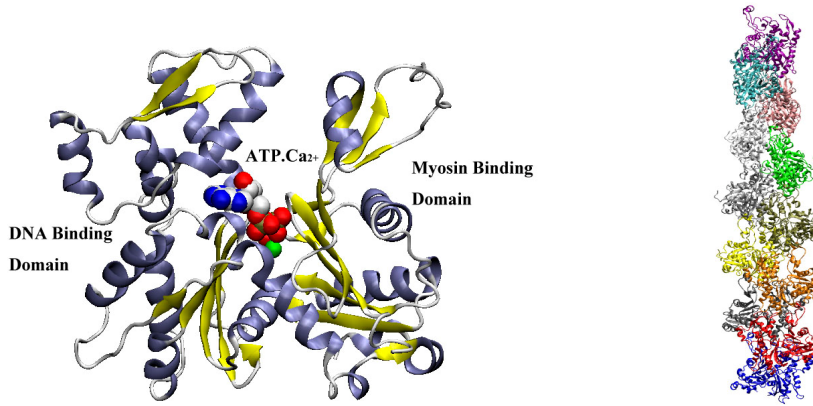
Figure 3: The atomic structures of G- and F-actin. The left panel shows the cartoon model of G-actin (1ATN.pdb). G-actin consists of two domains: the DNA-binding domain and the myosin-binding domain. The right panel shows the refined structure of F-actin from G-actin monomers (1O1G.pdb).

The G- to F-actin transition is growing on the basis of cellular functions, but the structural background of this transition remains enigmatic. A number of efforts have been attempted to make G-actin monomers fit into the F-actin helix by *in silico* refinement methods [5, 31] (see Figure 3). As a result, distinct F-actin structures have been produced by these methods, but the refinement of the F-actin helix is still not well determined because of the limited resolution (6 to 8 Å) of the fiber diffraction patterns from cryo-electron microscopy. In the refinements, four subdomains of G-actin are constrained to remain connected, but allowed to move as independent solid bodies in the minimization process based on the observed fiber diffraction pattern [30]. After that, the stereo-chemistry of the residues that connected the four subdomains is further optimized [32]. Finally, the thin filament (F-actin) with actin molecules (42 kDa) arranged on six left-handed turns, repeating every 36 nm is obtained in a helical polymer with a 2.75 nm rise per subunit (G-actin) [33]. Each of actin monomers turns 166° from the previous one, which lends the filament to a double helical shape. Each monomer interacts with four neighboring monomers, and thirteen G-actin monomers constitute one helix turn. In 2010, Wakabayashi *et al.* [34] present a F-actin model through the cryo-electron microscopy (cryo-EM) in the presence of phosphate and with some α-helical backbones and large side chains visualized. This EM-map based complete atomic model of F-actin proposes a possible

molecular mechanism for the biochemical events e.g. actin polymerization. The high resolution structure of F-actin is solved in the presence of ADP by Oda *et al.* [35]. This F-actin model was created by using X-ray fiber diffraction intensities with radial direction in 3.3 Å and 5.6 Å along the equator. Conformation of this helical F-actin with its actin monomer in a flat form and two major domains of every monomer rotated to each other about 20 degrees is essential for the stable formation.

1.4 Lymn-Taylor cycle

Myosin possesses a product-related ATPase enzyme mechanism analogous to that of the G-proteins due to the presence of the functional loops, i.e. P-loop and the SW-I and SW-II loops, which are residing in the myosin nucleotide-binding pocket. During this process, myosin undergoes numerous structural changes, which movements are coupled in different ways. The muscle contraction from consuming a molecule of ATP, hydrolyzing it to perform a power stroke was first postulated by Lymn-Taylor in 1971[24] (see Figure 4). Myosin, with its closed cleft in the absence of a nucleotide, strongly binds to the actin filament, creating a so called "rigor" complex. The transition from state I to state II of the Lymn-Taylor cycle shows that the rigor actomyosin complex is easily dissociated by ATP binding to the myosin. This ATP binding not only opens the actin-binding cleft caused by the conformational changes in the nucleotide-binding pocket, but it also results in the myosin to lose its strong actin-binding affinity [26, 37]. From state II to state III, the myosin converter rotates the lever arm by 60° and pushes it to the up-lever conformation followed by the activation of the ATPase function in the absence of actin. This process is the so-called recovery step with no accompanying conformational changes in the actin-binding cleft [38-40]. From state III to state IV, the ATP hydrolysis results in a stable state of myosin·*ADP.Pi* complex. This up-lever and open cleft state of myosin starts to weakly bind to actin. In the last step, from state IV to state I, the actin binding triggers three processes that are accompanied by structural changes in myosin: the further closure of the cleft with the increasing binding affinity of actin, the release of the products, and

the lever arm moving from the up to the down position. This process is referred to as the power stroke, which moves the myosin approximately 10 nm along the actin filament [41]. After the power stroke, the ADP release takes place followed by the phosphate release from the myosin by which myosin will return to its rigor conformation [24]. ATP binding will subsequently release myosin from the actin filament to start a new cycle.
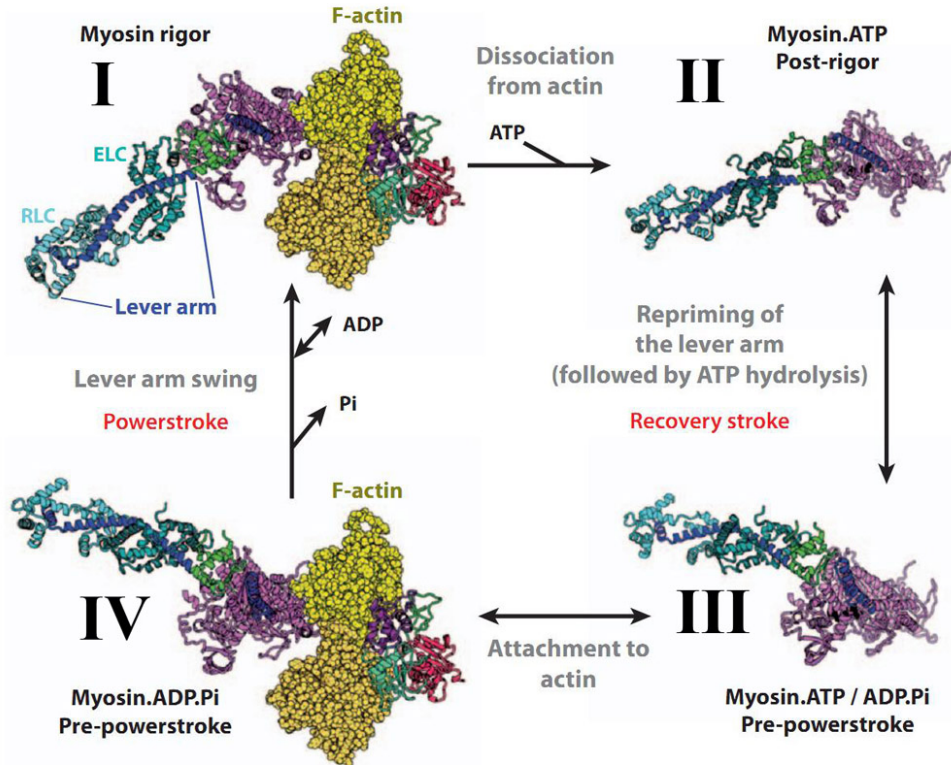
Figure 4: The Lymn-Taylor cycle with annotation. The myosin II S1 fragment is depicted in three structural states: rigor, post-rigor, and pre-power stroke. The lever arm is composed of a heavy-chain helix (blue) surrounded by the light chains (cyan) of myosin II. The lever arm position is controlled by the position of the converter (green), which swings relative to the rest of the motor domain (light purple). The structure of the relay-helix (blue) within the motor domain is likewise indicated. Three actin monomers are indicated to represent the F-actin filament. Two of actin monomers are shown space-filled with yellow spheres, and one is shown as a ribbon diagram. The original figure was published by Sweeney HL *et al*., 2010. Figure 4 in the original paper [24].

1.5 Actin-myosin binding interface

The actin attachment induces a weak to strong binding transition with over five orders of magnitude difference in affinity (from state III to state VI in the Lymn-Taylor cycle) [42].

Coupling of myosin II atomic structure with the decorated actin constructed from cryo-electron microscopy has produced a 13Å resolution density map (EM-map) [43]. A near-atomic view of the actin-myosin II rigor conformation can be achieved by combining structural data from the head of myosin V, which is crystallized in the absence of actin and supposed to be a strong actin-binding state [44, 45] (see Figure 5).
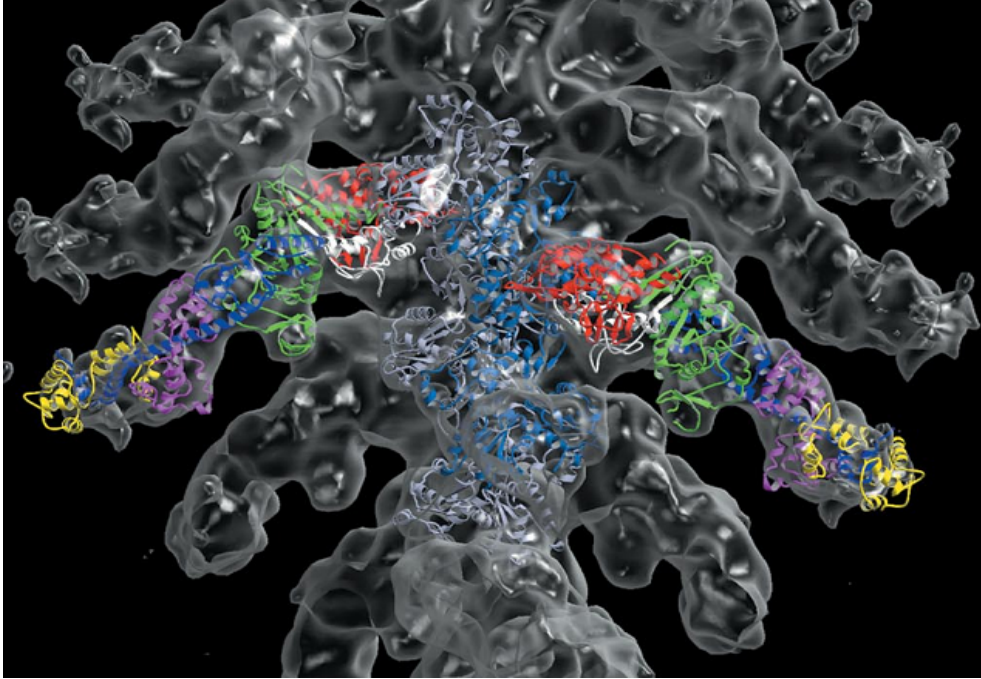
Figure 5: Fit of crystallographic molecular models of F-actin and myosin S1 into the reconstructed density map. Molecular models of myosin and F-actin are docked into the experimental density map, which is generated by the cryo-electron microscopy reconstruction. The myosin has been kept in its original X-ray crystallographic conformation (post-rigor). The original figure was published by Holmes KC *et al.*, 2004. Figure 4 in the original paper [45].

By flexible fitting the atomic structures into the EM-map, it has been revealed that two actin monomers from the actin filament named AC1 and AC3 are not only involved in the entire binding region of one myosin, but they also increase the contact surface to stabilize the strong actin-binding state [46]. Milligan *et al.* [47] suggested that there might be an interaction between loop 3 of myosin and actin AC1, which is also demonstrated by Mornet *et al.* in the zero-cross-linking experiment [48]. Holmes *et al.* showed similar features to previous results. Namely that there is a hole in the EM-map and a "finger" below the hole is referred to as the "Milligan contact" [45]. However, Liu and his colleagues presented a contradictory conclusion from their

experiments without "Milligan contact" [49]. When replacing the DNase I binding loop with the original loop from the crystal structure of actin, an interaction between the DNase I binding loop and myosin S1 was revealed [43]. Strong interactions between actin AC1residues 44-49 with myosin S1 residues 543-554 were confirmed by H-bonds in the binding surface [47]. Tyr91 and Arg95 of the actin AC1 can also interact with loop 3 of the myosin S1 [43]. However, other contacts between Glu576 (myosin S1) and Arg95 (AC1), Lys569 (myosin S1) and Glu99 (AC1) were identified by Milligan *et al.* [47].

In addition, the actin AC3 contains a larger interaction surface with myosin S1 than the actin AC1. The contact surface indicates several potential electrostatic interactions and H-bonds. Hydrophobic interactions between the residues 528-544 of myosin lower 50 KDa domain and residues 349-353, 146-149 of the actin AC3 are maintained [43]. Potential electrostatic interactions and H-bonds between myosin S1 and AC3 are mainly found in loop 2, loop 4, and the cardiomyopathy loop (CM-loop) of myosin. Loop 2 contains two lysines (chicken myosin) that make interactions between Lys640 (myosin S1) and Asp25 (AC3), Lys642 (myosin S1) and Glu334 (AC3). Loop 4 in the form of a β-hairpin running from residue 366 to 377 in myosin S1 also establishes electrostatic interactions and H-bonds with residues from actin AC3. The CM-loop makes a salt-bridge between Glu411 (myosin S1) and Lys336 of the actin AC3 [43]. A critical mutant of R403Q in the human cardiac myosin occurs in a severe familial hypertrophic cardiomyopathy and results in a decreased level of the actin-activated myosin ATPase due to a weaker binding affinity to actin [50]. Arg403 (myosin S1) is far away from actin AC3 and forms a salt-bridge with Glu631 in loop 2 (myosin S1) [43]. Therefore, the function of Arg403 is probably to stabilize the folded conformation of loop 2 in the actomyosin. Lys415 (myosin S1) forms a stable H-bond with Glu334(AC3), which was found by Liu *et al.* from MD simulations [49], but this interaction was not discovered from the results of the molecular dynamics flexible fitting (MDFF) in Holmes's model [51] (see Table 1). Onishi *et al.* introduced mutants to heavy meromyosin (HMM) by phosphorylation in three regions, residues 546-548 (hydrophobic region), residues 407,409 and 412-414 in the CM-loop, and

residues 652 and 653 in loop 2. They all have significantly lower the actin-activated ATPase in HMM or even utterly extinguish it [52]. Those mutants with possible contacts with residues of the actin AC3 were also found in the results of the MDFF refinement [43, 51] (see Table 1).

Table 1: Interactions and Bonds formed between the myosin S1 and two actin monomers based on cryo-electron microscopy in strong atin-binding state

Potential electrostatic interactions		Potential H-bonds	
AC3 - S1	AC1 - S1	AC3 - S1	AC1 - S1
ASP24-GLU629	LYS50-GLU576	GLY23N -LYS637O	LYS50NZ-ASN552OD1
ASP25-LYS640	ARG95-GLU576	ASP24N -GLY635O	ARG95NE-PR0570O
ARG28-GLU629	GLU99-LYS569	LYS328NZ- ARG371O	ARG95NH1-ALA575O
ARG147-GLU373		TYR337OH-ASN410OD1	ARG95NH2-LYS572O
GLU167-LYS544		SER348OG-LYS637O	GLY46O-LYS553NZ
ASP311-ARG371		THR351N-PR0529O	GLY48O-LYS553N
LYS328-GLU372		THR351OG1-PR0529O	GLN49OE1-SER549OG
GLU334-LYS642		GLY23O-GLY638N	TYR91O-ALA571N
ARG335-GLU372		ARG28O-ASN410ND2	
LYS336-GLU411		SER145O-GLY643N	
		GLU167O-LYS544NZ	
		GLN314OE1-ARG371NH1	
		ILE329O-ARG371NH2	
		GLU334O-LYS642NZ	

The weakly and strongly bound forms of actomyosin are basically determined by various nucleotide bound to myosin [42], whereas the main discrimination is depended on open or closed conformation of the actin-binding cleft [43]. However, the precise mechanism of the open-close transition of the actin-binding cleft is not yet clear due to limited information on weak actin-binding state.

The weak actin-binding state is formed when the myosin head initially attaches to the actin. Many functional experiments have been carried out with expressed myosin mutants at the probable actin-binding sites of myosin [52, 53]. These results suggest that some residues are predominantly engaged in the weak actin-binding state, whereas other residues are involved in the weak-to-strong actin-binding transition. One *in silico* based speculations has shown that five surface loops of myosin are

possibly involved in the weakly bound form of actomyosin [54]. One loop bears three hydrophobic residues, the second is loop 2, the third is the cardiomyopathy loop (CM-loop), the fourth is loop 3 and the last is the proline-rich loop. Hydrophobic triplets (Val534-Pro536) of Dictyostelium myosin are close to Leu142-Ala144 of actin. Lys622 and Lys623 of the myosin loop 2 interact to Asp24 and Asp25 of actin. ARG402 of the myosin CM-loop fits to the C-terminal region of the actin. Arg562 of the myosin loop 3 binds to Glu99 and Glu100 of the other actin monomer [54]. However, these proposed interactions have not been proved experimentally, except a novel discovered actin-binding region of myosin named the 'activation loop', which interacts with the N-terminal region of actin [55]. This interaction accelerating the movement of the relay and stimulating myosin's ATPase activity results in efficient force generation, but it is not essential for the unloaded motility in the weak actin-binding state [55].

The weak-to-strong actin-binding transition prompts that actin bound initiates conformational changes on the actomyosin interface and travels elsewhere in myosin including actin-binding cleft closure, lever swinging and later phosphate release, to carry out various functions.

1.6 Coupling between the nucleotide-binding site and myosin functional domains

The Lymn-Taylor cycle is composed of several different mechanisms (e.g., recovery step, power stroke), which tightly couple kinetic and thermodynamic events with conformational changes in different myosin regions. Nucleotide-binding takes place at the active site in the center of the myosin catalytic domain. It promotes conformational transitions from the center of the myosin head to the other myosin functional domains. ATP-binding results in a loss of affinity in the actomyosin complex and dissociation of myosin from the actin filament. ATP hydrolysis and post-hydrolytic phosphate-release are accompanied by an increase in the actomyosin's binding affinity. In addition, changes in binding affinity are also relied on various pathways of phosphate release.

The binding between the nucleotide-binding site and the actin-binding surface is described as an antagonism. ATP binding to actomyosin causes actin dissociation, while actin binding to myosin accelerates phosphate and ADP release [56]. It has been shown that the open and the closure of the actin-binding cleft are involved in the communication between the nucleotide-binding site and the actin-binding surface of myosin [57]. In general enzymatic reactions, the product release usually occurs via the same route as the substrate binds. When ATP is used as a substrate for later hydrolysis, phosphate release proceeds to ADP release via a different route because of ADP binding blocks the front door where the phosphate enters. In this case, one of the possible phosphate release routes is through a rear opening switch that is situated in the tube of the actin-binding cleft called the back-door release. Phosphate releasing into the 50 kDa cleft alters conformation of the actin-binding cleft and thereby affects actin-binding affinity based on the MD simulations [58]. An alternative pathway results in the opening of the SW-I loop by a rearrangement of the actomyosin interface in the weak actin association. The opening of the SW-I loop breaks the salt-bridge between Arg238 and Glu459, which accelerates phosphate release through a newly formed "trapdoor" rather than the previously suggested back-door [37, 58]. This trapdoor was discovered in a nucleotide-free crystal structure of *Dictyostelium* myosin II, where both the SW-I and the SW-II loops have moved away from their positions and formed an open state [37]. This structure reveals significant changes in the actin-binding region, which suggests that the actin-binding cleft closure is mechanically coupled with the opening of the SW-I loop [56]. This fact has also been confirmed by fitting crystallographic structures into the high-resolution electron micrographs of actomyosin [43] and by a crystallographic study of myosin V [44]. According to the phosphate release induced conformational changes in the nucleotide-binding site and open-close of the actin-binding cleft of myosin, this coupling movement implies that the antagonism described at the beginning of this paragraph can be further elaborated as a scissors-like mechanism [59]. It means that the larger the space of the nucleotide-binding pocket expanded, the more narrowed space of the actin-binding cleft achieved and the higher binding affinity of the

actomyosin complex obtained.

The communication event between the catalytic site and the other myosin functional domains not only modulates myosin's affinity for actin, but also triggers the conformational change of the lever swing in the recovery step (from state II to state III in the Lymn-Taylor cycle). A proposed mechanism of conformational changes in the nucleotide-binding site coupled with the lever-swing linked converter domain rotation has been investigated on the basis of MD simulations [39]. Several regions of myosin are involved in the communication between the nucleotide-binding site and the converter domain. The conformational changes of the relay-helix and the relay-loop (collectively called the relay region) are initiated by the SW-II loop [40]. This coupling mechanism can be regarded as though the SW-II loop in an open state with the lever arm in its down position and vice versa. The Gly457/Ser456 peptide group belonging to the SW-II loop forms a hydrogen bond with the γ-phosphate during the transition of the recovery step. This bond formation is essential for the ATP hydrolysis step before myosin binds to actin. If the SW-II loop freely close and re-open with the lever still in down position, ATP hydrolysis and an unproductive product release would be wasteful processes. Therefore, it is implied that the SW-II loop has to be closed in order to couple with the conformational changes of the converter domain, which leads the lever arm to its up position [39]. The movement of the converter domain is the result of the angle changes between the SH1-SH2 regions, which is influenced by the movement of the relay region in close proximity [60]. The residue Gly680 residing in the SH1-SH2 region has been proposed as a pivoting point for the structural changes on the basis of the time-resolved fluorescence measurements of muscle fibers [61]. The mutation experiments showed that G680A lead to 20 to 30-fold slower in the nucleotide binding rate and approximately 10-fold increased mantADP affinity in the absence of actin. These results suggest that the function of the SH1-SH2 regions is involved in the communication between the nucleotide-binding site and the converter domain [62].

The operational work of lever movement should be harmonized with the ATPase cycle, actin binding and detachment cycle (see Figure 6). In ATP form of myosin, the myosin

lever moved from the "down" position to the "up" position in actin detached state is the so called recovery step (from state II to state III in the Lymn-Taylor cycle) [39]. After ATP hydrolysis, if the up-to-down lever swing occurs in an actin-detached state of myosin, which is called reverse recovery step (equal to the futile lever swing in Figure 6) [63]. This step is regarded as a rate-limiting step of the basal myosin ATPase to the subsequent rapid product release without produce useful mechanical work. The crucial difference of the down-to-up lever swing with ATP-bound of myosin in the recovery step is four orders of magnitude faster than up-to-down lever swing in the reverse recovery step [63]. If actin binds to the up-lever state of myosin to move the lever to the "down" position, then an effective power stroke is generated (from state IV to state I in the Lymn-Taylor cycle). It is worth noting that the reverse recovery step is a reversible step in the absence of actin while power stroke occurs in the actin-attached form and becomes irreversible. The ratio of the effective power stroke over reverse recovery step can be accelerated up to several orders of magnitude with actin binding in the ADP·Pi state of myosin. Therefore, the predominant reaction flux is diverted to an actin-attached myosin power stroke, even though it is thermodynamically less favorable pathway because of the low actin-binding affinity [55,64] (see Figure 6).
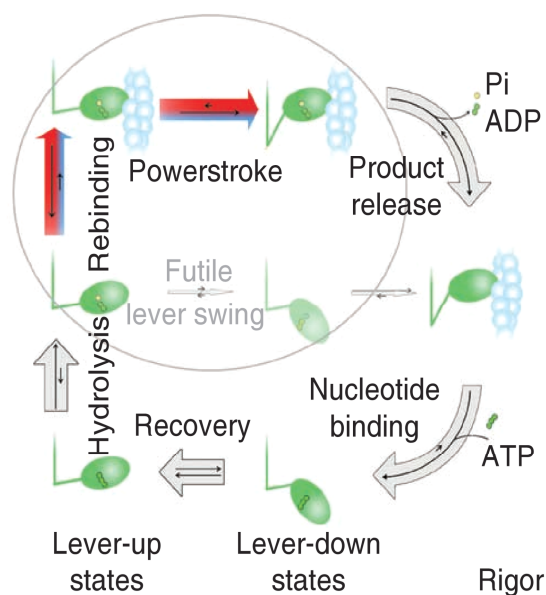
Figure 6: The chemomechanical cycle of actomyosin. When ATP binds to myosin, the myosin disassociation caused it lever swings to the up state (recovery step). After ATP hydrolysis, actin weakly rebinds in a step kinetically coupled to the power stroke. If the lever swing occurs before actin binding, then a futile lever swing (reverse recovery step) is generated. The original figure was published by Málnási-Csizmadia et al., 2012. Figure 1C in the original paper [55].

1.7 Scientific questions addressed in this thesis

When ATP binds to myosin, myosin disassociated from actin results in the down-to-up lever swing (recovery step). In the lever-up state of myosin, ATP hydrolysis occurs. After hydrolysis in the absence of actin the rate limiting up-to-down lever swing step (reverse recovery step) occurs. In this case, the lever swing is inefficient regarding the mechanical cycle since the swing happened in the actin detached state of myosin. If actin weakly binds to the up lever state of myosin, the rate of lever swing step is highly accelerated. The actin activation of the lever swing step channels the reaction pathway into the effective power stroke which mechanism is called kinetic pathway selection. We were mainly interested in the structural and energetic relationship of the myosin lever's back and forth swing in the recovery step and the power stroke. The following questions were stated upon the design of the project:

Firstly: A direct structural trajectory of the recovery step has been deduced from previous *in silico* simulations. But it neglected the dynamic behaviors of the protein that allows it to populate the neighboring conformational spaces along the suggested trajectory. Therefore, the question is how can we combine the intermediate states with suggested structural trajectory to discover the complete mechanism of the down-to-up lever swing in the recovery step?

Secondly: The precise mechanism of the up-to-down lever swing in the power stroke is still unclear due to limited structural information on the weak actin-binding state. What is the atomic structure of the weak actomyosin complex?

Thirdly: What conformational changes are induced by the actin binding in the different states of myosin?

Fourthly: What are the differences of the binding interfaces of the weak and the strong actomyosin complexes?

Fifthly: The effective pathways of the power stroke are not accessible due to the short lifetime of their intermediates. What is structural pathway and possible mechanism of the actin-binding induced myosin up-to-down lever swing in the power stroke?

CHAPTER 2: *In silico* approaches for structural investigations of proteins

In recent years, with the rapid development of computing power, the more extensive use of *in silico* methods is increasingly applied to a variety of life science researches. Since in my project, I mainly use molecular structural modeling, protein-protein docking, molecular dynamics (MD) and etc, in the following chapters I will summarize major principles and applications of these approaches.

2.1 Molecular structural modeling

In the biological research, some unstructured regions or subdomains can be predicted by at least one homologous protein with accurately known in its three-dimensional (3-D) structure, when their functional characterization of the 3-D structure cannot be determined from experiments [65]. Principle of comparative modeling predicts the 3-D structure of the protein from the sequence based on its alignment to one or more proteins of known structure. Many of the protein sequences can be modeled according to the target-template alignment. Results of prediction accuracy steadily have been increased because of the increased number of known protein structures as well as improvements in modeling softwares (see Table 2 in the Appendix). Nowadays, it is possible to accurately model more than half of the unknown protein sequences or subdomains based on the experimentally determined low-resolution structures [65, 66]. Even the forecast coarse structural features can be useful in showing some functional aspects of protein [67]. There are a number of computer modeling programs and web servers used for comparative modeling (see Table 2 in the Appendix). However, these automated web servers are not designed for unusual modeling cases or problematic alignments, such as modeling of the ligand binding, which needs other various modeling tools or softwares to solve these questions manually.

MODELLER is a program used for comparative modeling of unstructured proteins [65, 68]. New models with all non-hydrogen atoms by satisfaction of spatial restraints [69] according to some known protein structures can be generated through this program [70]. Three main steps of comparative modeling will be described in this program as below. The first step is to select protein structures as templates for the target sequence via some related web servers (see Table 2 in the Appendix). Two methods of protein selection are used in fold assignment identification. The first one is by using pairwise sequence-sequence comparison to select the same or similar sequence from the database for the target sequence [65]. The second one is by using profile analysis based on multiple sequence information to discover the important structural relationship when the sequence identity between the target and the template drops below 25% [69]. It is important to select templates with higher overall similarity for the given target sequence to generate the best 3-D structure model. Phylogenetic tree, environmental similarity (e.g., solvent, pH, ligands) and quality of the template structures are three critical factors that should be taken into account in the first step. In the second step, once templates have been selected, an algorithm should be selected to align the target sequence with the template structures. PAM [71] and BLOSUM [72] as two of dynamic algorithms by using standard substitution matrices are the preferred choice for the alignment. Usually, we can obtain a rational result, when the identity of alignment is higher than 40% between the template protein and target sequence [73]. When the sequence similarity falls below 30%, the alignment becomes problematic in the "twilight zone" with an increasingly large number of gaps and errors. The final step is to generate a structural model. The "modeling by assembly of rigid bodies" based on overlapping the protein structure into the conserved core regions and selecting variable loops from templates to decorate the backbone of the target sequence is one of the widely used approaches [74]. The other approach of modeling is called segment matching or coordinate reconstruction. A subset of atomic positions of the target sequence can be utilized as "guiding" positions for the template structures to identify and assemble all-atom short segments [75].

In comparative modeling, loops often play an important role in forming the active binding sites and defining the function of a given protein. But, loops are generally too short to provide sufficient information upon their folding. Even some identical sequences of loops in different proteins do not always adopt the same conformation. Although a number of *in silico* programs focused on the loop modeling have been developed, none of them can predict and build the target structure in a confident conformation, if the residue number is less than eight in the targeted loops [70]. In this thesis, the comparative modeling method can be utilized to build up the rational conformation of the missing loop 2 region of myosin and fix other missing residues, which of these missing parts are difficult to be resolved in the X-ray crystallographic diffraction experiments.

2.2 Protein-protein docking

Most of the proteins need to interact with other proteins and then form an active complex to achieve their function. Numerous methods are available for the study of protein complexes at different levels, but only a few of these techniques can provide high-resolution information at the atomic level. X-ray crystallographic diffraction and nuclear magnetic resonance (NMR) are two of the most popular experimental methods for solving the three-dimensional structure of proteins and for better understanding the biological function. The major challenge for X-ray crystallography studies in resolving the structures of complexes is difficult to achieve the crystallization due to the dynamics of the complex formation. On the other hand, the high molecular weight of the complex is the major problem for NMR in its size limitation. These obstacles make theoretical methods e.g., protein-protein interactions been well developed at the atomic level to study those unsolved protein complexes during the past few years [76, 77].

Docking as one of the most realistic methods of protein-protein interactions enables to predict the preferred orientation of one molecule to another when they bound to each other in a stable form [78]. Historically, docking research in the late 1970s

focused on refining a model of a complex structure by optimizing the separation of the interaction pattern while keeping their relative orientations fixed. Hence only a few configurations can be discriminated by *in silico* approaches after all heuristic constraints have been imposed [79]. Later, the relative orientation of each of the docking partners was allowed to change, but the internal geometry of structures, such as bond angles, bond lengths and torsion angles of the components were not changed during the generation of complex configurations, which was referred to as "rigid docking" [80]. With further increases of computational power, the internal geometrical changes of the interacting partners that were used to be extremely expensive in calculation became actual implementation, which is now referred to as "flexible docking" [81].

The molecular docking process can be divided into two steps. First, it creates an optimum number of configurations based on the experimentally determined binding modes from the search algorithm and then uses scoring functions to evaluate these experimental binding modes. Scoring functions are fast approximate mathematical methods used to predict the strength of the non-covalent interaction (also referred to as the binding affinity) between two molecules after docking process. Mostly one of the molecules is a small organic compound such as a drug and the other is the drug's biological target such as a protein receptor [82]. Scoring functions have also been developed to predict the strength of other types of intermolecular interactions between two proteins [83]. Nowadays, scoring functions are used in docking methods in two different ways. First, a reduced function is used to direct the search, and then a more rigorous one is used to rank the resulting structures. Second, a full scoring function is used to rank the complex conformation in the initial run. Then, the system of the complex conformation is modified by the search algorithm to re-rank those new generated complexes with the same scoring function again. Finally, the best ranked complex is yielded after several runs of repeatedly filtering processes. A parallel approach for analyzing protein-protein interactions is involved in most of docking programs. The procedure starts with one protein fixed in a certain space, and the second one is rotated and translated around it. When serials of different

configurations are obtained, a score function is used to calculate the various elementary terms, such as surface complementarities, electrostatic interactions, van der Waals repulsion, and so on. In order to receive a high resolution or plausible complex conformation, the current problem for all of applied methods is that the searching for entire conformational space of complex geometry makes the computer calculation extremely expensive.

As one of the new protein-protein docking programs, HADDOCK, not only uses the same search algorithm and scoring functions in each step as other programs, but it is also involving a novel high ambiguity driven docking algorithm [84, 85]. It means that the docking procedure can integrate some biochemical or biophysical interaction data from experiments such as the data of chemical shift perturbation obtained from NMR titration or mutagenesis experiments [86]. This is the so-called ambiguous interaction restraints (AIR) docking, which introduces the restraints to each of the residues in the "active" state or the "passive" state according to those known experimental results. For example, the "active" residues corresponding to all residues in the case of NMR titration data shows a significantly chemical shift perturbation upon complex formation as well as a highly solvent accessibility in the protein [87]. HADDOCK requires structural information from PDB files of the individual proteins and ambiguous interaction restraints from corresponding experiments. Three main stages are included in the docking process. First, randomizing the orientations of the interaction partners and minimizing the rigid body energy of the interactions. Second, semi-rigid annealing is simulated in torsion angles. Thirdly, Cartesian space of the docked complex is refined in the explicit solvent. After that, these targeted structures are ranked according to their intermolecular energies including electrostatics (E_{elec}), van der Waals (E_{vdw}), AIR energy terms (E_{AIR}) and average buried surface area (E_{BSA}). Finally, newly built structures with a high score of energetically ranking are clustered by using the pairwise-based backbone RMSD alignment according to those residues on the binding interface of the complex.

In this thesis, as we mentioned the second question in section 1.7, the precise mechanism of the up-to-down lever swing in the power stroke is still unclear due to

limited information of weakly bound form of actomyosin complex. According to this question, we try to build up an actomyosin complex in the weak-binding state to discover what kinds of subdomains or residues are involved in the weak actin-binding surface. Meanwhile, the actomyosin complex in the strong actin-binding state is also necessary to be created, which is not only regarded as a reference by comparing to the Holmes's rigor model to evaluate the results of our docking process, but it also can be considered as the end state of the power stroke by comparing to the beginning state (weak actin-binding state) to discriminate the function of residues on the binding surface of myosin.

2.3 Molecular dynamic simulations

Molecular dynamics (MD) is a combination of physics, mathematics and chemistry. MD is a molecular simulation method, which mainly relies on Newtonian mechanics to simulate the movement of the molecular system. Samples of the molecular system were collected in an ensemble constituted by the different states. Taken samples are used to calculate the configuration integral of the system. Furthermore, thermodynamic quantities and other macroscopic properties of the system can be calculated based on the results of the configuration integral [88]. MD as a comparable mean to experiment and theory is considered to be the "third way" in performing scientific research. MD can provide a number of valuable predictions, which are difficult to carry out by experimental approaches in some aspects [89]. Today, MD as a theoretical approach is possible to treat a system with millions of particles in the investigations of dynamic properties of structures under a variety of conditions [90]. Application of MD is ranged from studies of protein-protein binding, enzyme reaction mechanisms, refinement of structures to the analysis of the transition pathways.

The fundamental idea of MD simulations is based on the theory of statistical mechanics, which is used to investigate and predict macroscopic phenomena from the properties of individual molecules in the system [91, 92]. Each of the macroscopic system can be described by its thermo-dynamical state and represented by parameters

of temperature (T), pressure (P), and the number of particles (N). However, we can describe all possible macroscopic systems or thermodynamic states with different microscopic states, which are the so-called ensembles. The thermodynamic state referred to as the microcanonical ensemble (NVE) is characterized by a fixed number of atoms (N), a fixed volume (V), and fixed system energy (E), corresponding to an isolated system. A process may be seen as an exchange of potential and kinetic energy, with total energy being conserved. The canonical ensemble (NVT) is marked by a fixed number of atoms (N), fixed volume (V), and fixed temperature (T), which this simulation is sometimes called constant temperature molecular dynamics (CTMD). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat. In addition, the isobaric-isothermal ensemble (NPT) is defined by a fixed number of atoms (N), fixed pressure (P), and fixed temperature (T) [93]. In NPT, besides the thermostat, a barostat is considered, which is analogous to most of the laboratory conditions with a flask exposed to ambient temperature and pressure [94]. Besides that the forced MD is quite useful in revealing conformational changes in a protein at the atomic level based on applying forces to a structure by pulling it along desired degrees of freedom. Two typical ways of forced MD are popular used by now, which are pulling velocity constant and pulling force constant [95]. Umbrella sampling as one of the most prevalent forced MD move the structure along the desired reaction coordinate by distances, angles and so on [96]. All of the transition configurations in high or low energy will be adequately sampled in order to calculate the free energy changes by the potential mean force [97].

In order to start an MD simulation, an initial structure is needed from certain existing online services, e.g., the protein data bank (PDB). Atomic coordinates of the structure can be resolved by crystallization, NMR or other similar techniques, but some missing coordinates need to be fixed by utilizing molecular modeling approaches. The potential energy of molecules in this state contains a number of minima or sub-state of conformations in their global energy landscape and spatial constraints. The potential energy typically at the local minimum has a higher value than the potential energy of the global minimum. However, generally, the local minimum

energy states are stable and can represent the probable state of a structure interest in a statistical system. Energy minimization was able to perform to optimize the initial structure into a realistic relaxed conformation by eliminating existing spatial constraints [98]. A few hundred or thousand steps of minimization are typically calculated depending on the size of the system. A combination of different algorithms for the minimization has been proved to be the best choice for most of biological molecules. Two prominent algorithms are the steepest descent (SD) and conjugate gradient (CG) algorithms in the energy minimization [99, 100]. The SD moves Cartesian coordinates of all atoms in a direction parallel to the downhill. The successive steps of the SD are orthogonal in both gradients and the direction. The SD as a robust algorithm is able to find the local minimum even if the starting point is far away from the minimum. If the minimum lies in a long and narrow valley, the SD will perform many small steps to correct the previous errors of movements, since it has an inflexible determination on the searching direction path. Distinct from the SD, the sequential steps of the CG are orthogonal in gradients, but conjugate in directions. These directions have property for a quadratic function of variables, which make the minimum easy to be reached by steps.

Numerous software packages e.g., AMBER, CHARMM and GROMACS are capable of performing energy minimization and MD simulations [101-103]. AMBER is the collective name for a suite of programs that allows users to carry out molecular simulations, particularly for proteins, nucleic acids, carbohydrates and so on. [101]. Three main steps of principal flow contained in AMBER are preparation and minimization of the candidate molecules, MD simulation, and MD trajectory analysis. Autonomous operations are able to be carried out in separate programs with discrete encoding-codes, so that AMBER has its own distinctively important advantages. It not only permits separate program modules to be upgraded individually and to be written with different coding practices, but also allows the principal simulation codes (e.g., Sander and PMEMD) in AMBER are able to be compiled to improve the computing power for parallel operations. The main preparation programs for the candidate molecules are the antechamber and the LEaP. Result of preparation process

contains a coordinate file and a parameter-topology file, which are the basic requirements of the input preparation for each of the individual molecules or proteins to start their calculations and analyses. The Cartesian coordinates of all atoms in molecules are the main information in the coordinate file. Further information including atom names and masses, force field parameters, lists of bonds, angles, dihedrals and so on needed to compute energies and forces are contained in the parameter-topology file. The main calculation module for MD simulation is called Sander, which can be compiled parallel by using the MPI programming interface and coded by Fortran 90. Each processor deals with certain atoms, but all processors know the coordinates of all atoms. Every update step of certain atoms is performed by their own processors in the MD simulation. Meanwhile, a binary tree to communicate the updated positions to all processors is used in preparation for the next update step.

All of the results can be analyzed according to different modules in AMBER on the basis of the MD trajectories. The ptraj analysis module is composed of self-designed scripts, which support analyses for dynamic properties, conformational changes and so on. For instance, the script for the covariance matrix is extremely useful to estimate the dynamic correlation properties e.g., analyzing the possible pathways for the conformational changes of structures.

In addition, the transition pathway of conformational changes is quantified with a series of images connected to each other by virtual "springs" which keep the images from sliding down the energy landscape onto adjacent images that is can be carried out by the module of the nudged elastic band (NEB) in AMBER [104]. The NEB method derived from previous plain elastic band method provided by Karplus et al. [105] adds the spring forces to the potential energy surface and minimized the energy of the system. Advantage of the NEB method is that the corner cutting problem in the energy landscape of the plain elastic band method is prevented by truncating the spring forces in directions perpendicular to the tangent of the path. Implementation of NEB module in AMBER requires no hypothesis for a starting path. However, shrewd judgment of temperature and length of time are essential to populate the minimum energy path. Usually, the structure-based transition pathway of protein is not enough

to explore the potential mechanism of the conformational changes combined with related energy profile. Therefore, a free energy quantity is able to calculate potentials of mean force by using the umbrella sampling within AMBER [106]. In the umbrella sampling module, an artificial restraint is exerted to the system by biasing it to sample possible coordinates in a defined range of value. Values' distribution of these coordinates during the simulation should be recorded. After the NEB calculation for the structural transition of the recovery step, the entire pathway was divided into 29 small 'windows', which were considered as the coordinates' ranges of our interest in the umbrella sampling. Individual simulations were carried out at these 'windows' with coordinates' overlapping to their neighbors. After each of these 'windows' achieved to their equilibration without biased force, a potential of mean force can be constructed via the WHAM method [107] to represent the free energy profile along the structural transition of the recovery step.

In order to distinguish the weakly and strongly bound actomyosin complex in two states, binding free energy can be calculated with the approach of molecular mechanic Poisson-Boltzmann surface area (MM/PBSA) [108]. Binding free energy is calculated for each of the components with different equations shown as below. First, the general equation for the binding free energy is calculated on the basis of the equation (1):

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \qquad (1)$$

Binding free energy is generally contributed by enthalpy and entropy, (equation 2), which are evaluated as a sum of changes in the molecular mechanics (G_{MM}) upon the actomyosin complex formation, solvation free energy (ΔG_{sol}), and entropy ($-T\Delta S$) (equation 3).

$$\Delta G_{bind} = \Delta H - T\Delta S \qquad (2)$$

$$\Delta H = \Delta G_{MM} + \Delta G_{sol} + T\Delta S \qquad (3)$$

Molecular mechanical energy (ΔG_{MM}) is calculated on the basis of the equation (4),

where the ΔG_{MM} can be further divided into a van der Waals (ΔG_{vdw}) and a Coulomb term (ΔG_{ele}).

$$\Delta G_{MM} = \Delta G_{vdW} + \Delta G_{elec} \qquad (4)$$

The solvation free energy (ΔG_{sol}) is separated into a polar (electrostatic contribution to solvation free energy ($\Delta G_{sol\text{-}pol}$)) and a non-polar part (non-polar contributions to the solvation free energy ($\Delta G_{sol\text{-}np}$)) (equation 5).

$$\Delta G_{solv} = \Delta G_{sol\text{-}pol} + \Delta G_{sol\text{-}np} \qquad (5)$$

The pbsa module in AMBER is used to evaluate the electrostatic contribution to the solvation free energy ($\Delta G_{sol\text{-}pol}$).

In this thesis, as we mentioned above in section 2.3, the weak and strong actin-binding states of the actomyosin complex have been built up by the docking process. Trajectories abstracted from normal MD simulation for both states can be analyzed by binding free energy calculations (MM/PBSA) to evaluate the discrimination between the weak and strong actin-binding states quantitatively. The main purpose of this project is to discover the relationship of the myosin lever's back and forth swing between the recovery step and the power stroke. The umbrella sampling and nudged elastic band method (NEB) as forced MD approaches enables to figure out the down-to-up lever swing related structural changes and potential intermediate states in the recovery step. The dynamic properties of two end-states and intermediate states e.g., conformational changes and inter-motional correlations can be determined via analysis script based on the MD trajectories. Combining structural changes with dynamic properties of each state, we can propose structural pathways and possible mechanism for the recovery step and the actin-binding induced myosin up-to-down lever swing in the power stroke.

According to the addressed questions in the first chapter and mainly applied *in silico* approaches to these questions in this chapter, I draw a flow chart of corresponding

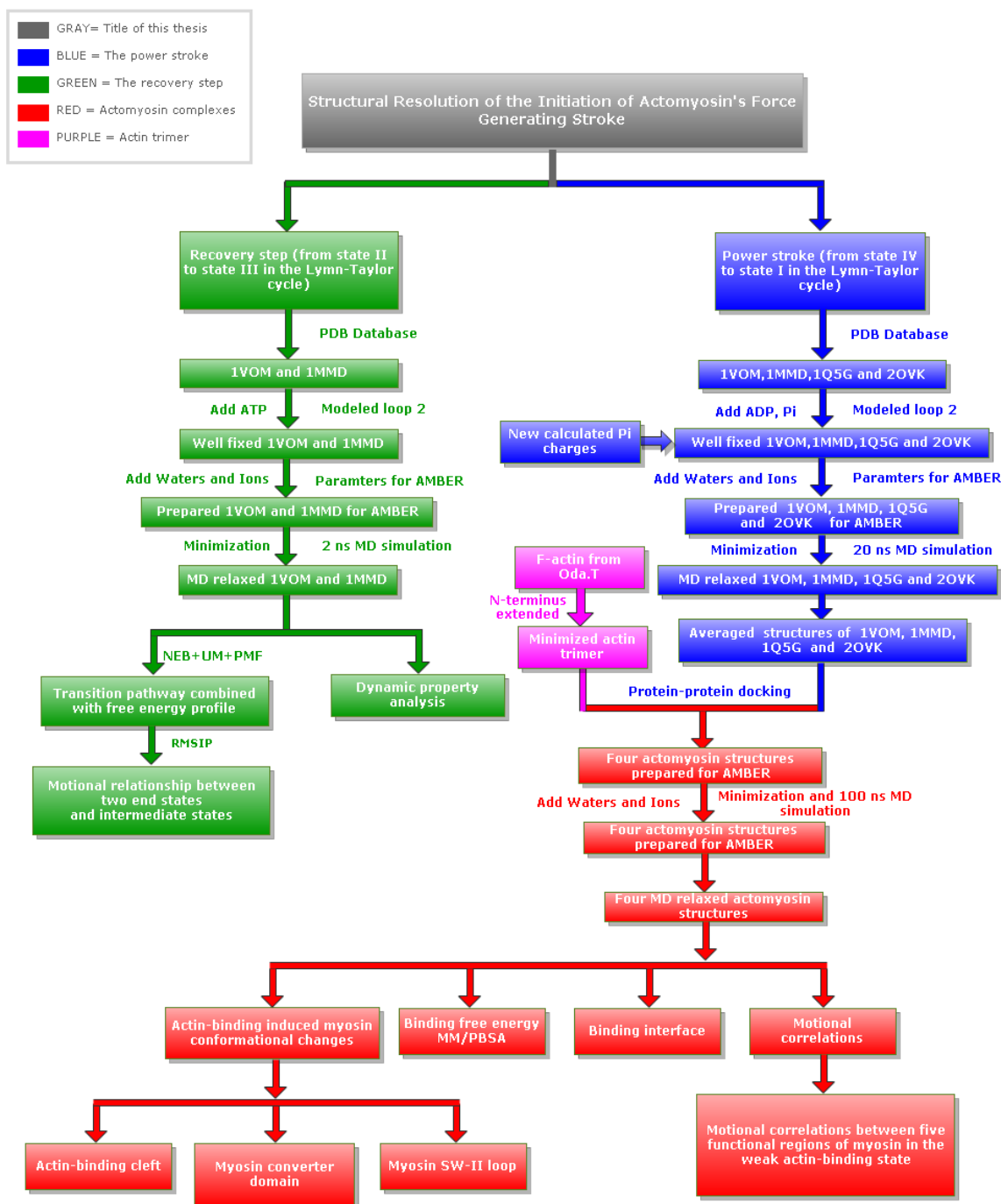procedures regarding to this thesis listed as follows (see Figure 7).

Figure 7: Flow chart of corresponding procedures regarding to this thesis.

CHAPTER 3: Materials and methods

3.1 Protein structures preparation

According to the procedure of the recovery step in Figure 7, we modeled two end states of the recovery step based on the *Dictyostelium* myosin II motor domain from the PDB database. The starting structure of the recovery step is 1MMD (PDB entry) in open actin-binding cleft and down-lever state with missing residues, which was fixed according to the molecular modeling with a detailed description in the next section [109]. The end structure is 1VOM (PDB entry) [110] in open actin-binding cleft and up-lever state with its missing parts referred to the structure M754 determined and kindly provided by Jon Kull [43] (see Table 3). Mg·ADP analogue as a substrate in the nucleotide-binding pocket was converted to Mg·ATP based on the conformational overlapping for both of the end states. Parameters of ATP and ADP in the AMBER force field were kindly contributed from Carlson *et al.* [111]. In order to discover the transition mechanism of the recovery step, mutants F481A/F482A and F652A were introduced to the structures of 1VOM and 1MMD, respectively.

According to the procedure of the power stroke in Figure 7, we modeled four myosin structures and the actin trimer. We used 1VOM as the starting structure of myosin in the power stroke for docking to the actin trimer. 1MMD with open actin-binding cleft and down-lever arm was also used for docking to the actin trimer as one of the possible transition state. Mg·ADP analogue as a substrate in the nucleotide-binding pocket is converted to Mg·ADP·Pi based on the conformational overlapping for the starting structure (1VOM) and the transition structure (1MMD) of myosin in the power stroke. The end state of the power stroke in closed actin-binding cleft, down-lever and nucleotide-free state was represented by 1Q5G (PDB entry) from *Dictyostelium* myosin II regarded as rigor-like state [37]. 2OVK (PDB entry) from squid myosin was modeled in the absence of nucleotide, closed actin-binding cleft and down-lever state also to be deemed as a rigor-like structure [112] (see Table 3). The structure of actin trimer for docking to myosin was refined from F-actin with its

missing N-terminus extended and acetylated [35]. In order to discover the actin-binding effect on the conformational changes of myosin, R520Q, R562Q were introduced to the actin trimer docked 1VOM 1MMD, 1Q5G and 2OVK actomyosin complexes.

Table 3: Crystal structures of myosin II used in this project

PDB entry	Resolution [Å]	Construct[1]	Ligand	State[2]	Scientific Name
1VOM	1.9	762	MgADP.VO$_4$	C/C	*Dictyostelium discoideum*
1MMD	2	762	Mg.ADP.BeF$_3$	C/O	*Dictyostelium discoideum*
1Q5G	1.9	771	-	O/O	*Dictyostelium discoideum*
2OVK	2.6	839	-	O/O	*Loligo pealei*

[1] The last resolved amino acid is given. [2] The conformational states of SW-I/SW-II loop are indicated. "O" refers to the open state, whereas "C" refers to the closed state.

3.2 Preparation of the nucleotide-binding pocket of myosin

The proportion of $H_2PO_3^-$, as the most dominant ionized state at pH=7.2 has been verified in quantum mechanical calculations under the CHARMM force field [113, 114]. We calculated the partial charges and coordinates, which can be used in the AMBER force field. RESP charges for each of the atoms in the $H_2PO_3^-$ group were calculated by using the Gaussian 03 [115] and ANTECHAMBER [116]. The calculations were carried out for a set of molecular conformations of the $H_2PO_3^-$ group by covering the accessible space in the system. The density functional algorithm with the B3LYP exchange and the 6-312+g(d,p) basis set were selected for calculation in the Gaussian 03. The resulting charges were averaged over all of the conformations. Charges of the oxygen atoms (O2 and O3) and hydrogen atoms (H1 and H2) were additionally averaged due to the local symmetry. RESP charges of each atom in the $H_2PO_3^-$ group are listed in the Table 4. By combining with ANTECHAMBER to calculate the corresponding parameters of protonated phosphate for the AMBER usage, the detailed steps can be ascertained in Figure 8.

Table 4: Atomic charges of the H$_2$PO$_3^-$ group

Atom No.	Atom Name	Atom Type	RESP Charge
1	O1	O2	-0.81334
2	P1	P	1.28996
3	O2	OH	-0.67792
4	H1	HO	0.34627
5	O3	OH	-0.67792
6	H2	HO	0.34627

RESP charges of the H$_2$PO$_3^-$ group were used for parameterization in the AMBER force field. Two oxygens (O2 and O3) and two hydrogens (H1 and H2) of phosphate were averaged assumed to the local symmetry.

```
Step 1:
*Generate a mol2 file for the protonated phosphate (Pi.mol2) in gaussian03.
Step 2:
*Generate a gaussian input file: Pi.gau from a mol2 file Pi.mol2
>antechamber -fi mol2 -fo gzmat -i Pi.mol2 -o Pi.gau (Running in AMBER)
Step 3:
*Run gaussian03 to get the output file (Pi.gout), the detailed setup can be seen in section 3.2.
Step 4:
*Generate an ac file Pi.ac from gaussian output file Pi.gout
>antechamber -fi gout -fo ac -i Pi.gout -o file Pi.ac -c resp (Running in AMBER)
Step 5:
*Extract resp charge to a charge file Pi.crg from Pi.ac
>antechamber -fi ac -i Pi.ac -c wc -cf Pi.crg (Running in AMBER)
Step 6:
*Read in Pi.crg and Pi.mol2 and generate an ac file Pi-1.ac
>antechamber -fi mol2 -i Pi.mol2 -c rc -cf Pi.crg -fo ac -o Pi-1.ac (Running in AMBER)
Step 7:
*Judge the atom type using "atomtype".
>atomtype -i Pi-1.ac -o Pi-1_gaff.ac -p gaff (Running in AMBER)
Step 8:
*Generate residue topology file Pi-1_gaff.prepc using "prepgen".
>prepgen -i Pi-1_gaff.ac -o Pi-1_gaff.prepc -f car (Running in AMBER)
Step 9:
*Use parmchk to check the missing force field parameters and generate additional force field
parameter file (frcmod) Pi.frcmod
>parmchk -i Pi-1_gaff.prepc -o Pi-1.frcmod -f prepc (Running in AMBER)
Step 10:
*With Pi-1_gaff.prepc and Pi.frcmod, we can use leap to generate topology files for AMBER.
>loadamberparams $AMBERHOME/dat/leap/parm/gaff.dat
>loadamberparams Pi-1.frcmod
>loadamberprep Pi-1.prepc
>saveamberparm PI PI.prmtop PI.prmcrd
```

Figure 8: Corresponding parameters of protonated phosphate were calculated

on the basis of ANTECHAMBER.

In addition, all crystal waters in the structure of 1VOM and 1MMD were stripped except two crystal waters formed interactions with magnesium ions in the nucleotide-binding site (see Figure 9). When these two crystal waters were stripped from the nucleotide-binding pocket, the conformation of ADP and phosphate were destroyed in the MD simulations.
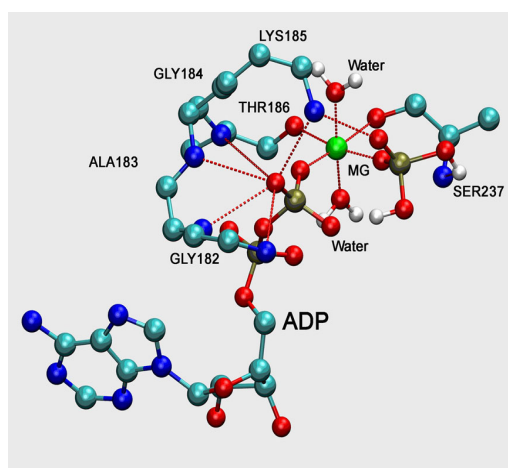

Figure 9: Hydrogen-bond network in the nucleotide -binding pocket of myosin. Two crystal waters formed interactions with magnesium, which contacted with residue Thr186, Ser237, β-phosphate of ADP and the $H_2PO_3^-$ group. The hydrolyzed γ-phosphate was in a protonated state with one hydrogen bond formed with β-phosphate of ADP and an electrostatic effect on the magnesium.

3.3 Comparative modeling of the myosin loop 2 in 1MMD

The amino acid sequence of the structure 1MMD was retrieved from the NCBI protein database. The BLASTP method was used to search for suitable templates from the protein structure database (PDB) for the missing loop 2. Three homologous *Dictyostelium* myosin II structures with higher evaluation scores were selected for the alignment of loop 2. These were 2AKA and 1G8X in the up-lever states (wild type) and 1W9L in the down-lever state (mutant, not wild type). Sequence alignments of loop 2 were carried out with the ClustaIW algorithm based on the pairwise sequence-sequence comparison by using the Gonnet series matrix with the "gap open" and "gap elongation" penalties of 10 and 0.2 respectively [109]. The result was refined to ensure the best alignment in these conserved residues of the myosin loop 2 through the JOY program [110]. Loop 2 started and ended with relatively stable alpha-helices was a hairpin-like β-strand conformation with four polar and some

solvent-accessible residues (see Table 5).

Table 5: Sequence alignment and analysis for the *Dictyostelium* myosin loop 2

615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634
p	n	I	*a*	s	r	*a*	k	k	g	*a*	n	*f*	i	*t*	*v*	*A*	*a*	q	Y
a	a	a	a				b	b			b	b			a	a	a	a	a

The JOY annotation is as follows: lowercase red letter: α-helix, lowercase blue letter: β-strand, uppercase letter: solvent-inaccessible residue, lowercase letter: solvent-accessible residue, bold: hydrogen bond to main-chain amide, underline: hydrogen bond to main-chain carbonyl. The *italic* letters represent loop 2 region. "a" represent the α-helix and "b" represent the β-sheet.

Based on the selection and refinement of these templates, the complete 3-D structure of the 1MMD was constructed according to the coordinate reconstruction of segment matching by MODELLER 9.2v and an integrated analytical front-end application for bioinformatics called "Friend" [68, 119]. When we opened the graphical interface of the "Friend" program, three template structures and sequence needed to load. 1MMD sequence was loaded by using the menu "File->Load->Alignment from text box" followed by paste of sequence in fasta format into the poped-up form. Three templates were loaded by using the menu "File->Load->Structure from PDB" followed by typing a 5-letter code in the poped-up form. Sequence alignment was prepared manually by using menu "Tools->Pairwise->Align". Then, in order to deal with the modeling step, we used the menu "Tools->MODELLER" with names of three templates and sequence marked. Interface of MODELLER as a pop-up window with many spaces were required to fill out. We specified the three templates and sequence for our modeling with hydrogen and without crystal waters. Ten models were built from three templates and refined according to the scoring function. Validations of the models were carried out after the refinement process by using Ramachandran Map. Two angle distributions (Φ and ψ) of modeled 1MMD and loop 2 with the highest modeling score are in the left and right panels (see Figure 10). Loop 2 shows its 92.9% residues are in a favored region while the 1MMD is in an acceptable conformation with 88.1% residues in a favored region.
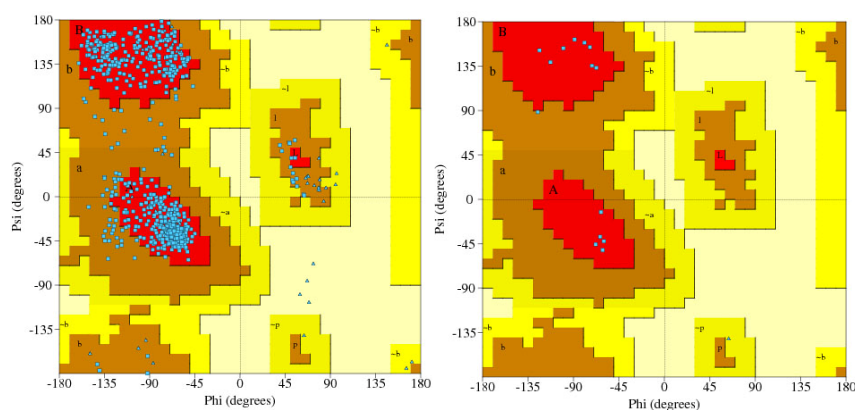
Figure 10: Comparative modeling of the myosin loop 2. In the left panel, the myosin motor domain (762 residues) was analyzed with 88.1% residues in a favored region, 11.8% residues in an allowed region and 0.1% residue in an outlier region. In the right panel, loop 2 (14 residues) was analyzed with 92.9% residues in a favored region, 7.1% residues in an allowed region and none of residues in an outlier region.

3.4 Protein-protein docking

Four averaged myosin structures (1VOM·ADP·Pi, 1MMD·ADP·Pi, 1Q5G and 2OVK) after 20 ns MD simulations were docked to the refined actin trimer (Ac-DEDE-actin) by using the docking program of HADDOCK V2.1 [84]. (A) 3A-1VOM$_{dock}$: Actin trimer docked to 1VOM·ADP·Pi (open actin-binding cleft and up-lever) as the beginning state of the power stroke (weak actin-binding state). (B) 3A-1MMD$_{dock}$: Actin trimer docked to 1MMD·ADP·Pi (open actin-binding cleft and down-lever) as one of the possible transition states of the power stroke (weak actin-binding state). (C) 3A-1Q5G$_{dock}$: Actin trimer docked to 1Q5G (closed actin-binding cleft and down-lever) as the end state of the power stroke (strong actin-binding state). (D) 3A-2OVK$_{dock}$: Actin trimer docked to 2OVK (closed actin-binding cleft and down-lever) as the end state of the power stroke (strong actin-binding state), which as the squid myosin was different from the *Dictyostelium* myosin in the group (A), (B) and (C).

Those experimentally defined myosin-actin interface residues regarded as the flexible segments were divided into the active and the passive amino acids for protein-protein docking (see Table 6). Active residues of the *Dictyostelium* myosin loop 2 (residues 619-629) and the lower 50 kDa domain (residues 519-555) defined from experiments

were selected for the actin myosin binding [120-122]. Ambiguous interaction restraints (AIR) identified from residues was used to drive the docking process. Overall, six active and ten passive residues were selected from myosin, five active and five passive residues were selected from actin (see Table 6). Ambiguous interaction restraints (AIR) with a maximum effective distance of 8.0 Å was available for the active-active and active-passive residues, but not for the passive-passive residues. The best 200 of 1000 refined complexes were obtained from rigid body energy minimization and then submitted to the semi-rigid simulated annealing process. These structures were exposed to the 12 Å shell of TIP3P water solvent for MD simulations with a cutoff value of 5 Å. Calculations of the non-bonded energy were set up with a cut-off value of 9 Å in the OPLSX force field [123]. The best 100 complexes were selected according to the evaluation score of their average interaction energies and buried surface area. The HADDOCK score, as the main criteria for selection, was calculated on the basis of the equation 6. Finally, subsets were collected according to more than ten members of the complex in each of them after the backbone-based RMSD clustering with a cut-off value of 7.5 Å. All the above-mentioned parameters and processes were carried out through the HADDOCK online service, http://haddock.science.uu.nl/services/HADDOCK/haddock.php.

$$HADDOCK_{SCORE}=E_{VWD}+0.2E_{ELEC}+0.1E_{AIR+} E_{DESO} \qquad (6)$$

Table 6: Intermolecular restraint residues for HADDOCK docking

Ambiguous interaction restraints	
Myosin motor[1]	
Active residues[A]	**Lys622, Lys623, Val534, Phe535, Pro536, Ala538**
Passive residues[B]	**Gly621, Thr629, Glu530-Ser533, Thr539, Thr542**
Flexible segment[C]	**Gln621 to Lys629,**
Actin trimer	
Active residues	**Glu24, Glu25, Leu142, Ala144**
Passive residues	**Gly23, Lys28, Ser141, Ser145, Gly146,**
Flexible segment	**Gly23 to Lys28, Arg147 to Gly150**

[1] The residue numbers of myosin motor represent the *Dictyostelium* myosin. The residue numbers of actin are the same for two myosin species. [A] Active residues

represent those residues involved in the actin myosin binding. [B] Passive residues corresponding to all surface neighbors of the active residues were in the solvent accessible. [C] Flexible segments were defined with those regions connected to active and passive residues.

3.5 Molecular dynamic (MD) simulations and dynamic property analysis

Structures of pre-recovery state (1MMD·Mg·ATP), post-recovery state (1VOM·Mg·ATP) and four mutants (1MMD·Mg·ATP$_{F481A/F482A}$, 1MMD·Mg·ATP$_{F652A}$, 1VOM·Mg·ATP$_{F481A/F482A}$, 1VOM·Mg·ATP$_{F652A}$) were relaxed in the MD simulations. We named them as 1MMD$_{wild}$, 1VOM$_{wild}$, 1MMD$_{F481A/F482A}$, 1MMD$_{F652A}$, 1VOM$_{F481A/F482A}$ and 1VOM$_{F652A}$ in the subsequent chapter 4. Meanwhile, four myosin structures 1VOM·Mg·ADP·Pi, 1MMD·Mg·ADP·Pi, 1Q5G and 2OVK were relaxed in the MD simulations and averaged before docking to the actin trimer, which were named as 1VOM$_{md}$, 1MMD$_{md}$, 1Q5G$_{md}$ and 2OVK$_{md}$. Four actomyosin complexes obtained from the protein-protein docking were relaxed in the MD simulations, which were named as 3A-1VOM$_{md}$, 3A-1MMD$_{md}$, 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$. In addition, mutants of R520Q, R562Q were introduced to these four relaxed actomyosin complexes and relaxed via the MD simulations. These mutants were named as 3A-1VOM-520MU, 3A-1VOM-LP3MU, 3A-1MMD-520MU, 3A-1Q5G-520MU and 3A-2OVK-520MU. There were six structural trajectories as the first set for recovery step analysis and thirteen structural trajectories as the second set for the power stroke analysis.

The module of sander in the molecular dynamics package AMBER11 (compiled under NVIDIA card Tesla S2050) was used for minimization and molecular dynamic (MD) simulations [101]. All the above-mentioned structures were neutralized by adding discrete number of Na$^+$ ions in the most appropriate electro-negative areas around proteins according to different size or conformations of structures. After that, they were solvated in a truncated octahedron box of TIP3P water environment with a 12Å buffer cut-off value along each dimension [124]. Long-range electrostatic interactions were treated by the particle-mesh Ewald method in periodic boundary conditions [125]. Energy minimization was processed with 1000 steps of steepest

descent and conjugate gradient. The system was slowly heated up to 300 K and equilibrated at the constant temperature (NVT) and the constant pressure (NPT) conditions. The Berendsen coupling algorithm [126] was used for the temperature control. The SHAKE algorithm was applied for constraints on covalent bonds and all hydrogen atoms [127]. The root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) analyses were carried out by using ptraj module in AMBER and visualized with VMD graphical software [128]. Detailed general input files and scripts for minimization and simulation processes can be referred to Figure 11 and 12.

In the first set (recovery step analysis), after 1ns NVT and 1ns NPT, the final MD trajectories of six structures were collected via extra 2 ns NPT equilibration. Torsional angles mobility of structures was calculated by averaging the coordinates of 125 structures from the final 250 ps long equilibrium phases. We determined the amplitude (δ) of torsional mobility of the Φ, Ψ angles according to the equation (7):

$$\delta = \sqrt{\frac{\sum (c - \bar{c})^2}{n-1}} \tag{7}$$

The "c" is the actual torsion angle and the "$\bar{c}$" is the average of the Φ or Ψ angle of the given residue. The "n" is the number of data points. Structures were averaged on the basis of MD trajectories via scripts in Figure 12. This equation was compiled into an executive command and kindly offered by my colleague Zahoránszky KG. In the second set (power stroke analysis), after 5 ns NVT and 5 ns NPT, the final MD trajectories of four myosin structures were collected via 20 ns NPT equilibration and nine actomyosin complexes (four wild types and five mutants) were collected via 100 ns NPT equilibration. The final average structure and motional correlations can be received according to the MD trajectories via scripts in Figure 12.

```
*****input file for the minimization (named min.in)*****
Minimization (1000 steps of SD and 1000 steps of CG)
 &cntrl
   imin=1, maxcyc=1000, ncyc=1000,
   cut=12.0, ntb=1, ntc=1,ntf=1 /
***********************************************************************
*****input file for the heating process (named heat.in)*****
Heating process (50 ps long)
 &cntrl
   imin=0,irest=0,ntx=1, ntc=2,ntf=2, cut=12.0, ntb=1, ntpr=1000, ntwx=1000, ntwr=1000,
   ntt=3, gamma_ln=1.0, tempi=0.0, temp0=300.0, nstlim=25000,dt=0.002 /
***********************************************************************
*****input file for the NVT process (named NVT.in)*****
NVT process (nstlim=500000 means 1ns (500000x0.002fs) long in NVT equilibration, which can
be modified according to requirements of various simulation time)
     &cntrl
   imin=0,irest=1,ntx=5, ntc=2,ntf=2, cut=12.0, ntb=1, ntpr=10000, ntwx=10000, ntwr=10000,
   ntt=3, gamma_ln=2.0, temp0=300.0, nstlim=500000,dt=0.002 /
***********************************************************************
*****input file for the NPT process (named NPT.in)*****
NVT process (nstlim=500000 means 1ns (500000x0.002fs) long in NPT equilibration, which can
be modified according to requirements of various simulation time)
     &cntrl
   imin=0,irest=1,ntx=5, ntc=2,ntf=2, cut=12.0, ntb=2, ntp=1, taup=2.0, ntpr=10000, ntwx=10000,
   ntwr=10000, ntt=3, gamma_ln=2.0, temp0=300.0, nstlim=500000,dt=0.002 /
***********************************************************************
```

Figure 11: General input files for minimization and molecular dynamic simulations.

```
***** Parameter and coordinates preparation for AMBER******
>tleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff99SB
>loadamberprep ADP.prepi              (ADP.prepi is offered from Carlson et al.)
>loadamberprep ATP.prepi              (ATP.prepi is offered from Carlson et al.)
> loadamberparams ADP.frcmod          (ADP.frcmod is offered from Carlson et al.)
> loadamberparams ATP.frcmod          (ATP.frcmod is offered from Carlson et al.)
> loadamberprep Pi.prepc              (Pi.prepi is calculated from Gaussian 03)
> loadamberprep Pi.frcmod             (Pi.frcmod is calculated from ANTECHAMBER)
>X=loadpdb ***                        (X is a self-defined variable, *** is the name of a
given structure in the pdb form)
>addions2 X Na+ 0                     (add Na+ ions to neutralize the given structure)
>solvateoct X TIP3PBOX 12.0           (solvate structure in water with a 12Å cut-off)
>saveamberparm X X.prmtop X.inpcrd    (save all into a parameter file and a coordinate file)
*********************************************************************
***** Commands to execute *****
> $AMBERHOME/exe/mpirun -np 100 sander.MPI -O -i min.in -o min.out -p X.prmtop -c
X.inpcrd -r min.rst
> $AMBERHOME/exe/mpirun -np 100 sander.MPI -O -i heat.in -o heat.out -p X.prmtop -c
min.rst -r heat.rst –x heat.mdcrd
> $AMBERHOME/exe/mpirun -np 100 sander.MPI -O -i NVT.in -o NVT.out -p X.prmtop -c
heat.rst -r NVT.rst –x NVT.mdcrd
> $AMBERHOME/exe/mpirun -np 100 sander.MPI -O -i NPT.in -o NPT.out -p X.prmtop -c
NVT.rst -r NPT.rst –x NPT.mdcrd
*********************************************************************
*****RMSD calculation (RMSD.trajin)*****
trajin NPT.mdcrd
rms first out rmsd.dat @N,CA,C time 1.0
> $AMBERHOME/exe/ptraj X.prmtop < RMSD.trajin > RMSD.out
*********************************************************************
*****RMSF calculation (RMSF.trajin)*****
trajin NPT.mdcrd
rms first @N,CA,C time 1.0
atomicfluct out rmsf.dat @C,CA,N byres bfactor
> $AMBERHOME/exe/ptraj X.prmtop < RMSF.trajin > RMSF.out
*********************************************************************
*****Average strucutre calculation (average.trajin)*****
trajin NPT.mdcrd
rms first mass @N,CA,C time 1.0
average average_NPT.pdb pdb
> $AMBERHOME/exe/ptraj X.prmtop < average.ptrajin > average.out
*********************************************************************
*****Motional correlation calculation *****
trajin NPT.mdcrd
rms first *
matrix correl name correlation :1-1880@CA,C,N out correlation.dat
*********************************************************************
```

Figure 12: Scripts for minimization, MD simulations and RMSD, RMSF, average structure and motional correlation calculation.

3.6 Transition pathway and potential mean force calculation of the recovery step

Two averaged structures named as 1MMD$_{wild}$ and 1VOM$_{wild}$ were used for exploring the transition pathway of the recovery step according to the nudged elastic band method (NEB) [104]. The cutoff value of the non-bonded atoms was 15 Å in 0.1 M salt concentration. SHAKE algorithm was used to constrain the position of hydrogen atoms with a time step of 1 fs in a generalized Born (GB) environment [127]. The maximum distance for the pairwise summation of calculating the effective Born radii was set to 15. Langevin dynamic was used to control temperature (collision frequency 1000 ps^{-1}) and a 1 ns simulated annealing was used to sample the transition path. The first step involved 100 ps of MD simulations with spring constants of 5 kcal·mol-1·Å^{-2} during the temperature of the system increased from 0 K to 300 K. A 200 ps MD simulation step was applied at constant temperature (300 K) during the spring constants gradually increased to 20 kcal·mol-1·Å^{-2}. Then the system was heated to 400 K in 200 ps and cooled to 0 K over the next 200 ps. Finally, 300 ps quenched MD simulation was performed at 0 K. Detailed general input files and scripts for NEB simulations can be referred to Figure 13 and 14.

Two end states of the recovery step (1MMD$_{wild}$ to 1VOM$_{wild}$) with 28 intermediates obtained from NEB simulations were used to build up 29 continuous windows with harmonic-restraints imposed on defined strains for the umbrella sampling. Bending was induced by using a quadratic biasing potential with a force constant of 1 kcal·mol^{-1}·degree^{-2}. Windows were sampled from 220.69° to 122.73° according to the angle represented by atoms of 684CA, 689CA and 748CA. Every window sampled for 1 ns after 200 ps NPT equilibration resulted in a total 29 ns production time in the transition path of the recovery step [106].

```
*****NEB input file (addles.in)*****
file rprm name=(str1.prmtop) read
file rcrd name=(str1.inpcrd) pack=2 read
file wprm name=(neb.prmtop) wovr
file wcrd name=(neb.inpcrd) wovr
action
~use original mass
omas
pimd
~make 30 copies of atom 1 to 12172 (the whole system of myosin in 12172 atoms)
space numc=30 pick #prt 1 12172 done
*EOD
*****NEB heating-1 process for 100 ps (heat-1.in)*****
Myosin recovery step NEB heating for 100 ps
 &cntrl
  imin = 0, irest = 0, ntc=1, ntf=1, ntpr=50, ntwx=500, ntb = 0, cut = 15.0, rgbmax=15.0,
  igb = 1, saltcon=0.1, nstlim = 200000, nscm= 0, dt = 0.0005, ntt = 3, gamma_ln=1000.0,
  tempi=0, temp0=300, ineb = 1,skmin = 5,skmax = 5, nmropt=1 /
 &wt type='TEMP0', istep1=0,istep2=200000,
   value1=0.0, value2=300.0 /
 &wt type='END' /
*****NEB equilibration process for 200 ps (equil.in)*****
Myosin recovery step NEB equilibrate for 200 ps
&cntrl
  imin = 0, irest = 1, ntx=5, ntc=1, ntf=1, ntpr=1000, ntwx=5000, ntb = 0, cut = 15.0,
rgbmax=15.0, igb = 1, saltcon=0.1, nstlim = 200000, nscm= 0, dt = 0.001, ntt = 3,
gamma_ln=1000.0, temp0=300, ineb = 1,skmin = 20,skmax = 20 /
*****NEB annealing process for 700 ps (anneal.in)*****
Myosin recovery step NEB annealing for 700 ps
&cntrl
  imin = 0, irest = 1, ntx=5, ntc=1, ntf=1, ntpr=5000, ntwx=10000, ntb = 0, cut = 15.0,
rgbmax=15.0, igb = 1, saltcon=0.2, nstlim = 700000, nscm= 0, dt = 0.001, ntt = 3,
gamma_ln=1000.0, temp0=300, ineb = 1,skmin = 20,skmax = 20, nmropt=1   /
 &wt type='TEMP0', istep1=0, istep2=100000,
   value1=300.0, value2=400.0 /
 &wt type='TEMP0', istep1=100001,istep2=200000,
   value1=400.0, value2=400.0 /
 &wt type='TEMP0', istep1=200001,istep2=300000,
   value1=400.0, value2=300.0 /
 &wt type='TEMP0', istep1=300001,istep2=400000,
   value1=300.0, value2=0.0 /
 &wt type='TEMP0', istep1=400001,istep2=700000,
   value1=0.0, value2=0.0 /
 &wt type='END' /
```

Figure 13: General input files for NEB simulations.

```
*****NEB commands to execute*****
>tleap -s -f $AMBERHOME/dat/leap/cmd/leaprc.ff99SB
>loadamberprep ATP.prepi              (ATP.prepi is offered from Carlson et al.)
>loadamberparams ATP.frcmod           (ATP.frcmod is offered from Carlson et al.)
>str1=loadpdb 1MMD_wild.pdb
>str2=loadpdb 1VOM_wild.pdb
>saveamberparm str1 str1.prmtop str1.inpcrd
>saveamberparm str2 str2.prmtop str2.inpcrd
>diff str1.prmtop str2.prmtop
>cat str2.inpcrd >> str1.inpcrd
>$AMBERHOME/exe/addles <addles.in >addles.out
>$AMBERHOME/exe/sander.PIMD -O -i heat-1.in -o heat-1.out -p neb.prmtop -c neb.inpcrd -r
heat-1.rst -x heat-1.mdcrd
>$AMBERHOME/exe/sander.PIMD -O -i equil.in -o equil.out -p neb.prmtop -c heat-1.rst -r
equil.rst -x equil.mdcrd
>$AMBERHOME/exe/sander.PIMD -O -i anneal.in -o anneal.out -p neb.prmtop -c equil.rst -r
anneal.rst -x anneal.mdcrd
**************************************************************************
*****PMF calculation commands to execute*****
>run_script
#!/bin/bash
Set number=(220.69 122.73)
foreach N ($number)
  sed –e 's/minimium/'$N/g' umbrella.in > input.in
>$AMBERHOME/exe/sander -O -i input.in -o umbrella_$N.out -p umbrella.prmtop -c
umbrella-1.rst -r umbrella_$N.rst -x umbrella_$N.mdcrd
>./histogram_multi hist_metadatafile reaction.histogram
>$WHAM/wham P 220.69 122.73 29 0.0001 300.0 0 metadatafile recovery_step.pmf > wham.log
**************************************************************************
*****RMSIP calculation commands to execute *****
> $AMBERHOME/exe/ptraj X.prmtop < RMSIP.trajin > RMSIP.out
**************************************************************************
```

Figure 14: Scripts for NEB simulations, umbrella sampling, the potentials of mean force calculation and RMSIP calculation.

To determine the potentials of mean force (PMF) for the studied system, all explored configurations from umbrella sampling were analyzed with the weighted histogram analysis method (WHAM). Unit conversion from radians to degree was indispensable by multiply the force constant (rk2) with 0.0006092 (= $2(\pi/180)$ 2) for the torsion restraints [107]. Detailed general input files and scripts for umbrella sampling, PMF calculation can be referred to Figure 14 and 15.

```
*****Umbrella sampling for 1 ns (umbrella.in)*****
Myosin recovery step umbrella sampling for 1 ns
&cntrl
nstlim=1000000, cut=15.0, igb=1, saltcon=0.1, ntpr=100, ntwr=4000, ntt=3, gamma_ln=0.2,
ntx=5, irest=1, ntc=2, ntf=2, tol=0.000001, dt=0.001, ntb=0, nmropt=1, /
&wt type='DUMPFREQ', istep1=4000 /
&wt type='END' /
DISANG= angle.RST
DUMPAVE=angle.dat
*****************************************************************************
*****Umbrella sampling defined angel file (angle.RST)*****
#angle restraints
&rst iat=684,689,748, r1=360., r2=220.69., r3=220.69., r4=360., rk2 = 1., rk3 = 1., /
*****************************************************************************
*****RMSIP calculation (RMSIP.trajin)*****
trajin IM-1.mdcrd
rms first *
matrix covar name covar :1-759@CA,C,N out IM-1-matrix.dat
analyze matrix covar out IM-1-covar.dat vecs 10
*The IM-1-covar.dat can be used to calculate the RMSIP through a script (named as
xpm2normal.csh) written by Tsjerk AW in the University of Groningen, Netherlands*
*****************************************************************************
```

Figure 15: General input files for the umbrella sampling and RMSIP calculation.

3.7 Dynamical similarity assessment by essential motional analysis

Quantitative characterizations of the dynamical properties of two intermediate states (IM-1 and IM-2) and two end states (1VOM$_{wild}$ and 1MMD$_{wild}$) were determined from essential motional analysis by using a covariance matrix in the recovery step based on the fluctuations of the backbone atoms. Eigenvectors and the eigenvalues were obtained in a descending order by diagonalizing covariance matrix σ [129]. Trajectories of altered states equilibrated in 2 ns MD simulations were compared by means of the root-mean square inner product (RMSIP), which was determined on the basis of the former ten eigenvectors. The motional similarity of different essential subspaces was assessed according to the equation (8), which has been written to a script by Tsjerk AW [130]. Detailed input files and scripts for RMSIP calculation can be referred to Figure 14 and 15.

$$\text{RMSIP} = \sqrt{\left(\frac{1}{10}\sum_{i=1}^{10}\sum_{j=1}^{10}\left(\boldsymbol{h}_i\cdot\boldsymbol{u}_j\right)^2\right)} \quad (8)$$

3.8 Binding free energy calculation on two end states of the power stroke

In order to distinguish the weak and strong actin-binding states, binding free energy was calculated with the approach of molecular mechanic Poisson-Boltzmann surface area (MM/PBSA). The averaged structures of actin trimer, myosin and actomyosin complexes were obtained after MD simulations. We used these three structures to create three gas phase prmtop and inpcrd file pairs for the MM-PBSA calculation as well as actomyosin for the solvated complex based on scripts of parameters and coordination prepared for AMBER in Figure 12. Total eight files generated for the MM-PBSA calculation can be seen in Figure 16. 2500 snapshots were collected through the last 20 ns of the total 100 ns MD equilibration in actomyosin complexes. The time scale between snapshots was set to 8 ps due to the limited requirement of at least 5 ps according to the persistence of motional correlations [131]. The pbsa module in AMBER [108] was utilized to evaluate the electrostatic contribution to the solvation free energy ($\Delta G_{sol\text{-}pol}$) (see equation 9). The grid spacing of the cubic lattice was set at 2 Å. The dielectric constant values for the interior and exterior were 1 and 80. One thousand linear iterations were performed for the linear Poisson-Boltzmann equation. The non-polar contribution to the solvation free energy ($\Delta G_{sol\text{-}np}$) was calculated from the solvent-accessible surface area ($_\gamma SASA$) [132] according to the equation (10). The surface tension γ and the offset β were set to 0.00542 kcal·mol^{-1}·Å^{-2} and 0.92 kcal·mol^{-1}, respectively.

$$\Delta G_{solv} = \Delta G_{sol\text{-}pol} + \Delta G_{sol\text{-}np} \qquad (9)$$

$$\Delta G_{solv\text{-}np} = _\gamma SASA + \beta \qquad (10)$$

```
*****MM/PBSA calculation (example for 3A-1VOM_md) *****
*****Eight files generated for the MM-PBSA calculation were list as below*****
(3A-1VOM.prmtop, 3A-1VOM.inpcrd, 3A.prmtop, 3A.inpcrd, 1VOM.prmtop, 1VOM.inpcrd,
3A-1VOM_solvated.prmtop, 3A-1VOM_solvated.inpcrd)
***********************************************************************
***** Extract the coordinates from last NPT trajectory. (extract_coords.mmpbsa)*****
@GENERAL: PREFIX:snapshot, PATH:./, COMPLEX:1, RECEPTOR:1, LIGAND:1, GC:1,
AS:0,       DC:0,       GB:0,       PB:0,       COMPT:./3A-1VOM.prmtop,       RECPT:./3A.prmtop,
LIGPT:./1MMD.prmtop, MM:0, MS:0, NM:0.
@MAKECRD:      BOX:YES,      NTOTAL:29673,      NSTART:1,      NSTOP:2500,      NFREQ:1,
NUMBER_LIG_GROUPS:1,   LSTART:17444,   LSTOP:29637,   NUMBER_REC_GROUPS:1,
RSTART: 1, RSTOP:17443.
@TRAJECTORY: TRAJECTORY :./NPT.mdcrd
***********************************************************************
***** Binding energy calculation. (binding_energy.mmpbsa)*****
@GENERAL:   PREFIX:snapshot,   PATH:./,   COMPLEX:1,   RECEPTOR:1,   LIGAND:1,
COMPT:./3A-1VOM.prmtop, RECPT:./3A.prmtop, LIGPT:./1MMD.prmtop, GC:0, AS:0, DC:0,
MM:1, GB:1, PB:1, MS:1, NM:1.
@PB:  PROC:2,  REFE:0,  INDI:1.0,  EXDI:80.0,  SCALE:2,  LINIT:1000,  PRBRAD:1.4,
ISTRNG:0.0, RADIOPT:0, NPOPT:1, CAVITY_SURFTEN:0.0072, CAVITY_OFFSET:0.00,
SURFOFF:0.00, SURFTEN:0.0072.
@MM: DIELC:1.0.
@GB: IGB:2, GBSA:1, SALTCON:0.0, EXTDIEL:80.0, INTDIEL :1.0, SURFTEN:0.0072,
SURFOFF:0.
@NM: DIELC:4, MAXCYC:1000.0, DRMS:1.
@MS: PROBE:0.0
***********************************************************************
*****MM/PBSA commands to execute*****
$AMBERHOME/exe/mm_pbsa.pl extract_coords.mmpbsa > extract_coords.log
$AMBERHOME/exe/mm_pbsa.pl binding_energy.mmpbsa > binding_energy.log
***********************************************************************
```

Figure 16: General input files and scripts for MM/PBSA calculation.

The contribution of entropy (*-TΔS*) based on the ligand receptor association was performed with normal-mode analysis [133]. 1000 snapshots were collected and submitted to the molecular minimization with a distance-dependent dielectric constant E=4r. The end of convergence was not achieved until the RMSD of the gradient vector less than 1×10^{-4} kcal·mol^{-1}·Å^{-1}. Residues close to the binding surface with a distance less than 20Å were acquired to estimate the contribution of entropy. Detailed input files and scripts for MM/PBSA calculation are referred to Figure 16.

Experimental Section

CHAPTER 4: Functional role of the pivot in the seesaw mechanism of the myosin lever swing in the recovery step

4.1 Introduction

As described by the Lymn-Taylor cycle [24] (see Figure 4), three alternating events are well organized to produce the mechanical force. These are actin binding, myosin allosteric conformational changes and the ATP hydrolysis. When myosin moves its lever arm from the pre-recovery (down lever position) to the post-recovery (up lever position) orientation in the ATP-bound state, it undergoes a large conformational transition called the recovery step (from state II to state III in the Lymn-Taylor cycle). If the myosin moves its lever arm from the up (post-recovery) to the down (pre-recovery) position, which is called the reverse recovery step and occurred in the actin detached state. Besides that, the former occurs in the ATP-bound state, whereas the latter occurs in the ADP·Pi-bound state, which is the rate limiting step of the enzymatic cycle in the absence of actin [63]. Actin binding accelerates the rate of up-to-down lever swing in ADP·Pi-bound state of myosin by two orders of magnitude resulting in the occurrence of the power stroke. Actin binding, however, contributes very little to the rate constant of the conformational changes of myosin in the ATP-bound recovery step [134].

In order to unveil the mechanism of the myosin lever swing in the recovery step, the basic idea is to find out the related conformational changes in different functional regions of myosin. Until now, myosin II has been crystallized with ATP or different ADP analogues in both pre-recovery and post-recovery states, which are assigned to the state II and state III in the Lymn-Taylor cycle [39]. The most significant differences of conformation between these two states are showed in the SW-II loop

and the orientation of the lever arm, which is linked to the converter domain. From pre-recovery state to post-recovery state, the conformation of the SW-II loop moves from a partially open to a closed state that switches on the catalytic function of ATPase. Meanwhile, the converter domain rotated by ~60° relative to the rest of the myosin head is so as to push the lever arm from the down to the up position. To reveal the conformational transition of the recovery step, it is necessary to figure out the intermediate states along the transitional pathway. Unfortunately, these intermediates are not easily to be achieved and investigated through experiments.

Numerous *in silico* simulations have been performed in order to reveal the mechanism of the recovery step [39, 113, 114, 135, and 136]. Fischer *et al*. proposed a theoretical model based on the unconstrained minimum-energy pathway simulations [39]. Fischer's model elaborated a mechanism that the relay-helix movement in a seesaw model coupled with conformational changes of the SW-II loop and swing of the lever-arm was preliminary characterized by structural transition pathway. The closure of the SW-II loop pulls the N-terminus of the relay-helix with a hydrogen bond cluster formed in the γ-phosphate~Gly457~Asn475. This formed cluster is supported by a fulcrum composed of two hydrophobic residues in the relay-helix (Phe481, Phe482) and a neighbor residue Phe652, which of this phenylalanine cluster serves as a pivoting point in the middle of the relay-helix. This seesaw-like motion induces a movement of the SW-II loop to the C-terminus of the relay-helix and connected converter domain. In the second phase of the conformational transition, further seesaw-like motion is hindered because the relay-helix is not a freely pivoting beam. The broken of hydrogen bonds (486/490 and 483/487) accelerate the flip of the 486/487 peptide group and reorients Phe487 from a solvent-exposed to a buried position by forming an "aromatic switch" with Phe503 and Phe506. With these two prerequisites, the converter domain rotates to its final position by virtue of the started unwinding phase of the relay-helix (see Figure 17). This seesaw model suggests a coupling mechanism with two distinct phases in the recovery step. The seesaw-like motion unbend the relay-helix in the first phase and followed by the unwinding phase to induce the myosin lever further swing [39].
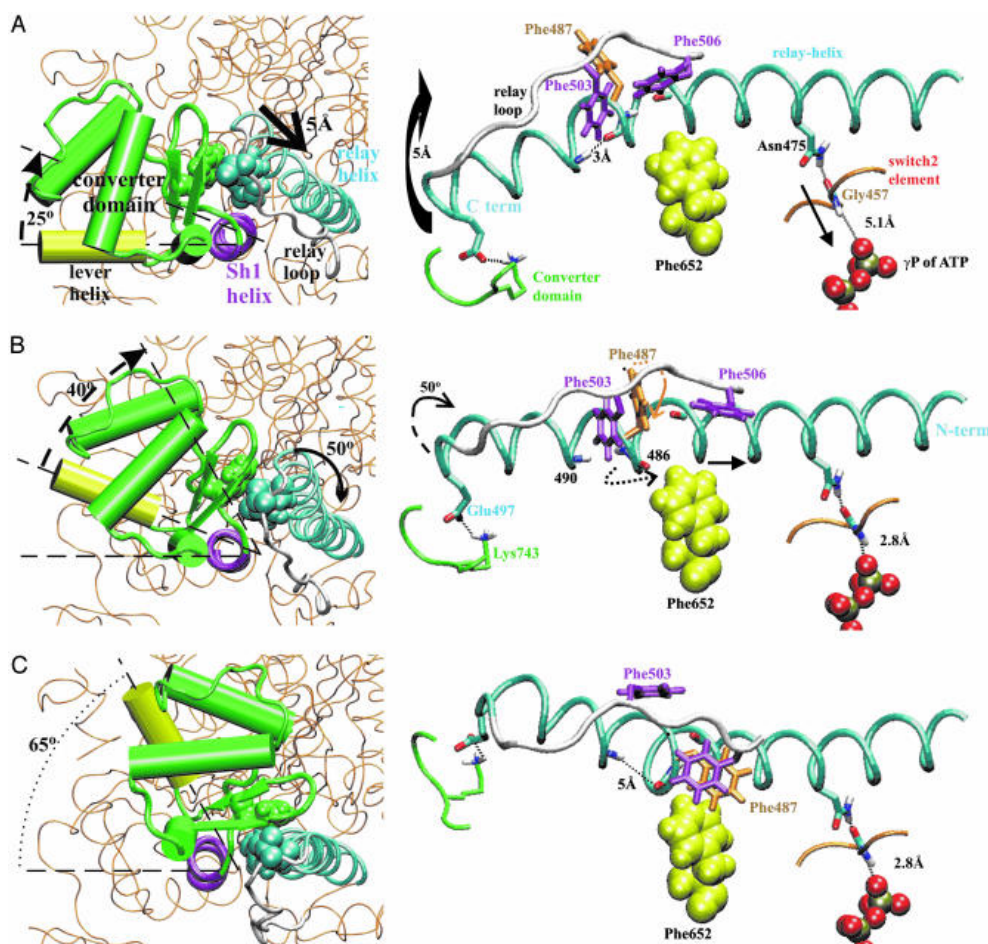
Figure 17: The schematic diagram of the coupling mechanism in the recovery step. The left three panels show the converter domain rotation. The right three panels show some changes along the relay-helix coupled to the small motion in the SW-II loop. (A) Gly457 moves toward the γ-phosphate and pulls the relay-helix unbending through the bond formed with Asn475. (B) The broken of the bond 486/490 favors the reorientation of Phe487, which unwinds the relay-helix and forms a hydrophobic interaction with Phe503, Phe506. (C) The converter domain rotates to its final position with an assistance of the unwinding phase of the relay-helix. The original figure was published by Fischer S *et al.*, 2005. Figure 2 in the original paper [39].

The steady-state and transient kinetic measurements suggest by Málnási-Csizmadia that the phenylalanine cluster mutants contribute a dramatic effect in the formation of the up-lever state with a reduced rate constant of the recovery step [137]. Besides that the phenylalanine cluster cause site-specific dynamic changes of the relay-helix unwinding region in the down-lever state without significant conformational changes based on the *in silico* simulations [137]. These results suggest that this phenylalanine fulcrum can constrain the relay-helix to unwind in the second phase of the recovery step, which is different from Fischer's model that the seesaw-like motion is hindered

in the second phase because the relay-helix is not a freely pivoting beam [137]. Because this seesaw model with two phases separated by one intermediate state neglects the dynamic behaviors of the protein that allows it to populate the neighboring conformational spaces along the suggested trajectory. Therefore, the question is how can we combine all existed intermediate states with suggested structural trajectory to unveil the complete mechanism of the down-to-up lever swing in the recovery step?

Results from our *in silico* simulations proved that the seesaw-unwinding model can be a basic mechanism for the recovery step. Strain along the relay-helix of myosin is rearranged by eliminating the pivoting point in the seesaw-like motions at the beginning stage of the recovery step. A three-phase model of the recovery step deduced on the basis of a free energy profile with the post-recovery state in a lower free energy is more preferred than the pre-recovery state. In the three-phase model of the recovery step, the formation of the hydrogen bond cluster (γ-phosphate~Gly457~Asn475) accelerates structure transformation to overcome the activation energy barrier in the first phase. Our intermediate state 1 (IM-1) has similar conformation to the unique intermediate state of Fischer's model, but the IM-1 state with 37% converter domain rotation is in a middle structural state of the recovery step was not demonstrated before. Phe487/Phe503 hydrophobic cluster closely packing with some indispensable structural rearrangements in our second phase of the recovery step was not regarded as separated phases in Fischer's model since there was no free energy profile of the recovery step in reference.

4.2 Results

4.2.1 MD simulations for pre-recovery and post-recovery states

In order to investigate the functional mechanism of the pivoting point in the relay-helix, two end structures ($1MMD_{wild}$ and $1VOM_{wild}$) and four mutants ($1MMD_{F481A/F482A}$, $1MMD_{F652A}$, $1VOM_{F481A/F482A}$ and $1VOM_{F652A}$) were analyzed our MD simulations. The potential energy profiles along the 2 ns long MD simulations demonstrate that the pre-recovery ($1MMD_{wild}$) and post-recovery ($1VOM_{wild}$) structures achieved stable conformations after the first 500 ps (see Figure 18). Similar potential energy profiles also can be obtained in four different mutants.
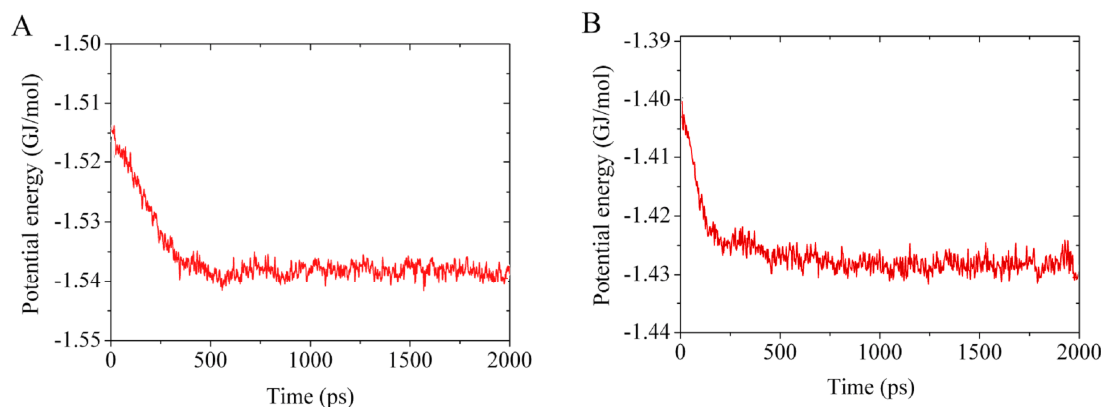
Figure 18: Molecular dynamic simulations for the pre-recovery and post-recovery states. Time courses correspond to the potential energy of two wild types during the MD simulations. Both the pre-recovery state (A) and the post-recovery state (B) relaxed to more stable conformations.

Torsional angles mobility of structures was calculated by averaging the coordinates of 125 structures from the final 250 ps long equilibrium phases. We analyzed the relaxed conformations from six structures (wild types and mutants) in both pre-recovery state and post-recovery states. Two wild types and six mutants in the pre-recovery structure do not show significant differences in their relay-helix regions (Figure 19A). However, the conformation of mutants in the post-recovery state shows deformation in the kink region of the relay-helix (Figure 19B). These results suggest that the phenylalanine cluster serves as a pivoting point in the middle of the relay-helix did

not implement its seesaw-like functions at the early stage of the recovery step (up-lever state), but it extremely destroyed the conformation of the relay-region at the end stage of the recovery step (down-lever state). Breaking the pivoting point hinders the movement of the relay-helix with coupled converter domain rotation. It can also constrain the relay-helix to unwind in the second phase of the recovery step. The seesaw-like motion started from a certain stage after the beginning of the recovery step until the very end of the recovery step from our results is different from Fisher's model that function of this phenylalanine cluster was implemented only in the first phase of the recovery step.
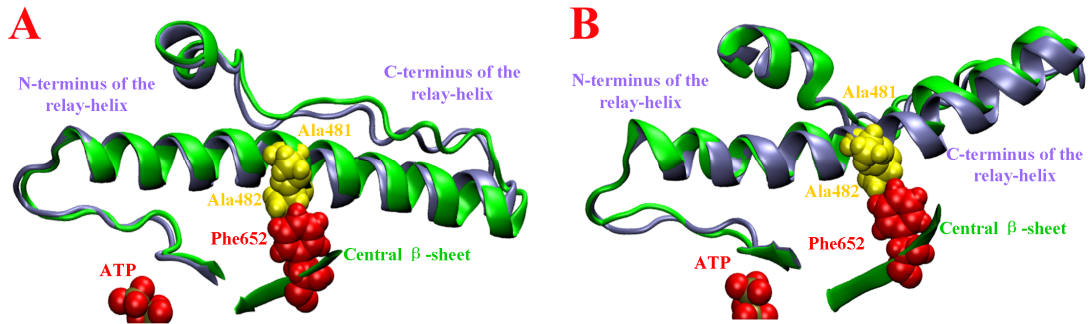
Figure 19: The conformation of myosin relay-helix in the pre-recovery and post-recovery states. A: Pre-recovery conformations of the wild type (Ice blue) and mutant (green) (F481A/F482A). B: Post-recovery conformations of the wild type (Ice blue) and mutant (green) (F481A/F482A).

As Fischer demonstrates that the seesaw-like motion is one of the main events occurred at the beginning of the recovery step, elimination of the pivoting point composed by the phenylalanine cluster (F481A/F482A) will destroy the conformation of the relay-helix region in the pre-recovery state. In order to probe the principle of no significant differences showed in the relay-helix region between the wild type and mutants in the pre-recovery state. We analyzed the dynamic property of the pre-recovery conformation by eliminating the pivoting point. We collected the backbone torsion angles Φ and Ψ of all residues according to equation (7). The averaged amplitudes of the torsion angle dynamics (δ) of the relay-helix in the pre-recovery state for the wild type and mutants are shown in Figure 20. The mobility of residues in the wild type and mutants is very similar to each other in the relay-helix

region except for Ψ491 and Φ492, which are twice as mobile as the others in this region. The mobility of the Φ and Ψ angles of the double mutants (F481A/F482A) at the pivoting point is not changed compared to the wild type. However, the averaged amplitudes of the mobility increase more than 50% at the position of (Ψ486/Φ487) and (Ψ488/Φ489) in the double mutants (F481A/F482A). Meanwhile, the averaged amplitudes of the Ψ491/Φ492 in the double mutants (F481A/F482A) drop to 50% compared to the wild type and a similar effect was detected in the single mutant (F652A). These results suggest that role of the fulcrum is more than just being a pivoting point in the seesaw mechanism, but it also can rearrange the strains along the relay-helix at the beginning stage of the recovery step.
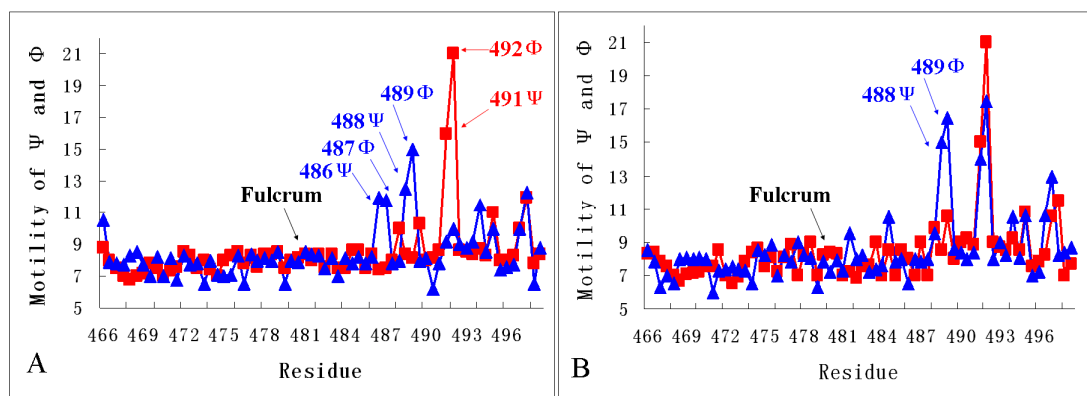
Figure 20: Averaged amplitudes of the torsion angle changes in the relay-helix of the pre-recovery state. The wild type is represented by (■). The triangles (▲) represent double mutants (F481A/F482A) in panel A, and single mutant (F652A) in panel B.

4.2.2 Exploring transition pathway with free energy profile in the recovery step

In order to investigate the mechanism of seesaw-like movement of the recovery step, $1MMD_{wild}$ and $1VOM_{wild}$ represented the pre-recovery state and post-recovery states respectively, were used to simulate the transition pathway by using the forced molecular dynamics (NEB and umbrella sampling). We explored a possible transition pathway of the recovery step with related intermediate states, which combined with related free energy profile were used to unveil the potential mechanism of the seesaw-like motion. In our results, the free energy of the pre-recovery state (-16.3

kcal·mol^{-1}) is 5.02 kcal·mol^{-1} higher than that of the post-recovery state (-21.32 kcal·mol^{-1}) and the activation free energy of the recovery step is 10.15 kcal·mol^{-1}. It means that the pre-recovery state is not as preferred as the post-recovery state. Two obvious free energy wells were discovered and named as intermediate state1 (IM-1: -17.4 kcal·mol^{-1}) and intermediate state2 (IM-2: -15.84 kcal·mol^{-1}). These two energy wells point out that there might be at least two relatively stable intermediate sates during the recovery step (see Figure 21). The corresponding phases can be dug out along the transition pathway, referring to the free energy distribution. According to these intermediate states and related free energy profile, three energy-based transition phases were determined along the pathway, which is a great assistance for us to analyze the conformational changes of myosin functional regions and to deduce the possible mechanism of the recovery step.
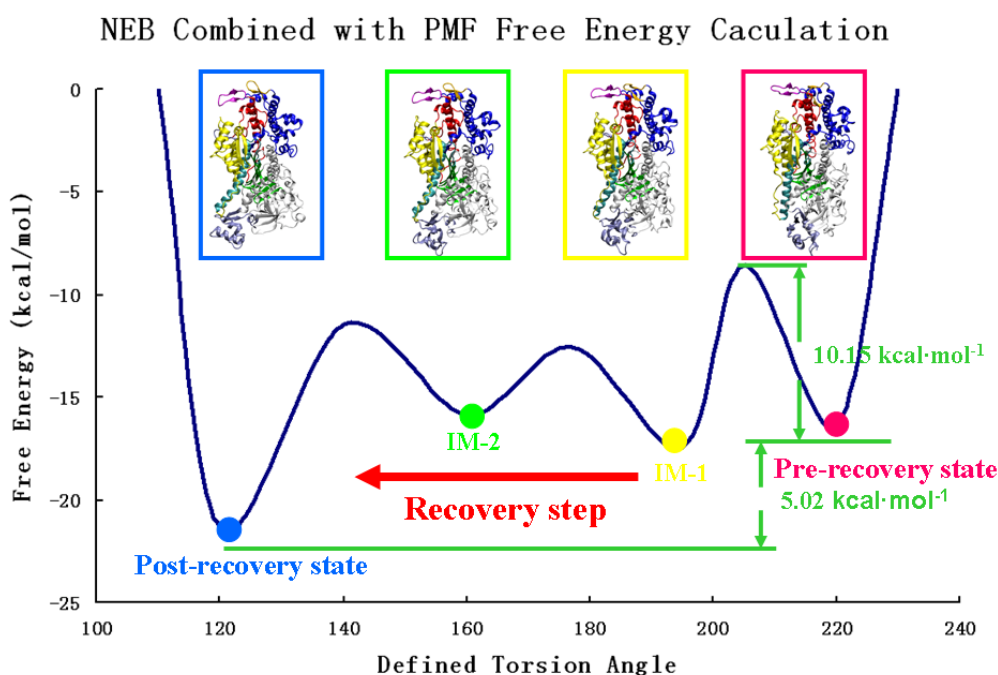

Figure 21: Free energy profile combined with structural changes in the recovery step. X-axis is a defined torsion angle represented by the alpha carbon atoms of residue Gly684, Arg689 and Ala748 over the recovery step. This angle represents the lever arm moving from the "down" to the "up" postion. Two intermediate states (called IM-1 and IM-2) with different free energies were colored by yellow and green points on the free enegy curve. Two end states of the recovery step with 5.02 kcal·mol^{-1} (pre-recovery state and post-recovery state) were colored by pink and blue points on the free enegy curve. Four box-colored myosin structures represent two end states and two intermediate states to their correspongding color points on the curve.

4.2.3 The dynamic similarity assessment of structures in the recovery step

In order to investigate the conformational changes of myosin along the transition pathway, the motional similarities of the two intermediates (IM-1 and IM-2) and the two end states were analyzed with the essential motional analysis on the basis of the root-mean square inner product (RMSIP). The motional similarity of two trajectories can be expressed by the RMISP value of 0~1, which presents a low (0) and high (1) motional similarity, respectively. Our results, with all the values ranging from 0.6 to 0.9 are acceptable for further evaluation since even in the case of two identical macromolecules, it is difficult to achieve with the similarity value of 1, if the trajectories collected with the simulation time less than 10 ns. The similarity of essential motions between pre-recovery and post-recovery states resulted in a value of 0.59 suggests that the same structure adopted different conformations (see Table 7). Transition from the pre-recovery state to the IM-1 state only accounted for ~25% in the rotation angle of the converter domain of the entire recovery step (see Figure 21). The essential motion similarity between the IM-1 state and the post-recovery state is 0.72 similar to the value of 0.69 calculated between the IM-1 state and the pre-recovery state. Result implies that the conformation of the IM-1 state is almost at the middle point of the recovery step in structure. By comparing with only ~37% converter domain rotation from the pre-recovery state to the IM-1 state, ~50% conformational changes in myosin functional regions were occurred in the first phase. Because the rotation of the converter domain contributes most of the conformational changes in the recovery step, ~50% conformational changes are not only from the rotation of the converter domain, but also they are from other myosin functional regions. Transition from the IM-1 state to the IM-2 state accounted for ~36% in the rotation angle of the converter domain of the entire recovery step. The essential motion similarity between the IM-2 state and the IM-1 state is 0.79 as similar as the value of 0.81 calculated between the IM-2 state and the post-recovery state. Result implies that the conformation of the IM-2 state is almost at the mid point of transition pathway from the IM-1 state to the post-recovery state, which indicates ~25%

conformational changes in myosin functional regions were observed in the second phase. In the third phase, we found ~27% converter domain rotation and ~25% conformational changes in myosin functional regions. The conformational transition from the IM-1 state to the post-recovery state suggests that the myosin conformational changes are mainly devoted by the converter domain rotation.

Our results showed that a three-phase model deduced on the basis of a related energy profile and two intermediates is different from the previously demonstrated seesaw-model by Fischer with only two phases [39]. By comparing with two different models and according to results from the essential motional analysis, we can speculate that our first phase with large conformational changes including converter domain rotation and structural rearrangement of other functional region, can be regarded as alike as the seesaw-like motion phase of Fischer's model. Conformational changes occurred in the second and third phases of our model are mainly dedicated from the converter domain rotation, which is almost equal to the unwinding phase of Fischer's model [39]. The separation of our second phase from Fischer's model is due to the related free energy profile and two intermediate states of the recovery step.

Table 7: RMSIP comparison in different states of the recovery step

	Pre-recovery	IM-1	IM-2	Post-recovery
Pre-recovery	0.88	0.69	0.65	0.6
IM-1		0.85	0.79	0.72
IM-2			0.86	0.81
Post-recovery				0.9

The number is the RMSIP value between different states of the recovery step, which was calculated on the basis of the MD trajectories and collected from the top ten eigenvectors.

4.2.4 Discussion

The phenylalanine fulcrum of the relay-helix serves as the pivoting point of the seesaw [39]. In our studies, two different mutations (F481A/F482A and F652A) were introduced to investigate the functional role of this pivoting point. Results verified the

essential role of this phenylalanine cluster in the recovery step by the two mutants with astonishingly identical effects. Lack of this phenylalanine fulcrum diverts the relay-helix into a false conformation in the post-recovery state, but no significant differences were observed between the structures of the wild type and the mutants in the pre-recovery state. It suggests that this fulcrum-based seesaw-like motion is more effective in the middle or at the end of the recovery step, but it is less significant in the beginning stage of the recovery step. We investigated the dynamic property of the residues in the relay-helix to check what kinds of influences occurred by the elimination of the pivoting point. The mobility of the backbone residues along the relay-helix is relatively low in the pre-recovery state, except for a highly mobile part ($\Psi491/\Phi492$) in the kink region. The flexibility of the pivoting point in these two mutants is not perturbed. However, highly mobile parts were shifted to $\Psi486/\Phi487$ and $\Psi488/\Phi489$ by a few amino acids away from the pivoting point in the recovery step. This means that constrains along the relay-helix are rearranged by the elimination of the pivoting point in the seesaw-like motions. These results suggest that the role of the fulcrum is more than just a pivoting point in the seesaw mechanism. It can also constrain the relay-helix to unwind in the later phase of the recovery step.

Experimental results showed that in the motor domain of the wild type the free energy change of the recovery step is around $\Delta G = -9.4 KJ \cdot mol^{-1}$ based on the rate constants of pre-recovery to post-recovery transition is ~1000 s^{-1}($k_{obs}=k_{3a}+k_{-3a}$) with the equilibrium constant close to 45 at a temperature of 296K in the ATP-bound state [138]. The calculated free energy of the pre-recovery state is higher than that of the post-recovery state, and it is also agreed with the experimental results that the post-recovery state is more preferable than the pre-recovery state. However, the calculated energy difference between the pre-recovery state and post-recovery state is -5.02 kcal$\cdot$mol^{-1} with twice as high as the one estimated from experiments. The reason might be the approximated nature of the reaction coordinates while using the PMF calculations. Fortunately, the free energy calculation of the recovery step has the same order of magnitude for both experimental and PMF results. Macromolecular

transition pathway simulations combined with the free energy calculations should be performed with the utmost caution when we use NEB with the umbrella sampling for the conformational discovery and the PMF method for the free energy calculation. Because a sophisticated conformational change cannot be easily described merely with distances, angles, and some sub-groups constrains by only using the umbrella sampling method [114]. With an assistance of the pre-predicted transition pathway by using the NEB method, a minor difference in the neighbor transition states can be easily and trustily simulated by umbrella sampling, which is really necessary to deduce a reasonable mechanism of the recovery step.

On the basis of our essential motional analysis, we deduced a three-phase structural model with the related energy profile which is different from Fischer's model that the seesaw-like motion occurred in the first phase is hindered in the second unwinding phase [39]. The first phase (from pre-recovery to IM-1) in our simulation is similar to the seesaw phase, which is a relay-helix coupled unbending transition. Our mutants results suggest that the fulcrum-based seesaw-like motion is more effective in the middle or at the end of the recovery step, but it is less significant in the beginning stage of the recovery step because the hydrogen bond cluster (γ-phosphate~Gly457~Asn475) is not formed in the pre-recovery state (from A to A' in Figure 22). The start of the seesaw-like motion is particularly dependent on the formation of the hydrogen bond cluster, which marked with transition A' state with the highest free energy in the recovery step. From state A to state A', myosin undergoes 18% conformational changes referred to the RMSD calculation between the state A' and the pre-recovery step and coupled with 14% converter domain rotation. This hydrogen bond cluster accelerates energy transmission to overcome the activation energy (from A to A' in Figure 22) and to achieve the purpose of later myosin conformational changes. After overpass this activation energy, the converter domain undergoes 23% seesaw-like motion coupled with myosin 32% conformational changes according to the RMSD calculation in the first phase (from the pre-recovery state to the IM-1 state or from A state to B state in Figure 22), which lead to the IM-1 state in a middle transition structure of the recovery step based on RMSIP calculation

of the essential motional analysis. Although by comparison the intermediate state of Fischer's model to our state B (IM-1) structurally, we find that they have quite similar conformation in both nucleotide binding pocket and the converter domain, but this intermediate state as the mid state with only 37% converter domain rotation was not demonstrated in Fischer's model.

The second phase from IM-1 to IM-2 (from B to C in Figure 22) is unveiled in our simulation based on the unearthed lower energy well, which was only elaborated in the Fischer's model with limited conformational changes [39]. Fischer suggests that some indispensable side-chain rearrangements occurred in this phase is resulted from substantial conformational changes sterically hindered as long as the relay-helix seesaw-like motion is not completed. We found also some structural rearrangements and a special hydrophobic interaction formed between Phe487 and Phe506 in this phase, which results in a close compact between the relay-loop and the relay-helix. In this phase, 31% myosin conformational changes are observed by referring to the RMSD calculation combined with 36% converter domain further-up rotation. We speculate that myosin conformational changes occurred in this phase are mainly resulted from the contribution of newly formed Phe487-Phe506 hydrophobic cluster impelled the converter domain further-up rotation.

The last phase is from IM-2 to post-recovery step (from C to D in Figure 22) in our model. Interaction of the Phe487/Phe503-based hydrophobic cluster is gradually in enhanced and shifted from a solvent-exposed to myosin buried position, which causes myosin undergoing 19% conformational changes according to the RMSD calculation and the C-terminus of the relay-helix further-up rotation provoking 27% converter domain unwinding via the SH1 helix movements. However, Fischer did not regard Phe487/Phe503 hydrophobic cluster shifting in our third phase and it closely packing with assistance of some indispensable structural rearrangements in our second phase as separated phases since they do not have any results consulting to the related free energy profile of the recovery step.
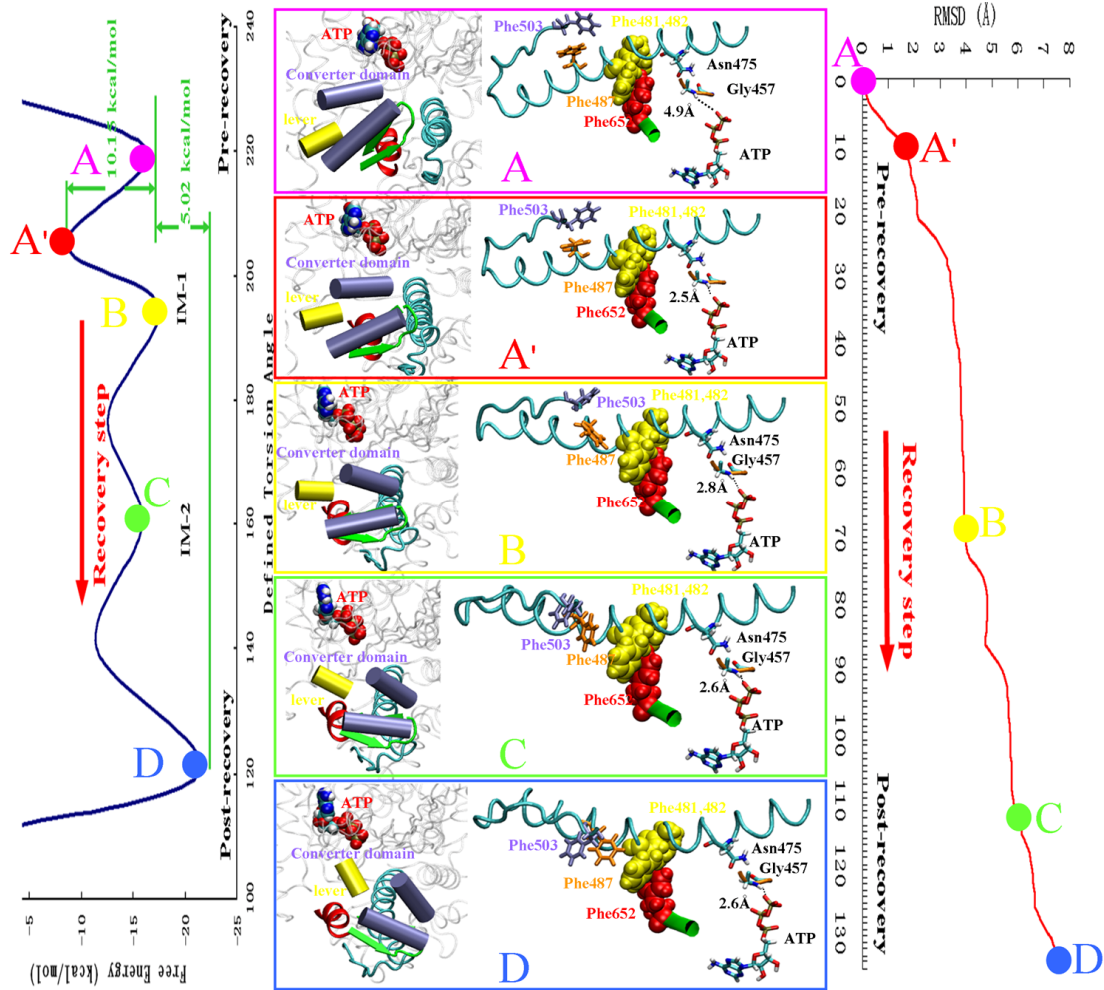
Figure 22: A three phase coupling mechanism is combined with free energy in the recovery step. The left panel shows the free energy profile with two intermediate states and one transition state. Five color points represent pre-recovery state (A), transition state (A'), intermediate state-1 (IM-1: B), intermediate state-2 (IM-2: C) and post-recovery state (D). Similar color points represent same states based on the RMSD calculation in the right panel. In the middle panel, conformation of the relay-helix and the converter domain of myosin boxed in different colors corresponding to five various states in both left and right panel. (From A to A'): Gly457 moved toward the γ-phosphate forms a hydrogen bond cluster (γ-phosphate~Gly457~Asn475). (From A' to B): This hydrogen bond cluster pulls the relay-helix unbending based on the seesaw-like motion. (From B to C): Phe503 rotation formed a hydrophobic interaction with Phe487 results in a close compact between the relay-loop and the relay-helix to impel the converter domain further-up rotation. (From C to D): Interaction of the Phe487/Phe503-based hydrophobic cluster is gradually in enhanced and shifted from a solvent-exposed to myosin buried position, which causes the converter domain unwinding to its final position via the SH1 helix movements.

CHAPTER 5: The Communication mechanism upon the power stroke discovered from a novel functional region of myosin

5.1 Introduction

Myosin bound to the actin filament results in a lever swing, which is called power stroke [24]. The cyclic actin binding and release, alternating with the swing of the myosin lever, define the mechanical cycle of actomyosin. This mechanical cycle is coupled with a chemical cycle in nucleotide binding, hydrolysis and product release. This chemomechanical cycle of actomyosin described different myosin structures in PDB code are assigned to different states of the Lymn-Taylor cycle (see Table 8). After ATP hydrolysis, myosin in the up-lever state starts to bind to actin in a weak actin-binding state (actin-myosin·ADP·Pi) that are described in the state "VI" of the Lymn-Taylor cycle in Table 8 [139]. The actin binding induces the up-to-down lever swing of myosin depicted from state VI to state I of the Lymn-Taylor cycle in Figure 4 that is followed by the release of hydrolysis products [140, 141].

Table 8: Assignment to the different states of myosin in the Lymn-Taylor cycle

State (PDB code)	SW-I	SW-II	Actin-binding cleft	β-Sheet	P-loop	Relay-helix	Converter domain
"I" (1Q5G,2OVK) (Rigor, Apo-state)	Open	C1	Close	Twist	Up	Straight	Down
"II" (1MMD) (Pre-recovery) (ATP bound)	Closed	Open	Open	No- twist	Down	Straight	Down
"III" (1VOM) (Post-recovery) (ATP bound)	Closed	C2	Open	No- twist	Down	Kink	Up
"IV" (1VOM-m*) (Bound to actin) (ADP·Pi)	*Closed*	C2	*Closed*	No- twist	Down	Kink	Up

The values of the seven movable elements that have been identified for the states defined in the Lymn-Taylor cycle. C1 and C2 are two different forms of closed conformation found in the rigor and the post-recovery states. The assumed properties of conformation (shown in italics) are extrapolated from the adjoining states in the

Lymn-Taylor cycle. 1VOM-m*: 1VOM with its ATP was replaced by ADP·Pi.

After ATP hydrolysis, if the up-to-down lever swing occurs in an actin-detached state of myosin (called reverse recovery step (see Figure 23 pathway 1st)), the chemical-energy expenditure is not accompanied by the production of useful mechanical work. This lever swing in actin detached state is the kinetically rate-limiting step of the ATPase cycle if myosin does not attach to actin during the cycle [63]. We note that the up-to-down lever swing with ADP·Pi-bound of myosin in the reverse recovery step is four orders of magnitude slower than the down-to-up lever swing with ATP-bound of myosin in the recovery step. Furthermore, the reverse recovery step is a reversible step in the absence of actin while power stroke occurs in the actin-attached form and becomes irreversible. The ratio of the effective power stroke over the reverse recovery step is accelerated up to two orders of magnitude by the actin binding in the ADP·Pi state of myosin. Consequently, the predominant reaction flux is diverted to an actin-attached process, even though it is a thermodynamically less favorable pathway because of the low actin-binding affinity of the actomyosin complex. In addition, there is another possible pathway if the actin-binding cleft closes before myosin binds to actin (pathway 2nd in figure 23). This is a very low possibility pathway since the open-closed transition of the actin-binding cleft is thermodynamically unfavorable ($K < 0.01$) in actin-detached state [56].

The effective power stroke occurs only when the myosin head is bound to actin. An efficient power stroke is generated through various pathways, which can be determined by distinct energy profiles and mechanical performances of the system. Whatever, three coupled events are viewed as 1. Myosin head binding to actin, 2. Structural changes resulted from weak to strong actin binding, i.e. actin-binding cleft open-closed transition and 3. The lever swing induced force generation. The power stroke starts in a weak actin-binding state (actin·myosin·ADP·Pi) of myosin in an open cleft and up lever position. The pathway represented by pathway 3rd in figure 23 is the so called the weak-to-strong actin-binding transition with the actin-binding cleft

closure occurring first and then followed by the lever swing [140]. Alternatively, the lever swing might occur first while the actin-binding cleft is still in the open state and the myosin head binds weakly to actin (see Figure 23 pathway 4th). The subsequent strengthening of the actomyosin interaction serves as a stabilizing role for the post-power stroke state that promotes the cleft closure more rapidly than that of myosin detaching from actin [143]. Both the pathway 3rd and pathway 4th could convey significant fluxes based on the available kinetic and thermodynamic data, which can push the flux of the effective power stroke through parallel reaction pathways [64]. The ratio of the productive power stroke (pathway 3rd and 4th) over the reverse recovery step (pathway 1st and 2nd) is regulated by allosteric activation of actin (known as actin activation) [64]. Actin activation is not necessary for exerting the functions of myosin motor domain, but it determines the energy efficiency of the mechanochemical system. However, the molecular mechanism and structural background of the actin-induced conformational changes in myosin also remain unrevealed.
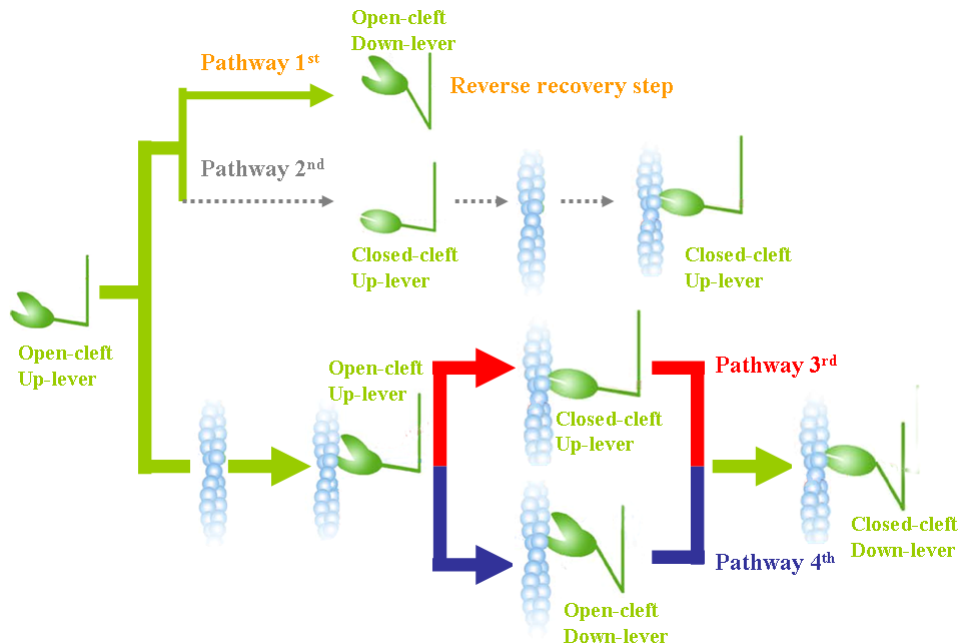
Figure 23: Four distinct pathways in the myosin power stroke. The initial state of the pathway is myosin in open cleft and up lever state which is detached from actin. If the lever swing occurs before the actin binding, a reverse recovery step is generated (pathway 1st). In the absence of actin, the actin-binding cleft closure starts before the actin binding, which is an unfavorable process (pathway 2nd). Pathway 3rd and 4th are parallel to each other. Both processes start in the actin bound myosin open cleft and up lever state and end by the actin bound myosin closed cleft and down lever state.

Atomic structure of actomyosin is still not available. Atomic structure of actomyosin is still not available. Only the high resolution electron micrographic model of rigor actomyosin complex has been determined in 14 Å resolutions [35, 43] and *in silico* docking simulations [45, 49, and 144]. The strength of actin-binding affinity is regulated by the closure of the actin-binding cleft that is formed by the upper and lower 50 kDa domains of myosin [143] (see Figure 24). The CM-loop (cardiomyopathy loop) and loop 4 from the upper 50 kDa domain of myosin are mainly bound to the C-terminal domain of actin (see Table 9). Mutations in these myosin regions abolish the interactions between the two loops and actin in the strong actin-binding state (closed cleft of myosin), while without significantly affects to the weak actin-binding state (open cleft of myosin) [144, 145]. In the lower 50 kDa domain of myosin, two loops (helix-loop-helix and loop 2) as the most predominant regions are also involved in actin-binding interface. Mutations in these loops lead to adverse impacts in both weak and strong actin-binding state and in actin activation [52, 53, 146 and 147]. On the other hand, actin also contains various myosin binding regions. The major myosin-binding area of actin is residing in subdomain I that include parts of the C- and N-terminus of actin. If the four conserved and negatively charged residues at the beginning of the N-terminal region of actin are deleted or mutated to neutral amino acids, the actin binding is not able to increase the basal ATPase activity of myosin [148, 149]. So far, these four residues are usually missing in the existing models of actomyosin structures preventing the studies to reveal the mechanism of actin activation related to the N-terminal segment of actin [44]. Residues principally involved in the binding interface between actin and *Dictyostelium* myosin II are listed in Table 9.

Our aim was to reveal whether the up-to-down lever swing in the power stroke occurs similarly to that of in the recovery step as it was supposed by Fischer and Holmes [39]. Furthermore, what conformational changes are induced by the actin binding in the different states of myosin? Based on the recent experimental technologies, the atomic resolution of the weak actomyosin complex cannot be approached, because of its short lifetime in the enzymatic cycle. Thus, we modeled its structural and dynamic features
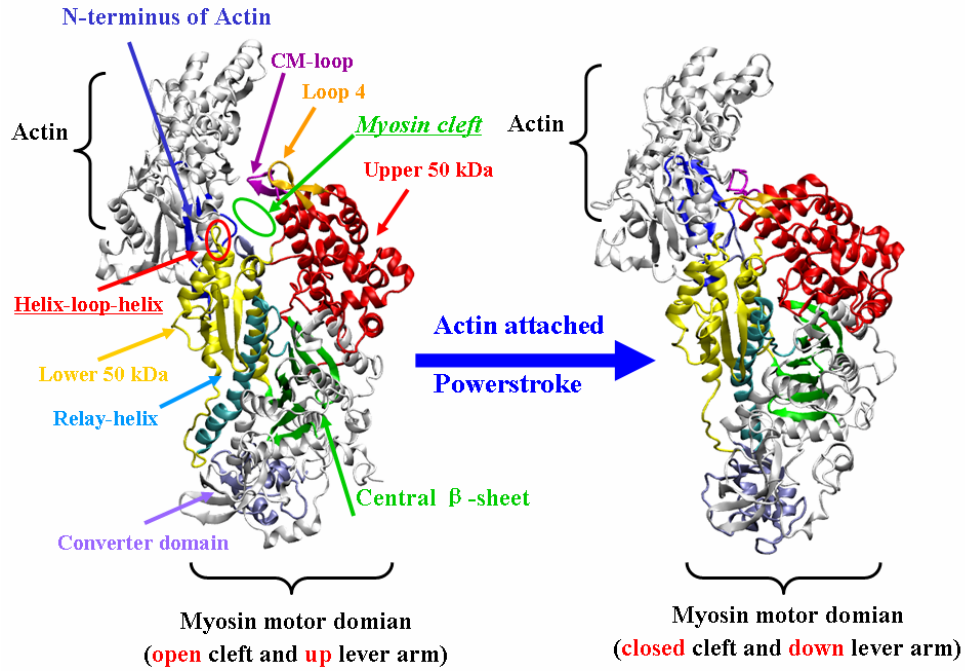
by *in silico* simulations.

Figure 24: Structural model of weak and strong actomyosin complexes. The left picture is actin-attached myosin open cleft and up-lever arm state (weak actin-binding state or the prepower-stroke actomyosin state). The right picture is actin-attached myosin closed cleft and down-lever arm state (strong actin-binding state or the postpower-stroke actomyosin state). The actin-binding cleft is highlighted by a green circle in the left picture. These color-codes for the functional subdomains of myosin are represented by drawing method of cartoon (Green: central β-sheet; Orange: Loop 4; Purple: CM-loop; Dark ice blue: loop 2, Cyan: relay-helix; Yellow: lower 50 kDa domain; Red: upper 50 kDa domain; Light ice blue: converter domain; and Blue: N-terminus of actin).

Table 9: Areas of contact between actin and *Dictyostelium* myosin II

Contact	Myosin	Actin	Comments
Loop 4	360-370	145-149, 328-335	Contact between the exposed loop 4 in the upper 50 kDa domain and actin.
CM-loop	390-410	322-333	Contact between the exposed CM-loop in the upper 50 kDa domain and actin.
Activation loop	519-523	1-4	Salt-bridge between Arg520 and four negtively charged residues at N-terminus of actin.
Lower 50 kDa domain	519-555	144-148, 339-354	Hydrophobic interaction between Lower 50 kDa domain and actin.
Loop 3	560-570	79-92, 91-100	Electrostatic interaction between loop 3 and the second actin monomer.
Loop 2	620-630	1-4, 22-27	Contact between two lysines and actin.

We remodeled the weak and the strong actomyosin complexes by using complete actin trimer structure which also contains the N-terminal residues. The weak and strong actin-binding states were obtained through protein-protein docking and long time-range molecular dynamics relaxations while the binding free energies were calculated from these models. The binding affinity of the actomyosin complex is represented by its binding free energy. The lower binding free energy of the strong actin-binding state is due to the more residues involved in the binding surface than that of in the weak actin-binding state. The binding interface analysis of different actomyosin complexes implies structural orientations and the interaction patterns of loop 2 probably coupled with actin-binding cleft closure. The recently discovered activation loop appears distinct characteristics in the weak and strong actin-binding state. The conserved positive tip of the activation loop (Arg520) interacts with four negatively charged residues in the N-terminus of actin in various binding patterns of weak and strong actin-binding state [55]. Three specific myosin conformational changes induced by actin binding were observed in the weak actin-binding state: 1. The partial closure of the actin-binding cleft, 2. The further up rotation of the lever arm and 3. Further closure of SW-II loop. The R520Q mutation in myosin prevented these conformational changes. Motional correlations suggested that motions of the activation loop are correlated with four mainly functional regions of myosin (loop 2, the N-terminus of the relay-helix, the SW-II loop and the "prestrut" region). When the R520Q mutation was introduced to myosin, the motional correlations between the activation loop and other four regions decreased to the same level as that of in the absence of actin. Two communicational pathways were speculated between the actin binding regions and the myosin nucleotide binding site, which might be a reasonable mechanism of the actin-induced myosin conformational changes at the initial stage of the power stroke.

5.2 Results

5.2.1 MD simulations of different actomyosin states

The missing residues of the applied four myosin structures (1VOM·Mg·ADP·Pi, 1MMD·Mg·ADP·Pi, 1Q5G and 2OVK) were replaced by homology modeling. Then they were minimized and relaxed by molecular dynamics. Resulted structures were named as $1VOM_{md}$, $1MMD_{md}$, $1Q5G_{md}$ and $2OVK_{md}$, respectively. In the course of total 100 ns long MD simulations the four structures converged within the first 3.5 ns (see Figure 25). Structural fluctuations were equilibrated in the RMSD value of 2.33±0.16Å, 2.27±0.17Å, 2.14±0.15Å and 2.68±0.16Å for $1VOM_{md}$, $1MMD_{md}$, $1Q5G_{md}$ and $2OVK_{md}$ respectively. All structures were relaxed to their equilibrium states up to 5.5 ns in MD simulations. Compare with crystal structures no significant backbone rearrangement occurred during the molecular dynamics simulation.
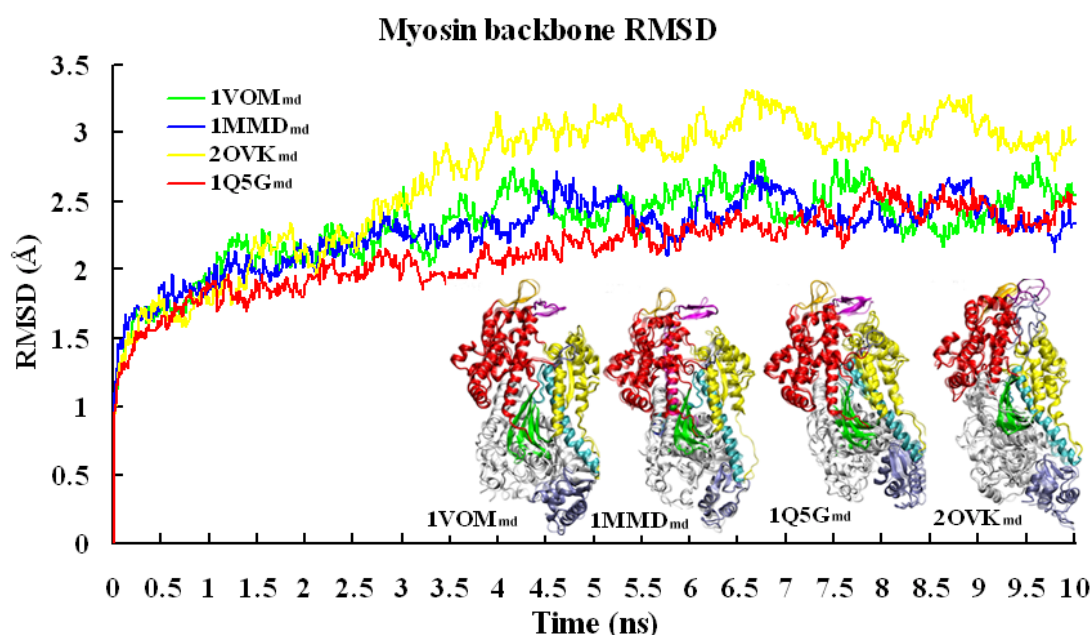
Figure 25: Four myosin structures before and after MD simulations. Different regions of myosin structures are color-coded (Red: upper 50 kDa domain; Yellow: lower 50 kDa domain; Green: β-sheet; Orange: loop 4; Purple: CM-loop; Ice blue: converter domain; Cyan: relay-helix; White: others). RMS deviations were drawn as a function to the first 10 ns MD simulations.

5.2.2 Protein-protein docking and validations in different actomyosin complexes

The relaxed myosin structures were docked to actin trimer as described in the Materials and methods section. In order to identify the structural background of actin induced myosin conformational changes, the specific actin-myosin interactions were characterized. First we analyzed these interfaces obtained from protein-protein docking procedures. Major docking results of 3A-1VOM$_{dock}$, 3A-1MMD$_{dock}$, 3A-1Q5G$_{dock}$ and 3A-2OVK$_{dock}$ are summarized in Table 10. The ensemble of populated structures in each of actomyosin complexes shows an approximate 5~6Å RMS deviations based on their backbones. The intermolecular energies E$_{VWD}$ (van der Waals), E$_{ELEC}$ (electrostatics) and E$_{DESO}$ (desolvation) are the main contributions to the HADDOCK$_{SCORE}$. The ambiguous interaction restraint energy (E$_{AIR}$) with similar values around 40 kcal·mol^{-1} in four actomyosin complexes is only a minor impact on the total energy. Averaged values of the BSA (buried surface areas) in the four actomyosin complex are between 1624 Å^2 and 1836 Å^2. The HADDOCK$_{SCORE}$ between 3A-1VOM$_{dock}$ and 3A-1MMD$_{dock}$ is principally different in the value of E$_{VWD}$, but not the value of E$_{ELEC}$, E$_{DESO}$, E$_{AIR}$ and BSA, which indicate that there is no different between these two actomyosin complexes in their binding patterns. 3A-1Q5G$_{dock}$ shows comparable values in the E$_{VWD}$, E$_{DESO}$ and E$_{AIR}$, but a higher value in the E$_{ELEC}$ (-262.4 ± 41.7 kcal·mol^{-1}) and a larger BSA (1757.7 ± 106.2 Å^2) to those above two complexes. The HADDOCK$_{SCORE}$ of 3A-1Q5G$_{dock}$ (0.92 ±0.1) is higher than that of 3A-1VOM$_{dock}$ (11.4 ±2.3) and 3A-1MMD$_{dock}$ (8.0 ±2.1), which suggests distinct binding patterns are existed between 3A-1Q5G$_{dock}$ and other two actomyosin complexes. 2OVK in the docked complex 3A-2OVK$_{dock}$ comes from the squid with highly structure homologous to the *Dictyostelium* myosin 1Q5G, which lead to a similar binding pattern between 3A-2OVK$_{dock}$ and 3A-1Q5G$_{dock}$.

Table 10: HADDOCK results of distinct actomyosins

Name[A]	3A-1VOM$_{dock}$	3A-1MMD$_{dock}$	3A-1Q5G$_{dock}$	3A-2OVK$_{dock}$
HADDOCK$_{SCORE}$[B]	11.4 ±2.3	8.0 ±2.1	0.92 ±0.1	-0.43±0.4
Rmsd (Å)[C]	6.2 ± 1.5	5.2 ± 1.1	4.9 ± 1.2	5.6 ± 1.5

N [D]	18	10	11	17
E_{VDW} [E] (kcal mol^{-1})	-37.1 ± 4.2	-39.8 ± 3.1	-40.3 ± 4.1	-40.9 ± 4.0
E_{ELEC} [E] (kcal mol^{-1})	-242.2 ± 33.3	-239.2 ± 26.8	-262.4 ± 41.7	-275.9 ± 96.0
E_{AIR} [F] (kcal mol^{-1})	45.1 ± 2.9	43.0 ± 3.0	39.2 ± 2.3	41.7 ± 3.1
E_{DESO} [G] (kcal mol^{-1})	92.4 ± 9.4	91.3 ± 11.7	89.7 ± 10.7	91.5 ± 9.1
BSA (Å^2) [H]	1624.3 ± 113.9	1674.8 ± 105.8	1757.7 ± 106.2	1836.4 ± 119.8
RMS Deviation (Å) [I]	7.2 ± 2.7	6.3 ± 2.3	4.7 ± 1.7	5.5 ± 2.0

[A] The actin trimer docked to four different averaged myosins. [B]The HADDOCK$_{SCORE}$ is calculated according to the equation (6). [C] RMS deviations are calculated on the basis of their myosin backbones. [D]Number of collected structures in given ensembles. [E]Intermolecular energies (van der Waals and electrostatics). [F] ambiguous interaction restraint energy. [G]Desolvation energy. [H] Buried surface area. [I] RMS deviations are calculated between the averaged myosin structure of each MD relaxed actomyosin ensembles and myosin in the Holmes's model without their C-terminus been considered.

Four actomyosin complexes were obtained by protein-protein docking, but the exact binding interface was difficult to explore due to the existing spatial resistance between the actin trimer and myosin. Therefore, 20 ns NPT structural relaxations were carried out in MD simulations by referring to Figure 11 and 12. In order to validate the MD relaxed actomyosin complexes, we introduced an overlapping method based on the electron microscopy density map (EM-map) of a modeled strong actin-binding state (postpower-stroke actomyosin state) named as Holmes's model [43]. The grid space consisting of F-actin and myosin was generated from EM-map based on Holmes's rigor actomyosin complex. If we overlap our relaxed actin timer into the grid space of F-actin, then the structural deviations between myosins in our four MD relaxed actomyosin complexes and the grid space of myosin from EM-map are able to be calculated. One thing to note is that results of calculated RMS deviations are merely based on myosin without its converter domain. By the way, in later parts of this thesis, some of our results will often refer or compare to this Holmes's model. The major structural deviations of the simulated structures from the rigor EM map were found in the position of loop 4 and CM-loop in the upper 50 kDa domain of myosin (see Figure 26). In 3A-1VOM$_{md}$, loop 4 and the CM-loop protrude out of the EM-map because of the open actin-binding cleft (see Figure 26, 3A-1VOM$_{md}$, arrow 1 and arrow 2). Parts of the relay region and myosin lever related

converter domain are also out of the EM-map (see Figure 26, 3A-1VOM$_{md}$, arrow 3 and arrow 4). These differences are what we expected due to the open cleft and up-lever conformation in the 3A-1VOM$_{md}$ is different from closed cleft and down-lever conformation in the Holmes's model (rigor state). In 3A-1MMD$_{md}$, the straight relay region and the converter domain are involved in the EM-map, but loop 4 and the CM-loop in upper 50 kDa domain with an open actin-binding cleft are similar to those of 3A-1VOM$_{md}$ (see Figure 26, 3A-1MMD$_{md}$, arrow 1 and arrow 2) and parts of the myosin N-terminus are shifted out of the EM-map (see Figure 26, 3A-1MMD$_{md}$, arrow 3). As we expected in 3A-1Q5G$_{md}$, which represents a rigor actomyosin complex, not only the relay region and the converter domain, but also loop 4 and the CM-loop are completely involved in the EM-map. These results prove that our non-directed protein-protein docking and MD simulations create a similar conformation like that determined via the molecular dynamics flexible fitting (MDFF) in the Holmes's model [43]. 3A-2OVK$_{md}$ also as rigor actomyosin complex appears in a similar conformation to 3A-1Q5G$_{md}$ with most of its structures deeply buried in the EM-map. Since loop 2 is ten residues longer in squid myosin than chicken and *Dictyostelium* myosin II, 3A-2OVK$_{md}$ was not able to completely fit the EM-map in this region (see Figure 26, 3A-2OVK$_{md}$ arrow 1). The extended loop 2 in the squid actomyosin complex (3A-2OVK$_{md}$) is difficult to fold into the EM-map based structural envelop without forced MD simulations.

Quantitative validation can be evaluated on the basis of the backbone RMS deviation of myosin among four MD relaxed actomyosin complexes and Holmes's model. In Table 10, 3A-1VOM$_{md}$ (7.2 ± 2.7 Å) shows a similar RMS deviation to 3A-1MMD$_{md}$ (6.3 ± 2.3 Å). 3A-1Q5G$_{md}$ better fit to the EM-map than 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$ has a lower RMS deviation value of 4.7 ± 1.7 Å similar to the 3A-2OVK$_{md}$ (5.5 ± 2.0). Long and unfolded loop 2 in 3A-2OVK$_{md}$ makes the RMS deviation value of this squid actomyosin complex larger than the *Dictyostelium* actomyosin 3A-1Q5G$_{md}$. As the EM-map was created from the chicken myosin in rigor state (Holmes's model), 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$ with their myosin in the rigor-like state (closed actin-binding cleft) are better overlapped than 3A-1VOM$_{md}$
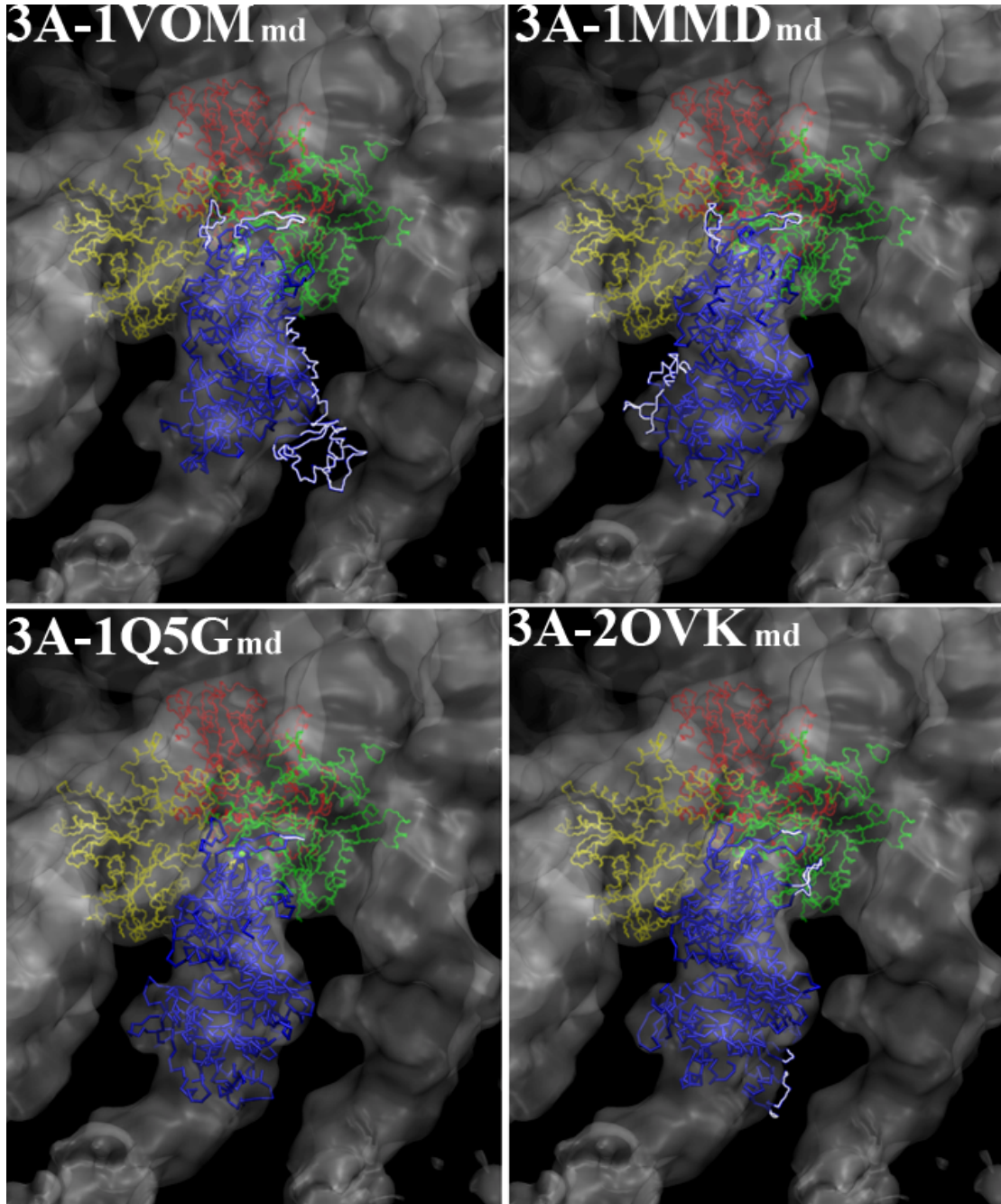
and 3A-1MMD$_{md}$ (open actin-binding cleft).

Figure 26: MD relaxed four docked actomyosin complexes. Actin-trimer is colored by green, red and yellow in each monomers and myosin is colored by blue in four actomyosin complexes. White arrows show differences between four MD relaxed actomyosin complexes and Holmes's model based grid space generated from the EM-map. The blue segments of myosin are entirely involved in the EM-map, but the white parts as are still outside of the EM-map.

To summarize the results in this section, the MD relaxed 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$ as the beginning and transition states of the power stroke are deemed to

be weak actin-binding states, whereas the 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$ with a better fitting to the Holmes's model are considered closely to the end states of the power stroke in a strong actin-binding state.

5.2.3 Evaluation of the weak and strong actin-binding states

In order to evaluate the differences of the binding affinity in the weak and strong actin-binding states, the binding free energy was calculated on the basis of the MM/PBSA method [101]. This method is highly insightful as it provides various components for the binding free energy, including the van der Waals energy and electrostatics interactions, the non-polar components of the solvation energy, and entropy [133].

Results of binding free energy and binding surface analysis were summarized in Table 11 and Table 12. In 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$ (open actin-binding cleft states), the number of residues involved in the binding surface is similar to each other (see Table 12), therefore, the binding free energy changes only influenced by the lever arm up-to-down movements which resulted in less than 1 kcal/mol difference (see Table 11). The binding free energy around -11~-13 kcal·mol^{-1} in both 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$ is much lower than that of 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$, which are due to fact that more than ten extra residues are involved in the rigor states (see Table 12). In addition, 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$ have similar numbers of residues and atoms in the binding surface to those in the Holmes's model. These results suggest that the higher binding affinity comes mainly from the movements of loop 4 and the CM-loop in the upper 50 kDa domain of myosin. This actin-binding cleft closure on behalf of the weak-to-strong actin-binding transition with ~4 kcal·mol^{-1} of free energy decreases is also confirmed with previous kinetic experimental results [150]. The binding free energy of the 3A-2OVK$_{md}$ complex is ~1 kcal/mol higher than that of the 3A-1Q5G$_{md}$ complex, which is probably caused by the different conformations of loop 2 regions between squid and *Dictyostelium* myosin. Table 11 reveals that four actomyosin complexes have non-negligible differences in the van der Waals energy

and electrostatic interactions, which are nearly -20~-30 kcal·mol^{-1} and -10~20 kcal·mol^{-1} of energy differences between the weak actin-binding states (3A-1VOM$_{md}$, 3A-1MMD$_{md}$) and the strong actin-binding states (3A-1Q5G$_{md}$, 3A-2OVK$_{md}$).

These binding free energy results suggest that the increasing actin-binding affinity in the power stroke comes mainly from the upper 50 kDa domain moving towards the lower 50 kDa domain. This was also demonstrated by previous results that the weak-to-strong actin binding is based on the actin-binding cleft closure, and is not related with the myosin lever's up-to-down movement [151].

Table 11: Binding free energies of different actomyosin complexes

Name[A]	ΔE_{vdw}[B] (kcal/mol)	ΔE_{ele}[C] (kcal/mol)	ΔE_{PBSA}[D] (kcal/mol)	ΔE_{SASA}[E] (kcal/mol)	ΔH[F] (kcal/mol)	$T\Delta S$[G] (kcal/mol)	ΔG_{bind}[H] (kcal/mol)
3A-1VOM$_{md}$	-127.3±9.6	-103.4±8.8	230.0.±16.6	-13.5±1.6	-14.1±1.3	-6.5±0.6	-7.6±0.8
3A-1MMD$_{md}$	-136.9 ±9.9	-111.3 ±7.6	248.7±17.4	-15.7±1.5	-15.2±1.3	-6.8±0.9	-8.4±1.1
3A-1Q5G$_{md}$	-152.2±11.7	-128.3±10.0	281.9±19.2	-20.9±1.3	-19.6±1.2	-7.1±0.8	-12.4±1.3
3A-2OVK$_{md}$	-167.3±12.4	-129.2±11.0	300.1±16.3	-22.4±1.6	-18.9±1.9	-7.3±1.1	-11.4±1.0

[A] Four MD relaxed actomyosin complexes. [B] Van der Waals contributions. [C] Electrostatic energy contributions. [D] Polar contributions to the solvation free energy. [E] Non-polar contributions to the solvation free energy are based on the solvent-accessible surface area. [F] Enthalpic contribution: $\Delta H = \Delta E_{vdw} + \Delta E_{ele} + \Delta E_{PBSA} + \Delta E_{SASA}$. [G] Entropic contribution (ΔS) is based on the normal-mode analysis. [H] Binding free energy is calculated on the basis of $\Delta G_{bind} = \Delta H - T\Delta S$.

Table 12: Comparisons of actomyosin complexes before and after MD simulations

Names		Myosin		Actin		Surf-area	HB	SB
		Ires	Inat	Ires	Inat			
3A-1VOM	Docked	33	146	42	176	1477.5	13	4
	MD relaxed	46	216	61	235	1997.1	30	24
3A-1MMD	Docked	39	195	42	195	1678.8	4	4
	MD relaxed	45	202	63	236	1994.9	24	21
3A-1Q5G	Docked	41	177	55	220	1753	19	12
	MD relaxed	51	238	62	271	2205.3	35	25
3A-2OVK	Docked	40	133	50	164	1468	16	5
	MD relaxed	60	244	69	294	2707	40	24
Holmes Model	EM-map	57	230	65	271	2243.9	41	23

Ires: Residue numbers are involved in the actomyosin binding surface. Inat: Atom numbers are involved in the actomyosin binding surface. Surf-area: The binding surface area of actomyosin complexes. HB: Formed hydrogen bonds are involved in the actomyosin binding surface. SB: Formed salt-bridge bonds are involved in the

actomyosin binding surface.

5.2.4 Binding interface analysis of different actomyosin complexes

We analyzed the interfacial residues of the actin-binding surface in different actomyosin complexes (see Figure 27). Loop 3, loop 4, the CM-loop, loop 2, and the lower 50 kDa domain as five main regions of myosin are involved in the actomyosin interface.
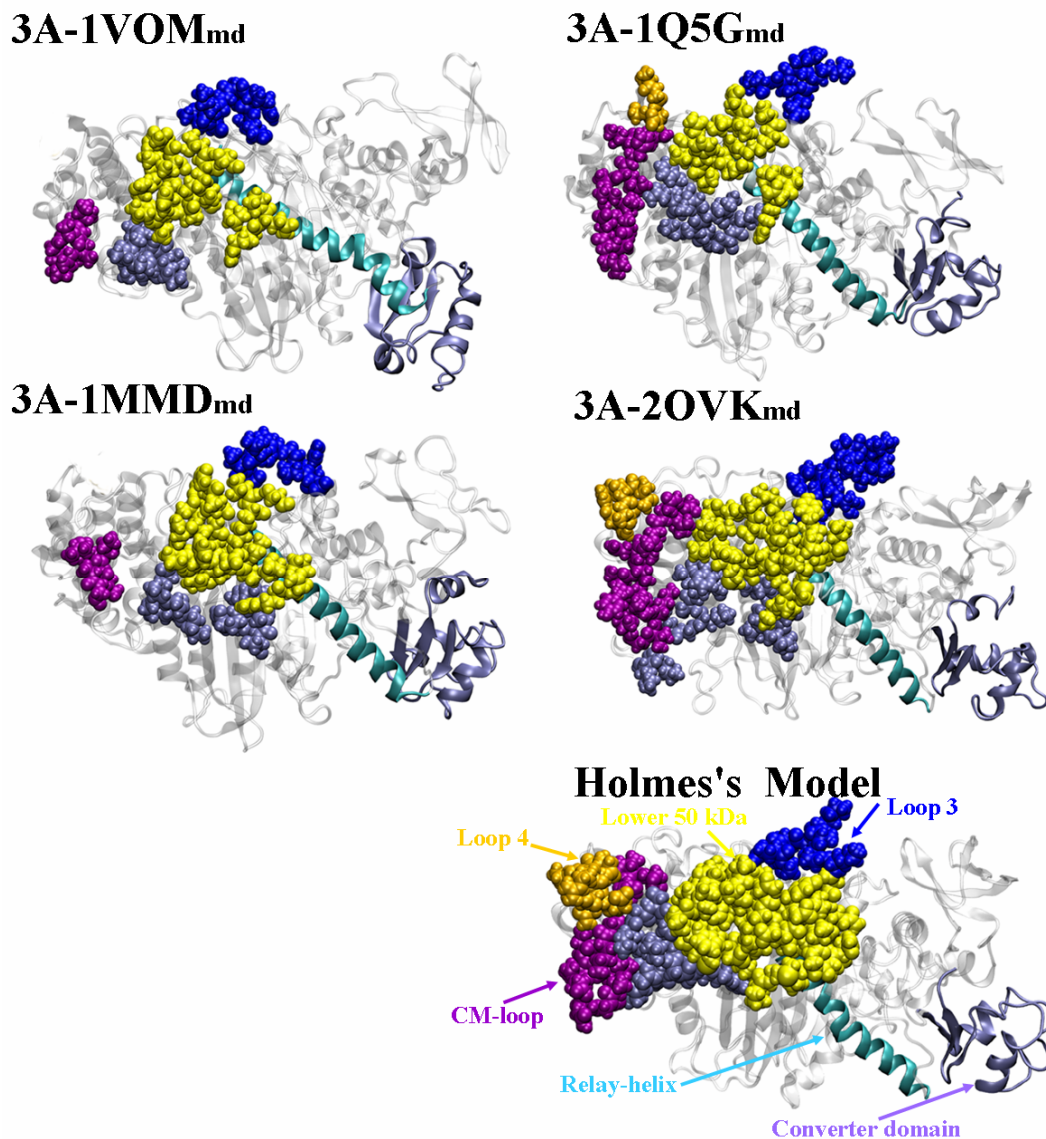

Figure 27: Actin-binding interface of myosins. These color-codes for the actin-binding interface of myosin are presented by the drawing method of space-filled and cartoon (Orange: loop 4; Purple: CM-loop; Dark ice blue: loop 2, Cyan: relay-helix; Yellow: lower 50 kDa domain; Light ice blue: converter domain; Blue:

loop 3).

Loop 3 (560-570 in *Dictyostelium* myosin) established significant contacts with the second actin monomer (AC1) in all of the four actomyosin complexes. This interaction is referred to the "Milligan contact", which has been confirmed by Milligan's zero-cross-linking experiments [47]. We found a salt-bridge interaction between Glu99/Glu100 of the second actin (AC1) and Arg562 of the myosin loop 3, which has been proved as the major contribution to the Milligan contact. DNase I binding loop of actin (residue 38-52) formed an interaction with loop 3 of myosin in the strong actin-binding states (3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$). In the weak actin-binding states (3A-1VOM$_{md}$ and 3A-1MMD$_{md}$), several H-bonds were formed between Asn541, Thr542 of myosin and Gln49, Lys50 of the DNase I binding loop of actin, which have also been mentioned by Khaitlina SY [152]. In addition, some other contacts in myosin/actin e.g., Lys546/Met44, Ser564/Arg95 and Lys565/Tyr91 were also found in the *Dictyostelium* actomyosin and Holmes's model, but they were not presented in the squid actomyosin (3A-2OVK$_{md}$) (see Table 13).

Table 13: Actin binding interface of myosin loop 3

Mysoin	Actomyosin Complexes				
	3A-1VOM$_{md}$	3A-1MMD$_{md}$	3A-1Q5G$_{md}$	3A-2OVK$_{md}$	Holmes's model
LP3-SB	ARG562-GLU99 92.57	ARG562-GLU99 93.72	ARG562-GLU99 ARG562-GLU100 96.89	LYS567-ASN92 87.54	LYS569-GLU99
LP3-HB		ARG562-THR126 88.47	PHE563-ARG95 92.7	LYS567-GLU99/100 93.68	PRO570-ARG95
	SER564-ARG95 88.25		SER564-ARG95 91.55		ALA571-TYR91
	LYS565-TYR91 91.5	LYS565-TYR91 89.3	LYS565-TYR91 92.1	LYS572-ASP1/GLU2 92.1	LYS572-ARG95
	GLU567-ARG95 79.75	GLU567-ARG95 82.96		LYS572-GLU100 83.7	ALA575-ARG95
				GLU577-ARG95 85.69	GLU576-ARG95

LP3-SB: Salt-bridge bonds were formed between the myosin loop 3 and actin. LP3-HB: Hydrogen bonds were formed between the myosin loop 3 and actin. Numbers represent the occurrence rate of bonds above 70% in the MD simulations. In

the type of the bonds formation, the front one and later one are residues from myosin and actin, respectively.

Loop 4 (360-370 in *Dictyostelium* myosin) and the CM-loop (390-410 in *Dictyostelium* myosin) showed a number of potential electrostatic interactions with actin interfacial residues in the strong actin-binding state, which also have been verified by Sutoh K [153]. The R405N mutation of skeletal myosin (identical to the R402N mutation in the *Dictyostelium* myosin) causes a severe familial hypertrophic cardiomyopathy and results weaker rigor actin-binding affinity [50]. This interaction between R402 of *Dictyostelium* myosin and actin was found only in our strong actin-binding state, but it was not involved in the weak actin-binding states (3A-1VOM, 3A-1MMD). In addition, in our simulations two residues (Gly401 and Leu399) of myosin formed hydrogen bonds with residues (Pro27 and Arg28) of actin also only in our strong actin-binding states (3A-1Q5G, 3A-2OVK) (see Table 14).

Loop 2 (620-630 in *Dictyostelium* myosin), as an important interfacial region of myosin to actin, is one of the major determinants for the weak and strong actin-binding structures. Experimental results showed that double-alanine mutation of loop 2 (K622A and K623A in *Dictyostelium* myosin) affects the actin-binding cleft closure and the weak-to-strong actin-binding transition [147]. So as to discover detailed functional mechanism of loop 2 *in silico* simulations, we investigated its structure orientations and the interaction patterns in all four MD relaxed actomyosin complexes. In Holmes's model, these two lysines (Lys636 and Lys637) homologous to Lys622 and Lys623 in *Dictyostelium* myosin play two distinct functional roles. The conformation of loop 2 was stabilized by Lys636 via formed H-bonds. Lys637 by moving its side chain diminished the interactions with Asp24/Asp25 of actin and formed a strong salt-bridge bond with the residue Glu538 of myosin (see Figure 28). In 3A-1Q5G$_{md}$, Lys623 formed a salt-bridge bond with Asp530 of myosin (homologous to Glu538 in Holmes's model) suggesting that the actin-binding cleft closure was related to structural orientations and the interaction patterns of myosin loop 2. Furthermore, Lys622 was stably connected to Asp24/Asp25 of actin with two salt-bridge bonds. These interactions were not found in 3A-2OVK$_{md}$ because of

unusual length and conformation of loop 2 in the squid myosin. Except that, the side chain of Lys623 moving to another direction formed a salt-bridge with Asp3 of the actin N-terminus. In 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$, Lys622 merely formed a salt-bridge bond with Asp25 of actin instead of formed with both Asp24 and Asp25. Lys623 revealed a similar interaction pattern as its roles in 3A-2OVK$_{md}$ to interact with Asp3 of the actin N-terminus (see Table 14).
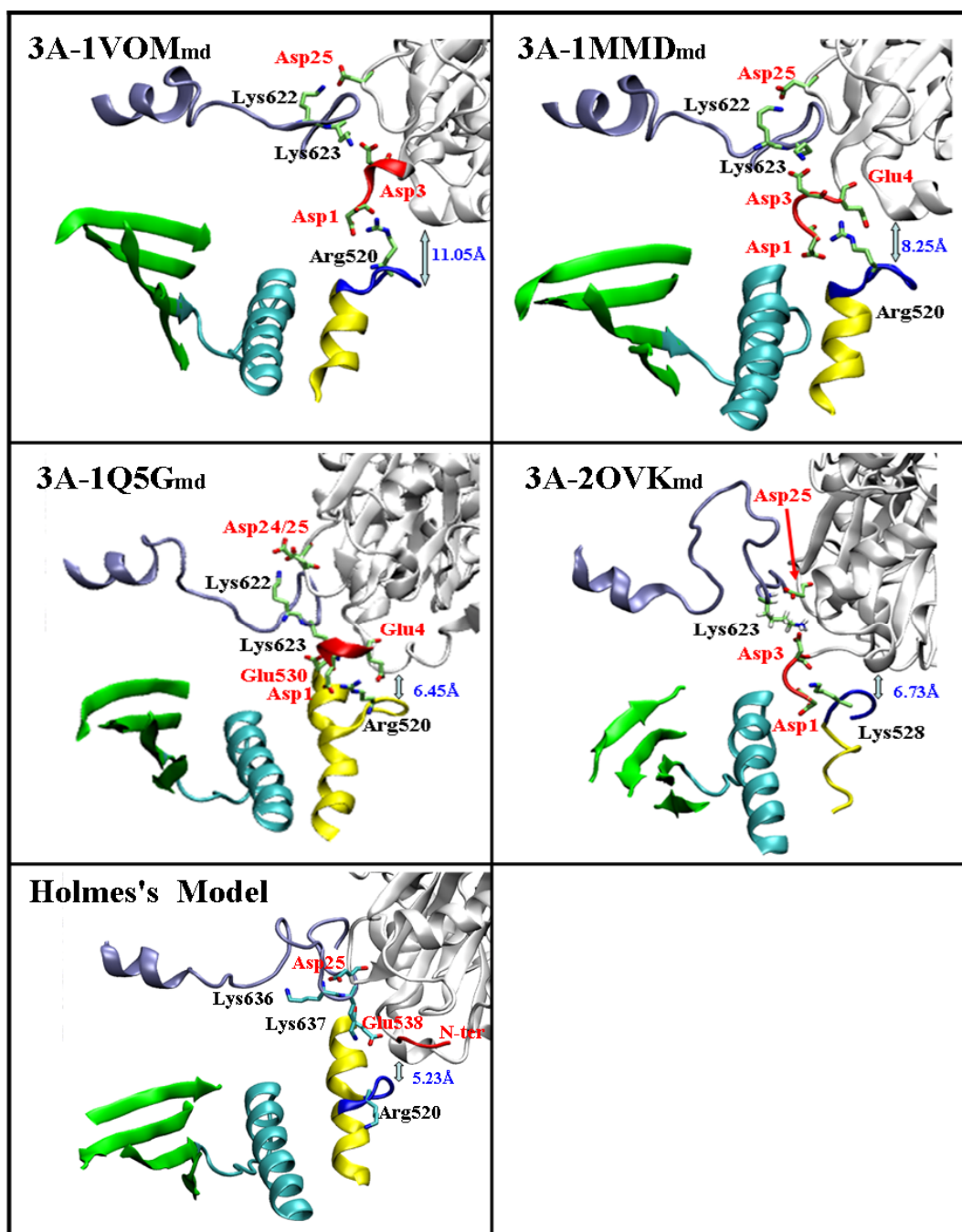

Figure 28: Interactions between N-terminus of actin and the myosin loop 2 as well as the activation loop. These color-codes are represented by drawing method of

cartoon (Dark ice blue: loop 2, Cyan: relay-helix; Yellow: lower 50 kDa domain; Green: central β-sheet; Blue: activation loop; Red: N-terminus of actin; White: actin monomer). These distances shown on this figure were measured between the activation loop of myosin and C-terminus of actin.

Table 14: Actin binding interface of myosin upper 50 kDa domain

Myosin	Actomyosin Complexes				
	3A-1VOM$_{md}$	3A-1MMD$_{md}$	3A-1Q5G$_{md}$	3A-2OVK$_{md}$	Holmes's model
LP2-SB					GLU629-ARG28
			LYS622-ASP24	LYS640-ASP3	
	LYS622-ASP25	LYS622-ASP25	LYS622-ASP25	LYS640-GLU4	LYS640-ASP25
	89	84.38	94.16	92.47	
		LYS635-ASP1	LYS635-ASP1	LYS640-ASP25	LYS642-GLU334
		89.4	92.34	90.34	
LP2-HB				SER641-ASP24	GLY635-ASP24
				72.38	
	LYS623-ASP3	LYS623-ASP3	LYS623-SER350	ALA642-ASP24	LYS636-GLY23
	82.17	78.24	85.42	75.94	
		LYS623-THR5		ALA643-ASP25	LYS637-SER348
		75.95		88	
				LYS652-GLU2/4	GLY638-GLY23
				92.98	
				LYS657-GLU2/4	LYS642-GLU334
				83.67	
				LYS660-GLU1/2	GLY643-SER145
				89.5	
LP4-SB					ARG371-ASP311
			GLU365-ARG147	GLU373-ARG147	GLU372-LYS328
			89.2	86.71	GLU372-ARG335
					GLU373-ARG147
LP4-HB					ARG371-LYS328
					ARG371-ILE329
					ARG371-GLN314
CM-LP-SB			ARG402-GLU93	LYS407-ASP25	
			88.21	91.3	
CM-LP-HB			GLY401-PRO27	GLN415-GLU334	GLU411-LYS336
					ASN410-TYR337
			92.91	87.24	ASN410-ARG28
			GLN407-GLU334	GLY416-GLU334	
			92.27	82.91	
			LEU399-ARG28		
			78.59		

LP2-SB: Salt-bridge bonds were formed between the myosin loop 2 and actin.

LP2-HB: Hydrogen bonds were formed between the myosin loop 2 and actin. LP4-SB: Salt-bridge bonds were formed between the myosin loop 4 and actin. LP4-HB: Hydrogen bonds were formed between the myosin loop 4 and actin. CM-LP-SB: Salt-bridge bonds were formed between the myosin CM-loop and actin. CM-LP-HB: Hydrogen bonds were formed between the myosin CM-loop and actin. Numbers represent the occurrence rate of bonds above 70% in the MD simulations. In the bonds formation, the front one and later one are residues from myosin and actin, respectively.

To summarize the results in this section, we found all those interactions in the myosin loop 3, loop 4 and the CM-loop, which were demonstrated in previous experimental kinetic studies using actin mutants [154-156] or myosin mutants [147, 157]. The conformation of loop 2 is stabilized by two lysines (Lys622 and Lys623 in *Dictyostelium* myosin) and interacted with Asp24/Asp25 of the actin. However, structure orientations and the interaction patterns of the myosin Lys623 to the lower 50 kDa domain of myosin or N-terminus of actin in distinct actomyosin complexes are probably coupled with actin-binding cleft closure. In other words, various features of the myosin Lys623 as one of the novel findings effectively interpret the potentially functional mechanism of loop 2 related to the weak-to-strong actin-binding transitions.

5.2.5 Interaction of myosin activation loop with the actin N-terminal region

According to previous results, residues 530-537 (helix-loop-helix region) belonging to the lower 50 kDa domain of *Dictyostelium* myosin preserve strong hydrophobic interactions with residues 143-147 of actin [45]. Except that, we showed that activation loop of myosin (residues 519-524) play a key role in the actin binding and the actin activation of myosin ATPase activity in our previous experiments in the weak actin-binding state (state IV in the Lymn-Taylor cycle, see figure 4) [55]. Two prolines Pro529, Pro530 of the activation loop in Holmes's model are homologous to Pro521, Pro522 of *Dictyostelium* myosin. Prolines are able to promote this activation loop interacting with Leu349-Phe352 of the actin C-terminus in 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$. This interaction was assumed in strong actin-binding state by Morales

MF [52, 53] (see Figure 28). However, this hydrophobic interaction has not been detected in the weak actin-binding states 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$. These results point out that this hydrophobic interaction probably impels myosin and actin more closely bound to each other in the weak-to-strong actin-binding transition, which is also a significant sign for the discrimination of the weak and the strong actin-binding state. These results are in agreement with previous hypotheses that this hydrophobic interaction creates expanding networks of the actin-myosin interface during the weak-to-strong actin-binding transition [158].

We found a novel actomyosin interaction as Arg520 in the activation loop of *Dictyostelium* myosin interacted with the negatively charged N-terminal segment of actin [55] (see Table 15). It is interesting that in the protein-protein docking procedures this interaction was not established, while it was formed spontaneously in the first few nanoseconds during MD simulations in the four actomyosin complexes. In Holmes's model, Arg528 of chicken myosin (homologous to Arg520 of the *Dictyostelium* myosin) has no interactions with negatively charged N-terminus of actin since this region of actin was not resolved in the model. In 3A-1VOM$_{md}$, the interaction between Arg520 of myosin and Asp1 of actin impels Pro522 of myosin away from Ser350 of actin with the distance of 11.05Å, i.e. a long distance between the activation loop of myosin and C-terminus of actin prevents them from interacting with each other. This salt-bridge (Arg520/Asp1) is also presented in the 3A-1MMD$_{md}$, but the down-lever position of myosin induces the activation loop further approaching to the C-terminus of actin with the distance of 8.25Å. Arg520 forms strong contacts with residues Asp1 and Asp3/Glu4 of actin in 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$, which promote the myosin activation loop approaching to C-terminus of actin with the value of 6.45Å, 6.73Å, respectively (see Figure 28).

To summarize the results in this section, activation loop described above formed strong hydrophobic interactions with the actin C-terminus in the strong actin-binding state has been demonstrated by recent experiments [52, 53] and *in silico* simulations [54], but the functions and interaction patterns of this activation loop were not discovered in the weak actin-binding state before. The most prominent points of this

section is that we found a novel actin-binding residue (Arg520) in the myosin activation loop, which appears various binding patterns in different actomyosin complexes. Salt-bridge interactions between Arg520 of myosin and four negatively charged residues in the N-terminus of actin might be a starting engine for the activation loop approaching to the actin C-terminus or other inner conformational changes of myosin in the power stroke.

Table 15: Actin binding interface of myosin lower 50 kDa domain

Myosin	Actomyosin complexes				
	3A-1VOM$_{md}$	3A-1MMD$_{md}$	3A-1Q5G$_{md}$	3A-2OVK$_{md}$	Holmes's model
520-LP SB	ARG520-ASP1 93.37	ARG520-ASP1 ARG520-GLU4 94.17	ARG520-ASP3 ARG520-GLU4 96.46	LYS528-ASP1 LYS528- GLU4 92.2	
520-LP HB	GLN521-GLN354 48.5	GLN521-GLN354 56.98	GLN521-SER351 GLN521-GLN354 89.11		PRO529-THR351
LOWER50-HB			GLU531-THR351 92.6		
	ASN537-ARG147 98.5	ASN537-LEU142 ASN537-TYR143 95.64 ASN537-ARG147 85.7	ASN537-GLY146 94.75 ASN541-LYS50 91.23	LYS544-GLY48 91.27	SER549-GLN49
			THR542-GLN49 90.64	ASN552-GLN49 ASN552-LYS50 93.5	ASN552-LYS50
					LYS553-GLY46 LYS553-GLY48
	LYS546-MET44 80	LYS546-MET44 84.13	LYS546-MET44 95.8		

520-LP-SB: Salt-bridge bonds were formed between Arg520 of the activation loop and actin. 520-LP-HB: Hydrogen bonds were formed between Arg520 of the activation loop and actin. LOWER50-HB: Hydrogen bonds were formed between lower 50 kDa domain and actin. Numbers represent the occurrence rate of bonds above 70% in the MD simulations. In the bonds formation, the front one and later one are residues from myosin and actin, respectively.

5.2.6 Actin induced conformational changes of the actin-binding cleft of myosin

In order to identify the structural background of actin binding induced myosin

conformational changes, the analysis is firstly focused on the actin-binding cleft of myosin in different actomyosin complexes. The actin-binding cleft of 3A-1Q5G$_{md}$ and 3A-2OVK$_{md}$ are in the closed form different from while it is open in 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$ [5, 159]. Evidences for actin-binding induced open-closed cleft transition were observed from previous experiments [57, 160]. We were interested in whether actin induces intermolecular strain in myosin which relaxes in nanosecond timescale.

In order to characterize atomic structure re-arrangement occurred in the actin-binding cleft, we measured the distance based on the backbone carbon atom between the alpha carbon atoms of Glu365 located in the myosin loop 4 and Asn537 located in the helix-loop-helix element of the lower 50kDa domain helix-loop-helix of the myosin according to previous measurements [37]. In 3A-1VOM$_{md}$, partial cleft closure was observed as the distance decreased from 21.7Å to 20.57Å (see Table 16). In 3A-1MMD$_{md}$, we only observed minor conformational changes in the actin-binding cleft closure. In comparison with no cleft closure occurred in the MD simulations of single myosin in the absence of actin, these results suggest that the actin binding causes partial cleft closure in the weak actin-binding states. In 3A-2OVK$_{md}$, the partially closed cleft of this the rigor actomyosin complex was almost completely closed in the first few nanoseconds of the MD simulations (from 18.25Å to 16.82Å). Similar effect was observed in the other rigor actomyosin complex (3A-1Q5G$_{md}$), however, the closure was even more complete (from 16.72 Å to 11.91 Å). Meanwhile, we compared our two strong actin-binding states with Holmes's model. The distance value is 17.23Å in Holmes's model similar to what we found in 3A-2OVK$_{md}$ (16.82Å), but it was more closed in 3A-1Q5G$_{md}$ (11.91Å).

Notable distance alterations taken place in the two strong actin-binding states suggest that the actin-binding cleft in the closed form is also affected by actin binding. The feasible cause of actin-binding cleft further closure in the strong actin-binding states is the docked myosin still in a rigor-like state with incompletely closed actin-binding cleft, which is different from the rigor state, the state of theoretically assumed end state of the power stroke. Degree of partially cleft closure of the myosin up-lever

conformation (3A-1VOM$_{md}$) is more conspicuous than that of the myosin down-lever conformation (3A-1MMD$_{md}$). It means that the lever up-to-down movement is directly related to the actin binding induced cleft closure. Therefore, it might be interesting to discover what conformational changes of myosin contribute to this actin-binding cleft partial closure in 3A-1VOM$_{md}$.

According to previous results of the actin-binding surface analyses, we realized that some surface loops of myosin i.e. the CM-loop, loop 2, loop 3 and the activation loop were involved in actin-binding [147, 157]. We found a strong salt-bridge bond cluster formed between Arg520 in the myosin activation loop and the four negatively charged residues at N-terminal region of actin in both weak and strong actin-binding states, which is different from previous concept that there are not any interactions unearthed between the myosin activation loop and actin in the weak actin-binding state. In this case, we suppose that the formation of this salt-bridge might be involved in the actin-binding cleft closure in the weak actin-binding states.

Table 16: Measurement in conformational changes of the actin-binding cleft

Name	Distance(Å)	Name	Distance(Å)	Name	Distance(Å)
1VOM	21.1	1VOM$_{md}$	21.7	3A-1VOM$_{md}$	20.57
1MMD	21.28	1MMD$_{md}$	21.37	3A-1MMD$_{md}$	21.14
1Q5G	16.67	1Q5G$_{md}$	16.72	3A-1Q5G$_{md}$	11.91
2OVK	18.3	2OVK$_{md}$	18.25	3A-2OVK$_{md}$	16.82
				Holmes's model	17.23
				3A-1VOM-520MU	23.66
				3A-1VOM-LP3MU	21.76

The conformational changes of the actin-binding cleft in different actomyosin complexes were calculated on the basis of distance between the backbone carbon atom from myosin Glu365 and atom from myosin Asn537.

Based on this hypothesis, we introduced a R520Q mutant called 3A-1VOM-520MU to test the interactive effects of the myosin activation loop with actin and a R562Q mutant in the myosin loop 3 called 3A-1VOM-LP3MU to compare the effects of different mutants in the weak actin-binding states (3A-1VOM$_{md}$). The same 100 ns

MD simulations as their wild type structure (3A-1VOM$_{md}$) have been performed on these two mutants. Distinctive features in conformational changes in the actin-binding cleft showed in these two mutants were different from their wild type structure. In 3A-1VOM-LP3MU, the measured distance of the actin-binding cleft is 21.76 Å, which is similar to the distance value of the MD relaxed single myosin 21.1 Å (see Table 16). The cleft partial closure is completely offset by the R562Q mutant. This suggests that the cleft closure is sensitive to the conformational changes of the myosin loop 3. In 3A-1VOM-520MU, originally formed salt-bridge interaction was broken immediately only after a short time of the MD simulation, which resulted in further open of the actin-binding cleft with the distance value of 23.66 Å. These results imply that the cleft closure is not only sensitive to the conformational changes of loop 3, but also more influenced by the salt-bridge cluster formed between the myosin activation loop and N-terminus of actin. Arg520 based salt-bridge interaction is not the only one, but is one of the main contributions to the actin-binding cleft partial closure in the weak actin-binding state.

To summarize the results in this section, the actin binding causes a partial closure of the actin-binding cleft in both weak actin-binding states and striking closure in both strong actin-binding states. This cleft closure is not only sensitive to the conformational changes of loop 3, but also strongly influenced by the salt-bridge cluster formed between the myosin activation loop and N-terminus of actin.

5.2.7 Actin-binding induced conformational changes of the converter domain

The actin-binding cleft closure and the lever arm up-to-down movement are the two most important events in the power stroke. According to Arg520- actin N-terminus salt-bridge interaction is one of the main contributions to the actin-binding cleft partial closure in the weak actin-binding states, we analyzed the lever arm movements through the rotation of the converter domain after actin binding. We defined an angle, which represents the rotation of the myosin converter domain. The most striking actin induced event was that upon actin binding the converter domain underwent a further

up movement from 122.73º (1OVM$_{md}$) to 105.03º (3A-1VOM$_{md}$) (see Table 17). In contrast, the converter domain underwent a further down movement from 203.96º (1MMD$_{md}$) to 227.22º (3A-1MMD$_{md}$). Meanwhile, the angle of the converter of 3A-1Q5G$_{md}$ (221.73º) upon actin binding was still kept in its original position as similar as 1Q5G$_{md}$ (220.69º). In the squid myosin, there was a considerable angle change of the converter domain towards the down position from 220.58º (2OVK$_{md}$) to 245.27º (3A-2OVK$_{md}$). The defined angle of the converter domain in 3A-2OVK$_{md}$ has a similar value to the measured angle in Holmes's model (237.9º). If we combined these results with previous conformational changes in the actin-binding cleft, we assume that the actin-binding cleft further closure is coupled with the lever arm further up movement after actin binds to myosin, which might be considered as the first phase of the power stroke. Owing to none of intermediate states has been uncovered yet. We are still not clear what had happened in the second phase. Although we observed the actin-binding cleft further closure coupled with lever arm down position movement in 3A-1MMD$_{md}$, this state was only available under the premise of which the lever movement prior to the cleft closure. So, these two parallel-coupled movements might be the second phase or the third phase, which has also been indicated by the actomyosin structure 3A-2OVK$_{md}$ with these two movements well coupled. The final phase can be speculated by analyzing the actomyosin structure of 3A-1Q5G$_{MD}$ in the actin-binding cleft closure occurred after the lever arm down position movement. In order to discover what the role of the activation loop plays in the lever arm movement, we also analyzed the angle changes in two mutants. In 3A-1VOM-520MU, only ~2 º angle change was emerged by comparing to the MD relaxed single myosin (1VOM$_{md}$) in the absence of actin. In 3A-1VOM-LP3MU, there was a further up rotation in the converter domain (111.07º), but the amplitude of this movement was a little bit smaller than that of the 3A-1VOM$_{md}$ actomyosin complex (see Table 17). The mutational results suggest that Arg520 in the activation loop induces the converter domain from an up- to a further-up rotation, but loop 3 is not involved in this converter domain further up movement (see Figure 29).
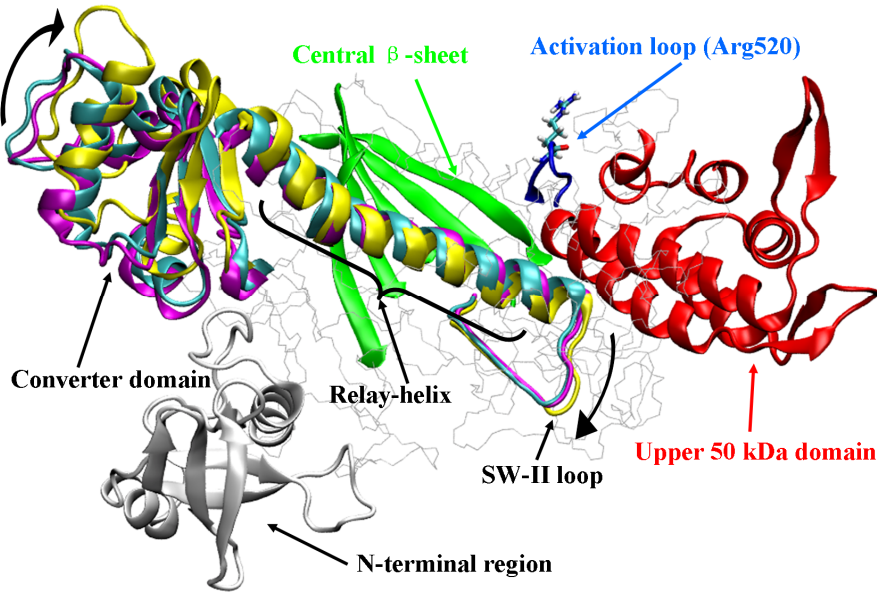
Figure 29: Conformational changes of the myosin functional regions in different actomyosin complexes. These color-code represent weak actin-binding complex 3A-1VOM$_{md}$ (yellow) and two mutants (3A-1VOM-520MU (magenta), 3A-1VOM-LP3MU (cyan)). (Green: central β-sheet; Red: upper 50 kDa domain; Blue: activation loop; and White: N-terminus of myosin).

Table 17: Angle changes of converter domain in different actomysin complexes

Name	Angle(°)	Name	Angle(°)	Name	Angle(°)
1VOM	121.52°	1VOM$_{md}$	122.73°	3A-1VOM$_{md}$	105.03°
1MMD	201.33°	1MMD$_{md}$	203.96°	3A-1MMD$_{md}$	227.22°
1Q5G	221.02°	1Q5G$_{md}$	220.69°	3A-1Q5G$_{md}$	221.73°
2OVK	221.14°	2OVK$_{md}$	220.58°	3A-2OVK$_{md}$	245.27°
				Holmes's model	237.9°
				3A-1VOM-520MU	120.12°
				3A-1VOM-LP3MU	111.07°

Angle changes of the converter domain were measured by two residues Gly684, Arg689 from "SH1 helix" and Ala748 at the end of the converter domain. Angle calculation was based on the CA atom of three residues with Arg689 in the middle.

5.2.8 Actin-binding induced conformational changes of SW-II loop of myosin

Based on the coupling mechanism of the converter domain rotation and relay-helix connected to the SW-II loop open-closed transition [39, 64], we did a structural comparison between the weak actin-binding state of the wild type (3A-1VOM$_{md}$) and

two mutants in these functional regions of myosin at an atomic level.

Lys462 residing in the myosin SW-II loop is extremely sensitive to its surroundings, which result in Lys462 to be a guard residue in controlling the Pi release upon the lever swinging of the power stroke [161]. In 1VOM$_{md}$, Lys462 with its side chain protrudes to the solvent without any bonds formation. In 3A-1VOM$_{md}$, Lys462 formed a stable salt-bridge with myosin Glu580 that is located at the end of the "wedge" β-hairpin region (571-575 of the *Dictyostelium* myosin). This bond formation assisted with four residues based a hydrophobic cluster contributes to the observed movements of the "wedge" region towards the lower 50 kDa domain upon the actin binding. The first two residues Ile529 and Leu543 (*Dictyostelium* myosin) are located in the helix-loop-helix. Phe466 is located at the N-terminal region of the relay-helix and Trp584 is close to the C-terminus of the "wedge" region. This hydrophobic cluster pushes the SW-II loop to a further closed conformation via the Lys462-Glu580 salt-bridge bond formation (see Figure 30). However, this hydrophobic cluster with only Ile529-Trp584 bond formed appears much weaker in the absence of actin (1VOM$_{md}$), which could not facilitate the Lys462-Glu580 bond formation and spread to the further closure of the SW-II loop. Lost interactions with the N-terminal region of actin in 3A-1VOM-520MU, the further closure of the SW-II loop was suppressed, because this hydrophobic cluster could not transform the information of the actin binding to the activation loop, since the mutation prevented the interaction. In order to know whether the SW-II loop further closure through the hydrophobic cluster is based on the interactions between the activation loop and N-terminus of actin? This idea is supported by the fact that the distance between Ile529 and Phe466 shortened to 3.1 Å from 6.99 Å in 3A-1VOM-520MU. In 3A-1VOM-LP3MU, the distance is 3.28 Å similar to its wild type structure (3A-1VOM$_{md}$). Phe466 is located between Ile529 and Leu543 in structures of 3A-1VOM$_{md}$ and 3A-1VOM-LP3MU, while Phe466 reorients its side chain to the direction without bond formation in 3A-1VOM-520MU.
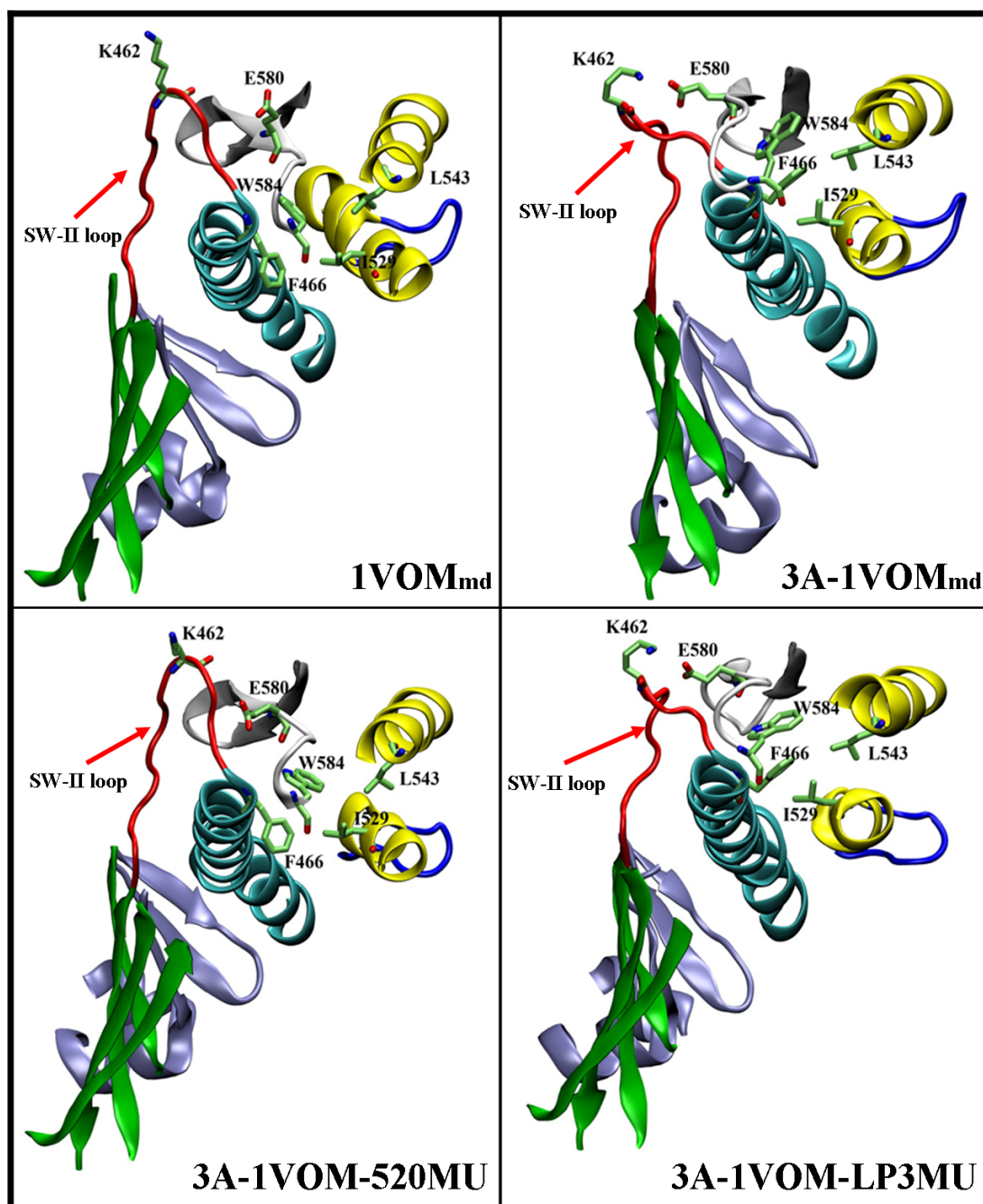
Figure 30：Actin binding induces conformational changes in the myosin SW-II loop. The relay-helix (cyan) is close to the lower 50 kDa domain (yellow). Converter domain and central β-sheet are colored by ice blue and green. Two loops are colored with red (the SW-II loop) and blue (the activation loop). Phe584 closed to the "wedge" region (white) forms a hydrophobic cluster with two residues (Ile529, Leu543) in the helix-loop-helix and one residue (F466) in the N-terminus of the relay-helix.

Consequences of the further closure of the SW-II loop affected by the Lys462-Glu580 salt-bridge and hydrophobic cluster probably bring about the conformational changes

of the nucleotide-binding pocket, which can be characterized through the distance changes between the SW-I loop and the SW-II loop of myosin (see Figure 31). $1VOM_{md}$ and 3A-1VOM-520MU display similar distances in the nucleotide-binding pocket with the value of 10.37±0.2Å and 10.21±0.19 Å respectively. However, noticeable changes in the nucleotide-binding pocket closure were observed in $3A-1VOM_{md}$ and 3A-1VOM-LP3MU with the values of 9.44±0.22Å and 9.53±0.18Å. In summary, we conclude that the salt bridge between the activation loop of myosin and N-terminus of actin directly regulates the further rotation of the converter domain, the partial closure of the actin-binding cleft and the further closure of the nucleotide-binding pocket.

To summarize the results in this section, the further closure of the SW-II loop in the first phase of the power stroke is pushed by a salt-bridge bond and a hydrophobic cluster between the activation loop and the relay-helix of myosin via interactions between Arg520 in the activation loop and the N-terminal region of actin.
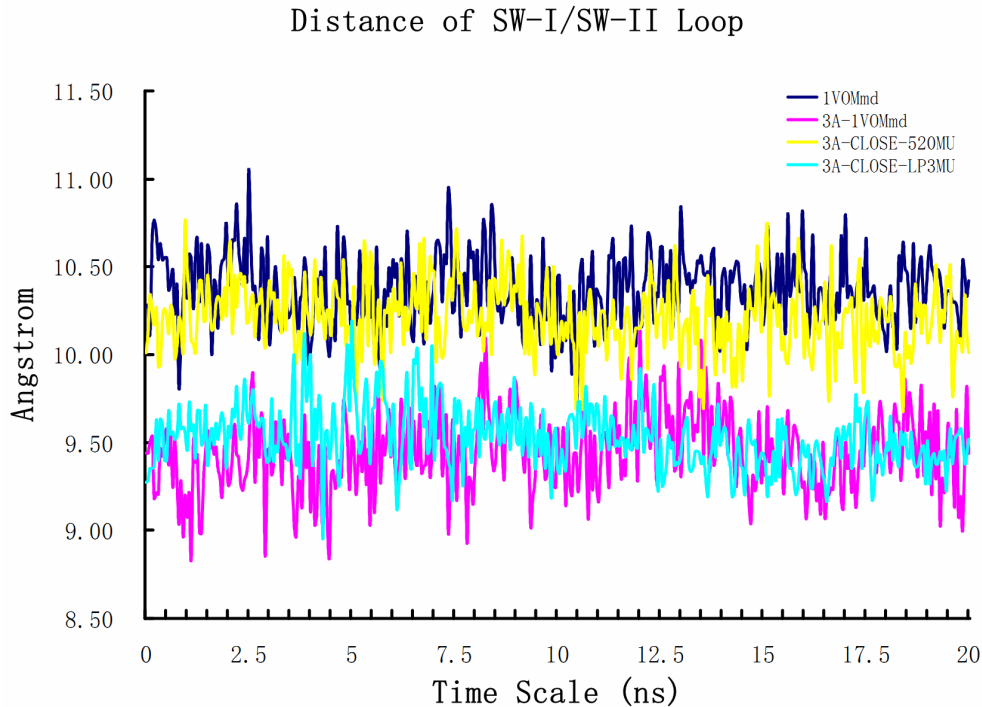
Figure 31: The distance changes between the SW-I loop and the SW-II loop. The distance was calculated between the backbone coordinates of the SW-I loop (235-241 in *Dictyostelium* myosin) and the SW-II loop (456-462 in *Dictyostelium* myosin). The colors represent $1VOM_{md}$ (blue), $3A-1VOM_{md}$ (purple), 3A-1VOM-520MU (yellow) and 3A-1VOM-LP3MU (cyan).

5.2.9 Motional correlations of myosin functional regions

According to the results of conformational changes in the actin-binding cleft, the converter domain and the SW-II loop, we suppose that there is a motion-coupled mechanism of these myosin functional regions. As we discovered the interactions between the activation loop and the four negatively charged residues in the N-terminal region of actin are intensely involved in the conformational changes in these myosin functional regions upon actin binding. In order to get further insights into the functional mechanism of the activation loop after actin binding, we analyzed the motional correlations obtained from the trajectories of the MD simulations. Motional correlation is defined as when the motion is increased in one region, the motions of the other regions are also increased (defined as correlation), nonexistent (defined as non-correlation), or decreased (anti-correlation).

In the absence of actin, we found relatively weak motional correlations between the activation loop (residue 519-523 in *Dictyostelium* myosin) and other parts of myosin in all studied structures (1VOM$_{md}$, 1MMD$_{md}$ and 1Q5G$_{md}$) (see Figure 32). When actin binds to myosin, some structural elements of myosin exhibit high motional correlations with the activation loop in three actomyosin complexes (3A-1VOM$_{md}$, 3A-1MMD$_{md}$ and 3A-1Q5G$_{md}$). Firstly, the lower 50 kDa domain is highly correlated with the activation loop in all three actomyosin complexes. Loop 4 (residue 360-370 in *Dictyostelium* myosin) from the upper 50 kDa domain is weakly correlated with the activation loop in the two open-cleft actomyosin complexes (3A-1VOM$_{md}$, 3A-1MMD$_{md}$), but the correlation becomes stronger in the closed-cleft actomyosin complexes (3A-1Q5G$_{md}$), which indicates that loop 4 is involved in the actin-binding cleft open-closed transition in the power stroke. Another region from the upper 50 kDa domain is the CM-loop (residue 390-410 in *Dictyostelium* myosin), which is intensely correlated with the activation loop in the two down-lever arm structures (3A-1MMD$_{md}$, 3A-1Q5G$_{md}$), but no correlation was observed in the up-lever arm structure (3A-1VOM$_{md}$), which suggest that different binding region and pattern of the CM-loop is related with the up and the down position of the myosin lever. We also

found that loop 2 (residue 620-630 in *Dictyostelium* myosin) is correlated with the activation loop in 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$, which is caused by the four negatively charged residues in the N-terminal region of actin interacted with loop 2 and the activation loop. Meanwhile, the correlation mechanism of loop 2 coupling with the activation loop in strong actin-binding state (3A-1Q5G$_{md}$) is based on the helix-loop-helix region. Further correlations were found in the N-terminus of relay-helix (residue 465-475 in *Dictyostelium* myosin), the "strut" region (residue 590-596 in *Dictyostelium* myosin), a small helix before the "strut" region that is called "prestrut" region (residue 586-590 in *Dictyostelium* myosin) and the SW-II loop (residue 456-462 in *Dictyostelium* myosin) in the three actomyosin complexes (3A-1VOM$_{md}$, 3A-1MMD$_{md}$ and 3A-1Q5G$_{md}$). The correlation between the activation loop and N-terminus of the relay-helix is due to a hydrophobic cluster formed between them e.g. Ile529 in the activation loop, Phe466 in the N-terminus of the relay-helix. The "strut" and the "prestrut" region located between the upper and lower 50 kDa domains might be involved in negotiating the open-closed actin-binding cleft transition [162, 163]. But the mechanism of correlations between the "strut-prestrut" region and the activation loop is still unclear. SW-II loop shows correlations with the activation loop only in the 3A-1VOM$_{md}$, but not in the two down-lever arm structures (3A-1MMD$_{md}$, 3A-1Q5G$_{md}$), which means the coupling movement between the activation loop and the SW-II loop merely existed in the up-lever arm state and not in the down-lever arm state. The most astonishing findings were that SW-II loop totally lost its correlation with the activation loop in 3A-1VOM-520MU, but that was still strongly kept in 3A-1VOM-LP3MU (see Figure 32).

Although the activation loop contains many motional correlations with different functional regions of myosin, motion of the SW-II loop is somehow directly influenced by the activation loop via the salt-bridge interaction with actin. For this reason, we investigated motional correlations between the SW-II loop and other functional regions of myosin (see Figure 33). In the absence of actin, the SW-II loop is only weakly correlated with the relay-helix in 1VOM$_{md}$ and 1MMD$_{md}$. Furthermore, it also correlates with the central β-sheet in 1Q5G$_{md}$. In the presence of actin, the

SW-II loop correlates to the activation loop in 3A-1VOM$_{md}$ and 3A-1VOMLP3MU, but this correlation is not observed in the 3A-1VOM520MU.
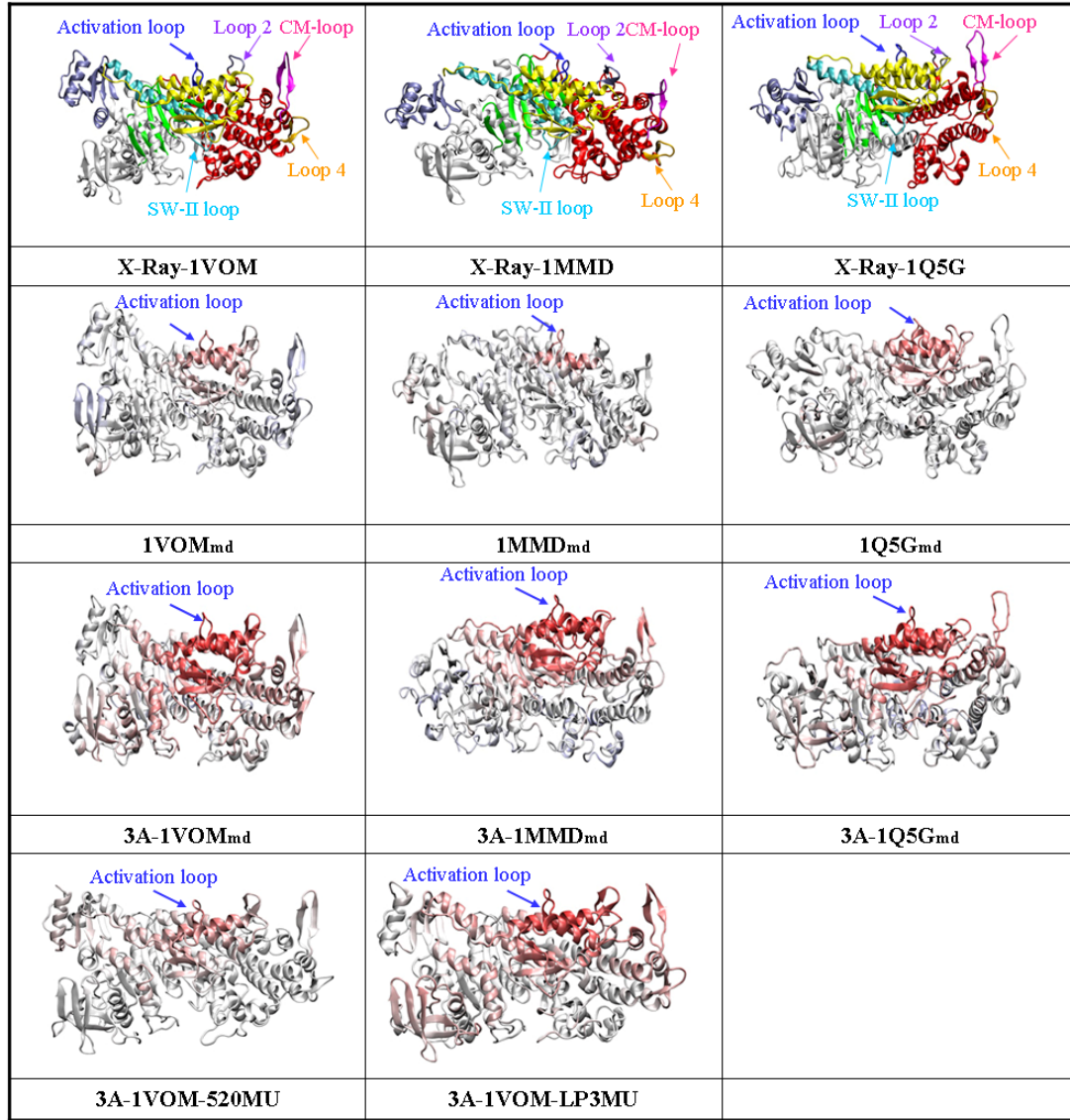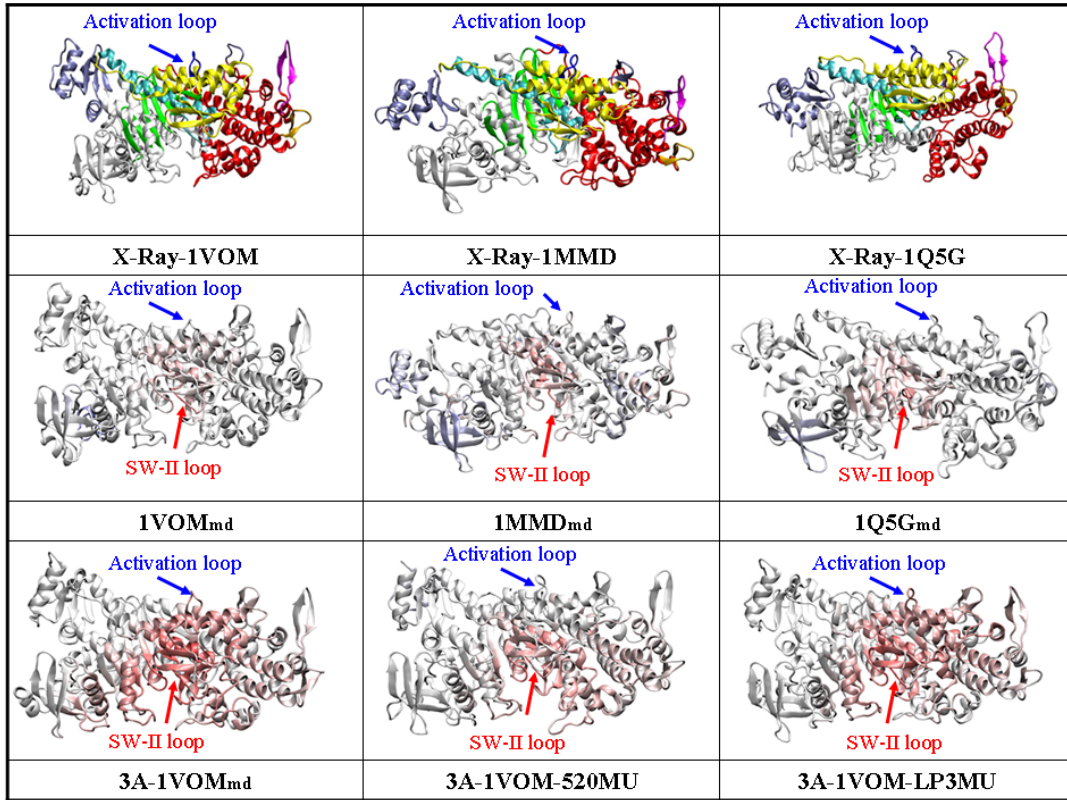

Figure 32: Motional correlations between the activation loop and other functional regions of myosin. In the first line, three colored pictures represent crystal structures of single myosin in different states (Red: upper 50 kDa domain; Yellow: lower 50 kDa domain; Green: β-sheet; Orange: loop 4; Purple: CM-Loop; Ice blue: converter domain; Cyan: relay-helix; White: others). The second third and fourth lines represent motional correlations between the activation loop and other functional regions of the MD relaxed myosin, actomyosin complexes in wild type and two actomyosin mutants. The red color in the second third and fourth line represents existed correlations and the white color means non-correlation.

Figure 33: Motional correlations between the SW-II loop and other functional regions of myosin. In the first line, three colored pictures represent structures as same as described in Figure 32. The second and third lines represent motional correlations between the SW-II loop and other functional regions of myosin, actomyosin complex in wild type and two actomyosin mutants in the weak actin-binding state. The red color in the second and third lines represents existed correlations and the white color means non-correlation.

To summarize the results in this section, the dynamics of activation loop correlate with that of loop 2, the N-terminus of relay-helix, the SW-II loop, the "strut", the "prestrut" region and so on in the weak actin-binding state. When the R520Q mutant was introduced to the weak actomyosin complex 3A-1VOM$_{md}$, the motional correlations between the activation loop and other functional region of myosin start to decrease or vanish. The motional correlation between the activation loop and the SW-II loop is extremely conspicuous in 3A-1VOM$_{md}$ and 3A-1VOMLP3MU, but this correlation is totally abolished in 3A-1VOM520MU. Therefore, we suppose that point-to-point motional correlations between the activation loop and the SW-II loop directly contribute to the SW-II loop further closure induced by the actin-binding via the interaction between Arg520 in the activation loop of myosin and four negatively

charged residues in the N-terminal region of actin.

5.3.0 Motional correlations of five myosin functional regions in the up-lever weak actin-binding state

In order to discover this potential coupling mechanism between the activation loop and SW-II loop, we focused on the motional correlation relationships in five functional regions, i.e. the activation loop, the SW-II loop, loop 2, the "prestrut" region and the N-terminal region of the relay-helix. We did this analysis in four structures, $1VOM_{md}$, $3A-1VOM_{md}$, 3A-1VOM520MU and 3A-1VOMLP3MU.

In the absence of actin, there are anti-correlations between the SW-II loop and the activation loop, the SW-II loop and the "prestrut" region, the SW-II loop and loop 2 (see Figure 34A). The N-terminal region of the relay-helix is correlated with the "prestrut" region because of a strong salt-bridge formed between Glu467 in the N-terminal region of the relay-helix and Lys587 in the "prestrut" region with the correlation value around 0.4. This result suggests that the structural movements of the "prestrut" region are coupled with the relay-helix even in the absence of actin via the salt-bridge. The effects of other functional regions of myosin are not very significant in the correlation value as their values are close to zero.

In the presence of actin (see Figure 34B), the correlation patterns changed in the $3A-1VOM_{md}$ are compared to that of the structure of $1VOM_{md}$. Five functional regions are almost coupled to each other with the average correlation values above 0.5. Besides that the correlation values are around 0.3 between the SW-II loop and the activation loop, the SW-II loop and the "prestrut" region, the SW-II loop and loop 2. The SW-II loop is anti-correlated with the above three regions in $1VOM_{md}$, but the correlation intensity between the SW-II loop and other three regions increased to ~ 0.3 in the actin-attached state is a considerable degree of enhancement by comparing with actin detached single myosin states. Four regions (N-terminal region of the relay-helix, the activation loop, the "prestrut" region and loop 2) are cross-correlated in the values above 0.5.
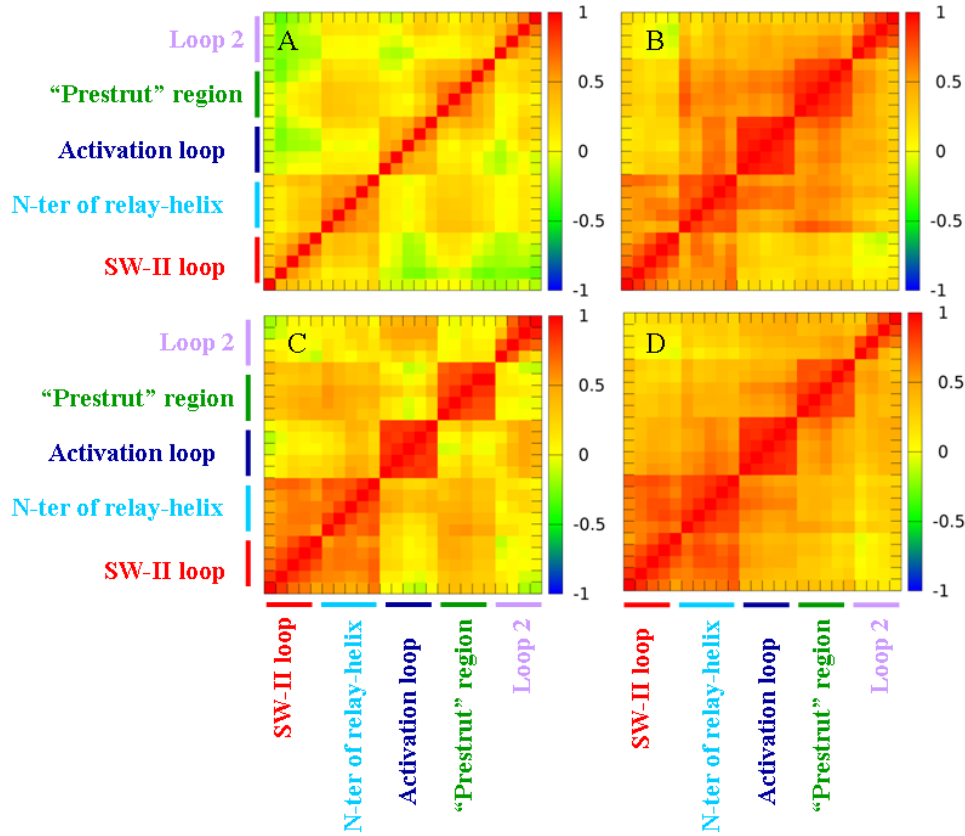
Figure 34: Motional correlations of five significant myosin functional regions. A: $1VOM_{md}$, B: $3A-1VOM_{md}$, C: 3A-1VOM520MU, D: 3A-1VOMLP3MU. The correlation value scale ranged from -1 to +1 represents from anti-correlation to correlation. Color from blue to yellow suggests motional correlation from anti-correlation to non-correlation. Motional correlations from weak to strong are represented by color from orange to red.

In 3A-1VOM-520MU (see Figure 34C), two motional correlations are decreased between the N-terminal region of the relay-helix and the activation loop, N-terminal region of the relay-helix and loop 2. The cause of reduced correlations in the first group (the relay-helix and the activation loop) is due to the N-terminal region of the relay-helix is away from the activation loop in the absence of the pushing force from the actin through the salt-bridge bond formed between Arg520 of the activation loop and the N-terminus of actin. In the second group (N-terminal region of the relay-helix and loop 2), there is an interaction between the N-terminus of actin and loop 2, which transfers motions of loop 2 through the salt-bridge bond formed between Arg520 and the N-terminus of actin up to the N-terminal region of the relay-helix. Correlation between the activation loop and the "prestrut" region is reduced to zero in

3A-1VOM-520MU. Two backbone H-bonds between Asp530, Ser533 of the helix-loop-helix and Lys588 of the "prestrut" region are formed in 3A-1VOM$_{md}$. However, such H-bonds are not formed in 3A-1VOM-520MU and 1VOM$_{md}$. R520Q mutation breaks the interactions between the activation loop and the "prestrut" region via this helix-loop-helix region. Correlation is reduced between the "prestrut" region and loop 2, which cannot be satisfactorily explained from directly structural coupling. However, we suppose that loop 2 correlation with the "prestrut" region is related to interactions between the activation loop and the N-terminal region of actin.

In 3A-1VOM-LP3MU (see Figure 34D), R562Q breaks the salt-bridge bonds formed between Lys562 of myosin and Glu99/Glu100 of actin, which shows similar correlation features in the five functional regions of myosin to those of the wild type 3A-1VOM$_{md}$.

5.3.1 Structural analysis of the myosin "prestrut" region

When actin binds to myosin, the "prestrut" region shows motional correlations with the SW-II loop, the N-terminal region of the relay-helix, loop 2 and the activation loop. This suggests an important functional role of the "prestrut" region in mediating the SW-II loop further closure through the activation loop. The structure of the "prestrut" region is conserved in all myosin families with the Glu-Lys-Asn-Lys-Asp sequence. Conformation of this structure looks like a claw surrounded by neighboring side-chains. In 1VOM$_{md}$, Glu586 and Lys589 form a salt-bridge bond to keep the "prestrut" region in a stable conformation, but this bond is not formed in two down-lever arm structures (1MMD$_{md}$ and 1Q5G$_{md}$)(see Table 18). Asp530 and Ser533 of the helix-loop-helix are connected to Asn588 of the "prestrut" region with hydrogen bonds. Lys587 is probably involved in the lever arm movement via a salt-bridge formed with Glu467 in the N-terminal region of the relay-helix. In addition, Asp590 does not have any interactions in single myosin (1VOM$_{md}$, 1MMD$_{md}$ and 1Q5G$_{md}$).

However, we found some changes in bond formations of the "prestrut" region after

actin binding (see Table 18). In 1Q5G$_{md}$, Glu586 has interactions with Arg418 that belong to the CM-loop of the upper 50 kDa domain, which suggest that the actin-binding cleft closure is probably related with bonds formation between the "prestrut" region and the CM-loop. Another salt-bridge interaction is observed between Glu586 of the "prestrut" region and Lys413 of the CM-loop in two partially closed cleft structures (3A-1VOM$_{md}$ and 3A-1MMD$_{md}$). The salt-bridge interaction between Lys587 of the "prestrut" region and Glu467 of the N-terminal regions of the relay-helix displays the same binding pattern as that is found in actin detached myosin. Asn588 and Lys589 have same interaction patterns with Asp530 and Ser533 of the helix-loop-helix region between 3A-1VOM$_{md}$ and 3A-1MMD$_{md}$. but Lys589 only interacts with Glu412 of the CM-loop in 3A-1Q5G$_{md}$. In three actomyosin complexes (3A-1VOM$_{md}$, 3A-1MMD$_{md}$ and 3A-1Q5G$_{md}$), there is a strong salt-bridge formed between Asp590 of the "prestrut" region and residue Lys265. Lys265 located in a loop, which has close backbone interactions with the SW-I loop (see Figure 35). This implies that this Asp590-Lys265 salt-bridge is probably important in regulating open-closed transition of the actin-binding cleft via the SW-I loop. This model might be confirmed by recent experimental findings as the deletion mutation of Asp590 in the "prestrut" region abolished the *in vivo* motor functions of myosin and showed defects in the strong actin-binding state [162, 164].

Table 18: Interactions of the "prestrut" region in different myosin structures

	"Prestrut" region				
	GLU586	**LYS587**	**ASN588**	**LYS589**	**ASP590**
1VOM$_{md}$	LYS589	GLU467	ASP530/SER533	LYS589	-
1MMD$_{md}$	-	GLU467	ASP530/SER533	-	-
1Q5G$_{md}$	-	GLU467	ASP530/SER533	SER533	-
3A-1VOM$_{md}$	LYS413	GLU467	ASP530/SER533	SER533/PHE535	LYS265
3A-1MMD$_{md}$	LYS413	GLU467	ASP530/SER533	SER533/PHE535	LYS265
3A-1Q5G$_{md}$	ARG418	GLU467/SER465	ASP530/SER533	GLU412	LYS265

To summarize the results in this section, the "prestrut" region has distinct interaction patterns in the absence or presence of actin. It looks like a claw interacting with the SW-II loop, the helix-loop-helix region in the absence of actin. However, upon actin

binding, the interactive areas of the "prestrut" region are further increased to the CM-loop and the SW-I loop in our three actomyosin complexes. It suggests an important functional role of the "prestrut" region in transducing the SW-II loop further closure, actin-binding cleft partial closure and the converter domain further rotation in the first phase of the power stroke.
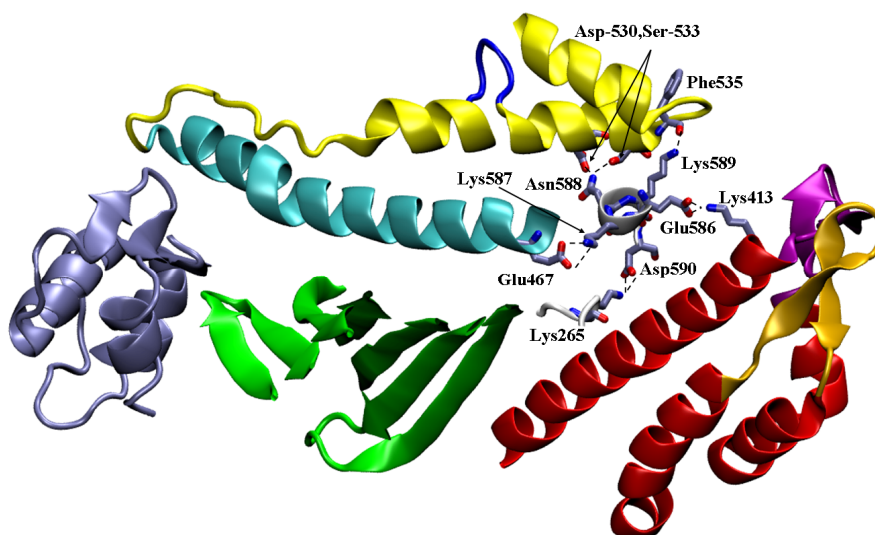
Figure 35: The "prestrut" region of myosin in the weak actin-binding state. These color-codes for functional domains of myosin were represented by the drawing method of cartoon. (Green: central β-sheet; Orange: loop 4; Purple: CM-Loop; Cyan: relay-helix; Yellow: lower 50kDa domain; Red: upper 50 kDa domain; Light ice blue: converter domain; Blue: activation loop)

5.3.2 Discussion

The effective power stroke is initiated by actin weakly binding to myosin in low actin binding-affinity state (open actin-binding cleft) [63]. The selection of the pathways to the effective power stroke is regulated by a kinetic pathway selection mechanism, which has been proved by several decades of investigations. However, structural mechanism of actin-binding induced myosin conformational changes and the potential mechanism of structure-based transition pathway of the power stroke have not been uncovered [165].

In this thesis, we describe a novel actin-binding site (Arg520 in the *Dictyostelium*

myosin) located in the activation loop of myosin that is responsible for transduction of actin-binding induced conformational changes towards myosin functional regions. As a highly conserved residue in myosin families, Arg520 interacts with the four negatively charged residues at the N-terminal region of actin in both weak and strong actin-binding states. The negatively charged N-terminus of actin is necessary for the force generation of actomyosin [148, 149]. In addition, Lys623 of loop 2 also interacts with this actin region in weak actin-binding states (3A-1VOM$_{md}$ and 3A-1MMD$_{md}$), but this interaction is disrupted in strong actin-binding states (3A-1Q5G$_{md}$) because of Lys623 forms a strong salt-bridge bond with Asp530 in the helix-loop-helix. Altered interaction profiles of Lys623 are able to assist the activation loop moving towards the C-terminus of actin in the power stroke. This effect might be confirmed by recent experiments in which loop 2 was shown to participate in actin activation [147]. Actin-binding induced three specific myosin conformational changes were observed in the weak actin-binding state. When R520Q mutation was introduced to myosin in order to prevent salt bridge formation with actin, three specific myosin conformational changes i.e. the partial closure of the actin-binding cleft, the further up rotation of the lever arm and further closure of SW-II loop were prevented or weakened. Motional correlations between the activation loop and other four regions decreased to the same level as that of in the absence of actin.

Two residues (Ile517 and Leu526) of the activation loop formed hydrophobic interactions with Leu469 of the myosin relay-helix in the actomyosin complexes. This bond formation causes a small local conformational change that is amplified in the relay-helix of myosin and resulting in a significant lever movement [166]. In addition, four residues (Ile529, Leu543 in the lower 50 kDa domain, Trp584 is closed to the "wedge" loop and Phe466 in N-terminus of the relay-helix) formed another hydrophobic cluster and mediate the correlation movements between the activation loop and the SW-II loop. The Lys462-Glu580 salt-bridge is substituted by His465-Glu580 in myosin VI as the recent X-ray crystal structure revealed [167]. Similar hydrophobic cluster is existed in myosin VI except that Leu543 in myosin II is replaced by phe543, which even contributes stronger effects to this hydrophobic

cluster than what we found in the myosin II complexed with actin. The recent myosin VI structure also represents the further lever up state of myosin. This experimental data validates our computational structural model of the actomyosin pre-power stroke state which shows the same further-up lever state of myosin.

Based on motional correlations, five functional regions of myosin (the activation loop, loop 2, the N-terminal region of the relay-helix, the "prestrut" region and the SW-II loop) can be divided into two groups in the weak actin-binding state. The first group consists of the N-terminal region of the relay-helix, the "prestrut" region and the SW-II loop. The N-terminal region of the relay-helix plays a central role in this group including Glu467 formed a salt-bridge with Lys587 of the "prestrut" region, and Asn475 interacted with Ser456 of the SW-II loop, which interaction was revealed in the recovery step by recent *in siliso* simulations [39]. The second group consists of the activation loop and loop 2, which are negotiated by the N-terminus of actin (see Figure 36).
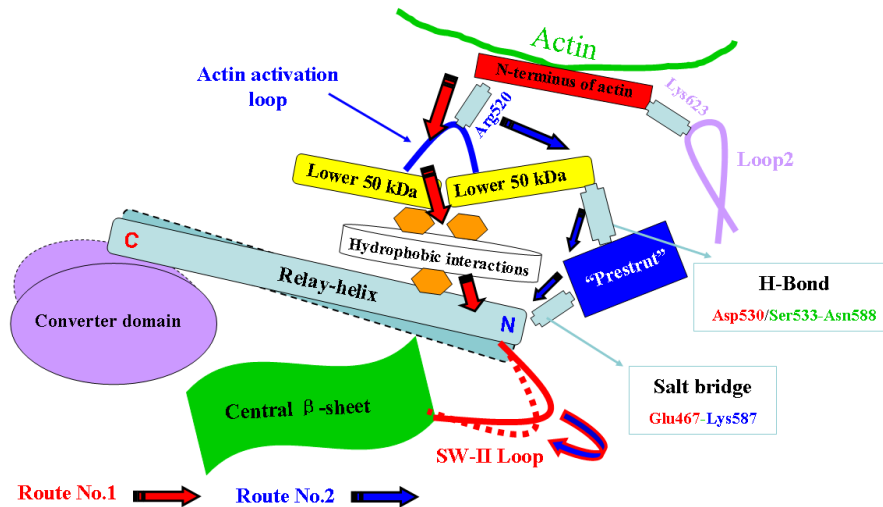
Figure 36: Two efficient routes between the activation loop and the SW-II loop. The first route starts from the activation loop to the SW-II loop through a hydrophobic cluster between the activation loop and the N-terminal of relay-helix. The second route starts from the apex of the helix-loop-helix to the "prestrut" region based on the bonds of Asn588-Asp530/Ser533, and then from the "prestrut" region to the SW-II loop based on the bond of Lys587-Glu467.

When actin binds to myosin, salt-bridge bonds are formed between Arg520 of the activation loop and four negatively charged residues in the N-terminus of actin also jointed with Lys623 of loop 2. Then, interactions start to push the activation loop

close to the relay-helix through a hydrophobic cluster and also close to the "prestrut" region via bonds formed between Asn588 of the "prestrut" region and Asp530/Ser533 of the helix-loop-helix region. Two deduced routes existed in these two groups are possible mechanism of actin-binding induced three specific myosin conformational changes prior to the power stroke. The first route starts from the activation loop to the SW-II loop through a hydrophobic cluster between the activation loop and the N-terminal of relay-helix. The second route starts from the apex of the helix-loop-helix to the "prestrut" region based on the bonds of Asn588-Asp530/Ser533, and then from the "prestrut" region to the SW-II loop based on the bond of Lys587-Glu467. Both routes are responsible for consequences of the SW-II loop further closure and converter domain further rotation. In addition, the second route is also possibly involved in the actin-binding cleft further closure according to the interaction between the SW-I loop and the "prestrut" region. In the pre-power stoke state of the actomyosin complex, these two communication pathway may play a crucial role in channeling the reaction pathway into the effective power stroke.

CONCLUSIONS:

According to raised scientific questions focused on the recovery step and acin-induced structural rearrangements of myosin pre-power stroke state stated in the chapter 1, we conclude our results to answer those questions.

First: Strain along the relay-helix of myosin is rearranged by eliminating the pivoting point in the seesaw-like motions at the beginning stage of the recovery step.

Second: A three-phase model of the recovery step deduced on the basis of a free energy profile with the post-recovery state in a lower free energy is more preferred than the pre-recovery state.

Third: In the three-phase model of the recovery step, the formation of the hydrogen bond cluster (γ-phosphate~Gly457~Asn475) accelerates structure transformation to overcome the activation energy barrier in the first phase.

Fourth: Our intermediate state 1 (IM-1) has similar conformation to the unique intermediate state of Fischer's model, but the IM-1 state with 37% converter domain rotation is in a middle structural state of the recovery step was not demonstrated before.

Fifth: Phe487/Phe503 hydrophobic cluster closely packing with some indispensable structural rearrangements in our second phase of the recovery step was not regarded as separated phases in Fischer's model. We could distinguish this phase based on the free energy profile of the recovery step.

Sixth: The weak and strong actin-binding states of actomyosin were obtained by protein-protein docking and long time-range molecular dynamics relaxations. The lower binding free energy of the strong actin-binding state is due to the involvement of more residues in the binding surface than that of in the weak actin-binding state.

Seventh: The conserved positive tip (Arg520) of the activation loop interacts with four negatively charged residues in the N-terminus of actin in various binding patterns of weak and strong actin-binding state.

Eighth: Three specific myosin conformational changes induced by the actin binding were observed in the weak actin-binding state: 1. the partial closure of the

actin-binding cleft, 2. the further up rotation of the lever arm and 3. further closure of SW-II loop. The R520Q mutation in myosin prevented these conformational changes.

Ninth: The motions of the activation loop are correlated with four mainly functional regions of myosin (loop 2, the N-terminus of the relay-helix, the SW-II loop and the "prestrut" region) in the weak actin-binding state. The motional correlations of R520Q mutation between the activation loop and other four regions in the weak actin-binding state decrease to the same level as that of in the absence of actin.

Tenth: Two communicational pathways were speculated between the actin binding regions and the myosin nucleotide binding site, which might be a reasonable mechanism of the actin-induced myosin conformational changes at the initial stage of the power stroke.

LIST OF FIGURES

LIST OF TABLES

APPENDIX

Table 2: Programs and web servers useful in comparative modeling

Name	Type[a]	WWW address
Databases		
GenBank	S	http://www.ncbi.nlm.nih.gov/Cenbank/
MODBASE	S	http://www.salilab.org/modbase/
PDB	S	http://www.rcsb.org/pdb/
TrEMBL	S	http://srs.ebi.ac.uk
Template search		
BLAST	S	http://www.ncbi.nlm.nih.gov/BLAST/
FastA	S	http://www.ebi.ac.uk/fasta33/
MATCHMAKER	P	http://bioinformatics.burnham-inst.org
Sequence alignment		
BLAST2	S	http://www.ncbi.nlm.nih.gov/gorf/bl2.html
CLUSTAL	S	http://www2.ebi.ac.uk/clustalw/
FASTA3	S	http://www2.ebi.ac.uk/fasta3/
Modeling		
COMPOSER	P	http://www.Tripos.com
ICM	P	http://www.molsoft.com
Insightll	P	http://www.accelrys.com
MODELLER	P	http://www.salilab.org/modeller/
QUANTA	P	http://www.accelrys.com
SYBYL	P	http://www.Tripos.com
WHAT IF	P	http://www.cmbi.kun.nl/whatif/

[a].S, server, P, program.

ABBREVIATIONS

ADP	adenosine 5' diphosphate
AIR	ambiguous interaction restraints
Arg, R	arginine amino acid
ATP	adenosine 5' triphosphate
BSA	buried surface area
CM-loop	cardiomyopathy loop
EM-map	electron microscopy density map
EMA	essential motional analysis
Gln, Q	glutamine amino acid
Glu, E	glutamic amino acid
IM-1	intermediate state 1
IM-2	intermediate state 2
Lys, K	lysine amino acid
MD	molecular dynamic
Mg	magnesium
MM/PBSA	molecular mechanic poisson-boltzmann surface area
NEB	nudged elastic band
Pi	inorganic phosphate group
PMF	potentials of mean force
Pro, P	proline amino acid
RMSD	root-mean-square deviation
RMSIP	root-mean-square inner product
RMSF	root-mean-square fluctuation
S1	myosin subfragment-1
SW-I	switch-1 loop
SW-II	switch-2 loop
WHAM	weighted histogram analysis method

REFERENCES

[1] Bustamante C, Chemla YR, Forde NR and Izhaky D. 2004. Mechanical processes in biochemistry. Annu. Rev. Biochem., 73: 705-748.

[2] George SB. 2000. Preface. Biochim. Biophys. Acta., 1496:1-2.

[3] Geeves MA and Holmes KC. 2005. The molecular mechanism of muscle contraction. Advances in Protein Chemistry, 71:161-193.

[4] Cope MJ, Whisstock J, Rayment I and Kendrick-Jones J. 1996. Conservation within the myosin motor domain: implications for structure and function. Structure, 4:969-987.

[5] Rayment I, Rypniewski WR, Schmidt-Base K, Smith R, Tomchick DR, Benning MM, Winkelmann DA, Wesenberg G and Holden HM. 1993. Three dimensional structure of myosin subfragment-1: a molecular motor. Science, 261: 50-58.

[6] Dominguez R, Freyzon Y, Trybus KM and Cohen C. 1998. Crystal structure of a vertebrate smooth muscle myosin motor domain and its complex with the essential light chain: visualization of the pre-power stroke state. Cell, 94: 559-571.

[7] Houdusse A, Kalabokis VN, Himmel D, Szent-Györgyi AG and Cohen C. 1999. Atomic structure of scallop subfragmentn S1 complexed with MgADP: a novel conformation of the myosin head. Cell, 97: 459-470.

[8] Vale RD and Milligan RA. 2000. The way things move: looking under the hood of molecular motor proteins. Science, 288: 88-95.

[9] Heintzelman MB and Schwartzman JD. 1997. A novel class of unconventional myosins from Toxoplasma gondii. J. Mol. Biol., 271:139-146.

[10] Cheney RE and Mooseker MS. 1992. Unconventional myosins. Curr. Opin. Cell Biol., 4:27-35.

[11] Sellers JR. 2000. Myosins: A diverse superfamily. Biochimica. Biophysica. Acta., 1496:3-22.

[12] Odronitz F, Kollmar M. 2007. Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species. Genome. Biol., 8(9):R196.

[13] Wilbur WJ and Lipman DJ. 1983. Rapid Similarity Searches of Nucleic Acid and Protein Data Banks. Proc Natl Acad Sci U.S.A., 80(3):726-730.

[14] Fujita-Becker S, Dürrwang U, Erent M, Clark RJ, Geeves MA and Manstein DJ. 2005. Changes in Mg2+-ion concentration and heavy chain phosphorylation regulate the motor activity of a class-I myosin. J. Biol. Chem., 280(7):6064-6071.

[15] Jontes JD and Milligan RA. 1997. Brush border myosin-I structure and ADP-dependent conformational changes revealed by cryelectron microscopy and image analysis. J. Cell. Biol., 39:683-693.

[16] Bähler M. 2000. Are class III and class IX myosins motorized signalling molecules? Biochim. Biophys. Acta., 1496:52-59.

[17] Schaub MC, Watterson JG, Waser PG. 1975. Conformational differences in myosin, IV. [1-3] Radioactive labeling of specific thiol groups as influenced by ligand binding. Hoppe. Seylers. Z. Physiol. Chem., 356(3):325-339.

[18] Reck-Peterson SL, Provance DW Jr, Mooseker MS and Mercer JA. 2000. Class V myosins. Biochim. Biophys. Acta., 1496: 36-51.

[19] Tang F, Kauffman EJ, Novak JL, Nau JJ, Catlett NL and Weisman LS. 2003. Regulated degradation of a class V myosin receptor directs movement of the yeast vacuole. Nature, 422:87-92.

[20] Altman D, Sweeney HL and Spudich JA. 2004. The mechanism of myosin VI translocation and its load-induced anchoring. Cell, 116:737-749.

[21] Bähler M. Are class III and class IX myosins motorized signalling molecules? 2000. Biochim. Biophys. Acta., 1496(1):52-59.

[22] Hettmann C, Herm A, Geiter A, Frank B, Schwarz E, Soldati T, Soldati D. 2000.A dibasic motif in the tail of a class XIV apicomplexan myosin is an essential determinant of plasma membrane localization. Mol. Biol. Cell., 11(4):1385-400.

[23] Holmes KC and Geeves MA. 2000. The Structural Basis of Muscle Contraction. Philosophical Transactions: Biological Sciences, 355(1396):419-431.

[24] Sweeney HL, Houdusse A. 2010. Structural and functional insights into the Myosin motor mechanism. Annu. Rev. Biophys., 39:539-557.

[25] Smith CA and Rayment I. 1996. X-ray structure of the magnesium (II).ADP.vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9 Resolution. Biochemistry, 35:5404-5417.

[26] Mocz G, Szilagyi L, Chen Lu R, Fabian F, Balint M, Gergely J. 1984. Effect of nucleotides, divalent cations and temperature on the tryptic susceptibility of myosin subfragment 1. Eur. J. Biochem., 145(2):221-229.

[27] Choi IG and Kim SH. 2006. Evolution of protein structural classes and protein sequence families. Proc Natl Acad Sci U.S.A., 103:14056-14061.

[28] McLaughlin PJ, Gooch JT, Mannherz HG, Weeds AG. 1993. Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. Nature, 364(6439):685-92.

[29] Schutt CE, Myslik JC, Rozycki MD, Goonesekere NC and Lindberg U. 1993. The structure of crystalline profilin-beta-actin. Nature, 365:810-816.

[30] Kabsch W, Mannherz HG, Suck D, Pai EF and Holmes KC. 1990. Atomic structure of the actin: DNase I complex. Nature, 347:37-44.

[31] Lorenz M, Popp D, Holmes KC. 1993. Refinement of the F-actin model against X-ray fiber diffraction data by the use of a directed mutation algorithm. J. Mol. Biol., 234(3):826-836.

[32] Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T and Warren GL. 1998. Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. Acta. Crystallogr. D., 54:905-921.

[33] Schröder RR, Manstein DJ, Jahn W, Holden H, Rayment I, Holmes KC and Spudich JA. 1993. Three-dimensional atomic model of F-actin decorated with Dictyostelium myosin S1. Nature, 364:171-174.

[34] Murakami K, Yasunaga T, Noguchi TQ, Gomibuchi Y, Ngo KX, Uyeda TQ, Wakabayashi T. 2010. Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release. Cell, 143(2):275-287.

[35] Oda T, Iwasa M, Aihara T, Maéda Y and Narita A. 2009. The nature of the globular-to fibrous-actin transition. Nature, 457(7228):441-445.

[36] Coureux PD, Sweeney HL and Houdusse A. 2004. Three Myosin V structures delineate essential features of chemo-mechanical transduction. EMBO Journal, 23(23):4527-4537.

[37] Reubold TF, Eschenburg S, Becker A, Kull FJ, and Manstein DJ. 2003. A structural model for actin-induced nucleotide release in myosin. Nature Structural Biology, 10(10):826-830.

[38] Li G and Cui Q. 2004. Mechanochemical Coupling in Myosin: A Theoretical Analysis with Molecular Dynamics and Combined QM/MM Reaction Path Calculations. Journal of Physical Chemistry B, 108:3342-3357.

[39] Fischer S, Windshugel B, Horak D, Holmes KC, and Smith JC. 2005. Structural mechanism of the recovery step in the Myosin molecular motor. Proc Natl Acad Sci U.S.A., 102(19):6873-6878.

[40] Koppole S, Smith JC, and Fischer S. 2006. Simulations of the myosin II motor reveal a nucleotide-state sensing element that controls the recovery step. Journal of Molecular Biology, 361:604-616.

[41] Cooke R. The mechanism of muscle contraction. 1986. CRC. Crit. Rev. Biochem., 21(1):53-118.

[42] Geeves MA, Fedorov R and Manstein DJ. 2005. Molecular mechanism of actomyosin-based motility. Cell. Mol. Life. Sci., 62:1462-1477.

[43] Holmes KC, Angert I, Kull FJ, Jahn W and Schroder RR. 2003. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. Nature, 425:423-427.

[44] Coureux PD, Wells AL, Ménétrey J, Yengo CM, Morris CA, Sweeney HL and Houdusse A. 2003. A structural state of the myosin V motor without bound nucleotide. Nature, 425:419-423.

[45] Holmes KC, Schroder RR, Sweeney HL and Houdusse A. 2004. The structure of the rigor complex and its implications for the power stroke. Philos. Trans. R. Soc. London. Ser. B., 359:1819-1828.

[46] Trabuco LG, Villa E, Mitra K, Frank J and Schulten K. 2008. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure, 16:673-683.

[47] Milligan RA. 1996. Protein-protein interactions in the rigor actomyosin complex. Proc Natl Acad Sci U.S.A., 93:21-26.

[48] Mornet D, Bertrand R, Pantel P, Audemard E and Kassab R. 1981. Structure of the actin-myosin interface. Nature, 292:301-306.

[49] Liu YM, Scolari M, Im W and Woo HJ. 2006. Protein-protein interactions in actin-myosin binding and structural effects of R405Q mutation: A molecular dynamics study. Proteins, 64:156-166.

[50] Roopnarine O and Leinwand LA. 1998. Functional analysis of myosin mutations that causefamilial hypertrophic cardiomyopathy. Biophys. J., 75:3023-3030.

[51] Lorenz M and Holmes KC. 2010. The actin-myosin interface. Proc Natl Acad Sci

U.S.A., 107(28):12529-12534.

[52] Onishi H, Mikhailenko SV, Morales MF. 2006. Toward understanding actin activation of myosin ATPase: the role of myosin surface loops. Proc Natl Acad Sci U.S.A., 103(16):6136-6141.

[53] Kojima S, Konishi K, Katoh K, Fujiwara K, Martinez HM, Morales MF, Onishi H. 2001. Functional roles of ionic and hydrophobic surface loops in smooth muscle myosin: their interactions with actin. Biochemistry, 40(3):657-664.

[54] Onishi H, Morales MF. 2007. A closer look at energy transduction in muscle. Proc Natl Acad Sci U.S.A., 104(31):12714-12719.

[55] Várkuti BH, Yang Z, Kintses B, Erdélyi P, Bárdos-Nagy I, Kovács AL, Hári P, Kellermayer M, Vellai T, Málnási-Csizmadia A. 2012. A novel actin binding site of myosin required for effective muscle contraction. Nat Struct Mol Biol, 19(3):299-306.

[56] Kintses B, Gyimesi M, Pearson DS, Geeves MA, Zeng W, Bagshaw CR, Málnási-Csizmadia A. 2007. Reversible movement of switch 1 loop of myosin determines actin interaction. EMBO J., 26(1):265-274.

[57] Conibear PB, Bagshaw CR, Fajer PG, Kovács M and Málnási-Csizmadia A. 2003. Myosin cleft movement and its coupling to actomyosin dissociation. Nat. Struct. Biol., 10(10):831-835.

[58] Lawson JD, Pate E, Rayment I and Yount RG. 2004. Molecular dynamics analysis of structural factors influencing back door pi release in myosin. Biophys. J., 86(6):3794-3803.

[59] Kühner S, Fischer S. 2011. Structural mechanism of the ATP-induced dissociation of rigor myosin from actin. Proc Natl Acad Sci U.S.A., 108(19): 7793-7798.

[60] Nitao LK, Loo RR, O'Neall-Hennessey E, Loo JA, Szent-Györgyi AG and Reisler E. 2003. Conformation and dynamics of the SH1-SH2 helix in scallop myosin. Biochemistry, 42(25):7663-7674.

[61] Burghardt TP, Garamszegi SP, Park S and Ajtai K. 1998. Tertiary structural changes in the cleft containing the ATP sensitive tryptophan and reactive thiol are consistent with pivoting of the myosin heavy chain at Gly699. Biochemistry, 37(22):8035-8047.

[62] Batra R, Geeves MA, Manstein DJ. 1999. Kinetic analysis of Dictyostelium discoideum myosin motor domains with glycine-to-alanine mutations in the reactive thiol region. Biochemistry, 38(19):6126-6134.

[63] Gyimesi M, Kintses B, Bodor A, Perczel A, Fischer S, Bagshaw CR and Málnási-Csizmadia A. 2008. The mechanism of the reverse recovery step, phosphate release, and actin activation of Dictyostelium myosin II. J. Biol. Chem., 283:8153-8163.

[64] Málnási-Csizmadia A, Kovács M. 2010. Emerging complex pathways of the actomyosin power stroke. Trends Biochem Sci, 35(12):684-690.

[65] Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F and Sali A. 2000. Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct., 29:291-325.

[66] Pieper U, Eswar N, Stuart AC, Ilyin VA and Sali A. 2002. MODBASE, a database of annotated comparative protein structure models. Nucleic. Acids. Res., 30(1):255-259.

[67] Baker D and Sali A. 2001. Protein structure prediction and structural genomics. Science, 294(5540):93-96.

[68] Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U and Sali A. 2006. Comparative protein structure modeling using Modeller. Curr Protoc Bioinformatics, Chapter 5:Unit 5.6.

[69] Sali A and Blundell TL. 1993. Comparative protein modeling by satisfaction of spatial restraints. J Mol Biol, 234(3):779-815.

[70] Fiser A, Do RK and Sali A. 2000. Modeling of loops in protein structures. Protein Sci, 9(9):1753-1773.

[71] Dayhoff M O, Schwartz R M and Orcutt BC. 1978. In Atlas of Protein Sequence and Structure. Nat. Biomed. Res. Found., 5(suppl. 3): 345-352,

[72] Henikoff S and Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci.U.S.A., 89:10915-10919.

[73] Saqi MA, Russell RB and Sternberg MJ. 1998. Misleading local sequence alignments: Implications for comparative protein modeling. Protein Eng., 11:627-630.

[74] Browne WJ, North ACT, Phillips DC, Brew K, Vanaman TC and Hill RC. 1969. A possible three-dimensional structure of bovine lactalbumin based on that of hen's egg-white lysosyme. J. Mol. Biol., 42:65-86.

[75] Bystroff C and Baker D. 1998. Prediction of local structure in proteins using a library of sequence-structure motifs. J. Mol. Biol., 281:565-577.

[76] Camacho CJ and Vajda S. 2002. Protein-protein association kinetics and protein docking. Curr. Opin. Struct. Biol., 12(1):36-40.

[77] Smith GR and Sternberg MJ. 2002. Prediction of protein-protein interactions by docking methods. Curr. Opin. Struct. Biol., 12(1):28-35.

[78] Lengauer T, Rarey M. 1996. Computational methods for biomolecular docking. Curr. Opin. Struct. Biol., 6 (3): 402-406.

[79] Levinthal C, Wodak SJ, Kahn P, Dadivanian AK. 1975. Hemoglobin Interactions in Sickle Cell Fibers: I. Theoretical Approaches to the Molecular Contacts. Proc Natl Acad Sci U.S.A., 72 (4): 1330-1334.

[80] Meng EC, Gschwend DA, Blaney JM, Kuntz ID. 1993. Orientational sampling and rigid-body minimization in molecular docking.Orientational sampling and rigid-body minimization in molecular docking. Proteins, 17(3):266-278.

[81] Kramer B, Rarey M, Lengauer T. 1997. CASP2 experiences with docking flexible ligands using FlexX. Proteins, S1:221-225.

[82] Jain AN. 2006. Scoring functions for protein-ligand docking. Curr. Protein Pept. Sci., 7(5): 407-420.

[83] Lensink MF, Méndez R, Wodak SJ. 2007. Docking and scoring protein complexes: CAPRI 3rd Edition. Proteins Structure Function and Bioinformatics, 69(4): 704-718.

[84] Dominguez C, Boelens R and Bonvin AM. 2003. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc., 125(7):1731-1737.

[85] Vries SJ, van Dijk AD, Krzeminski M, van Dijk M, Thureau A, Hsu V, Wassenaar T, and Bonvin AM. 2007. HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. Proteins, 69(4):726-733.

[86] Nilges M. 1993. A calculation strategy for the structure determination of symmetric dimers by 1H NMR. Proteins, 17(3):297-309.

[87] Hubbard SJ and Thornton JM. 1993. 'NACCESS', computer program. Department Biochemistry and Molecular Biology, University College, London.

[88] Balbuena and Seminario. 1999. Molecular Dynamics, 1st Edition From Classical to Quantum Methods. Print Book Elsevier Science.

[89] Allen MP, Tildesley DJ. 1989. Computer simulation of liquids. Clarendon Press.

[90] Markwick PR, McCammon JA. 2011. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. Phys Chem Chem Phys, 13(45):20053-20065.

[91] McCammon JA, Gelin BR and Karplus M. 1977. Dynamics of folded proteins. Nature, 267(5612): 585-590.

[92] McQuarrie D. 2000. Statistical Mechanics. University Science Books, 2nd edition.

[93] Perrot P. 1998. A to Z of Thermodynamics. Oxford University Press.

[94] Nash LK. 1974. Elements of statistical thermodynamics. Dover Publications, Inc, 2nd edition.

[95] Ulrich Nienhaus. 2005. Protein-Ligand Interactions : Methods and Applications. Methods in Molecular Biology, 305: 215-241.

[96] Virnau P, Muller M. 2004. Calculation of free energy through successive umbrella sampling. J Chem Phys., 120(23):10925-10930.

[97] Jerzy L. 1999. Computational Chemistry: Reviews of Current Trends. Computational Chemistry: Reviews of Current Trends. Vol8.

[98] Chadi. 1978. Energy-minimization approach to the atomic geometry of semiconductor surfaces. Physical. Review. Letters., 41(15):1062-1065.

[99] Davis ME and McCammon JA. 1989. Solving the finite difference linearized Poisson-Boltzmann equation: A comparison of relaxation and conjugate gradient methods. J. Comp. Chem., 10:386-391.

[100] Atich. 1989. Conjugate gradient minimization of the energy functional: A new method for electronic structure calculation. Physical. Review. B., 39(8):4997-5004.

[101] Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, and Kollman PA. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput. Phys. Commun., 91:1-41.

[102] Brooks B, Bruccoleri R, Olafson B, States D, Swaminathan S and Karplus M. 1983. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. J. Comput. Chem., 4:187-217.

[103] Bekker H, Berendsen HJC, Dijkstra EJ, Achterop S, van Drunen R, van Der Spoel D, Sijbers A, Keegstra H, Reitsma B and Renardus MKR. 1993. GROMACS: A parallel computer for molecular dynamics simulations. Proc of the 4 Int conf Physics Computing., 92:252-256.

[104] Caspersen KJ and Carter EA. 2005. Finding transition states for crystalline solid-solid phase transformations. Proc Natl Acad Sci U.S.A., 102(19):6738-6743.

[105] Elber R and Karplus M. 1987. A method for determining reaction paths in large molecules: Application to myoglobin. Chem. Phys. Lett., 139:375-380.

[106] Virnau P, Muller M. 2004. Calculation of free energy through successive umbrella sampling. J. Chem. Phys., 120(23):10925-10930.

[107] Kumar S, Rosenberg JM, Bouzida D, Swendsen RH and Kollman PA. 1992. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J. Comput. Chem., 13:1011-1021.

[108] Luo R, David L and Gilson MK. 2002. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. J. Comput. Chem., 23:1244-1253.

[109] Fisher AJ, Smith CA, Thoden JB, Smith R, Sutoh K, Holden HM and Rayment I. 1995. X-ray structures of the myosin motor domain of Dictyostelium discoideum complexed with MgADP.BeFx and MgADP.AlF4 .Biochemistry, 34:8960-8972.

[110] Smith CA and Rayment I. 1996. X-ray structure of the magnesium (II).ADP.vanadate complex of the Dictyostelium discoideum myosin motor domain to 1.9 Resolution. Biochemistry, 35:5404-5417.

[111] Meagher KL, Redman LT and Carlson HA. 2003. Development of polyphosphate parameters for use with the AMBER force field. Journal of Computational Chemistry, 24(9):1016-1025.

[112] Yang Y, Gourinath S, Kovács M, Nyitray L, Reutzel R, Himmel DM, O'Neall-Hennessey E, Reshetnikova L, Szent-Györgyi AG, Brown JH and Cohen C. 2007. Rigor-like structures from muscle myosins reveal key mechanical elements in the transduction pathways of this allosteric motor. Structure, 15(5):553-564.

[113] Yu H, Ma L, Yang Y and Cui Q. 2007. Mechanochemical coupling in the myosin motor domain. II. Analysis of critical residues. PLoS. Comput. Biol., 3(2):e23

[114] Yu H, Ma L, Yang Y and Cui Q. 2007. Mechanochemical coupling in the myosin motor domain. I. Insights from equilibrium active-site simulations. PLoS. Comput. Biol., 3(2):e21.

[115] Frisch MJ. 2004. Gaussian 03, Revision C.02 (Gaussian, Inc., Wallingford, CT).

[116] Wang J, Wolf RM, Caldwell JW, Kollman PA and Case DA. 2004. Development and testing of a general amber force field. J Comput Chem, 25(9):1157-1174.

[117] Chang PL, Rinne AW, Dewey TG. 2006. Structure alignment based on coding of local geometric measures. BMC Bioinformatics, 7:346.

[118] Mizuguchi K, Deane CM, Blundell TL, Johnson MS, Overington JP. 1998. JOY: protein sequence-structure representation and analysis. Bioinformatics, (7):617-623.

[119] Abyzov A, Errami M, Leslin CM, Ilyin VA. 2005. Friend, an integrated analytical front-end application for bioinformatics. Bioinformatics, 18:3677-3678.

[120] Chaussepied P and Morales MF. 1988. Modifying preselected sites on proteins: the stretch of residues 633-642 of the myosin heavy chain is part of the actin-binding site. Proc Natl Acad Sci U.S.A., 85(20):7471-7475.

[121] Bertrand R, Derancourt J and Kassab R. 1997. Probing the hydrophobic interactions in the skeletal actomyosin subfragment 1 and its nucleotide complexes by zero-length cross-linking with a nickel-peptide chelate. Biochemistry, 36(32):9703-9714.

[122] Bertrand R, Derancourt J and Kassab R. 1994. The covalent maleimidobenzoyl-actin-myosin head complex. Cross-linking of the 50 kDa heavy chain region to actin subdomain-2. FEBS Lett., 345(2-3):113-119.

[123] Jorgensen WL, and Tirado-Rives J. 1988. The OPLS potential functions for proteins. Energy minimization for crystals of cyclic peptides and crambin, J. Am. Chem. Soc., 110:1657-1666.

[124] Jorgensen WL, Chandrasekhar J, Madura J and Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys., 79:926-935.

[125] Darden TA, York DM and Pedersen LG. 1993. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: A comparison of the Ewald and truncated list methods. J. Chem. Phys., 99:8345-8348.

[126] Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A and Haak JR. 1984. Molecular dynamics with coupling to an external bath. J. Chem. Phys., 81:3684-3690.

[127] Jean PR, Giovanni C and Berendsen HJC. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Chem. Phys., 23:327-341.

[128] Humphrey W, Dalke A and Schulten K. 1996. VMD: visual molecular dynamics. J Mol Graph., (1):33-38.

[129] Li L, Uversky VN, Dunker AK and Meroueh SO. 2007. A computational investigation of allostery in the catabolite activator protein. J. Am. Chem. Soc., 129(50):15668-15676.

[130] Leo-Macias A, Lopez-Romero P, Lupyan D, Zerbino D and Ortiz AR. 2005. An analysis of core deformations in protein superfamilies. Biophys. J., 88(2):1291-1299.

[131] Wan S, Coveney PV, Flower DR. 2005. Peptide recognition by the T cell receptor: comparison of binding free energies from thermodynamic integration, Poisson-Boltzmann and linear interaction energy approximations. Philos. Transact. A. Math. Phys. Eng. Sci., 363(1833):2037-2053.

[132] Connolly ML. 1983. Analytical molecular surface calculation. J. Appl. Cryst., 16:548-558.

[133] Kottalam J and Case DA. 1990. Langevin modes of macromolecules: applications to crambin and DNA hexamers. Biopolymers, 29(10-11):1409-1421.

[134] Conibear PB, Málnási-Csizmadia A and Bagshaw CR. 2004. The effect of F-actin on the relay helix position of myosin II, as revealed by tryptophan fluorescence, and its implications for mechanochemical coupling. Biochemistry, 43:15404-15417.

[135] Koppole S, Smith JC, Fischer S. 2007. The structural coupling between ATPase activation and recovery step in the myosin II motor. Structure, 15(7):825-837.

[136] Mesentean S, Koppole S, Smith JC, Fischer S. 2007. The principal motions involved in the coupling mechanism of the recovery step of the myosin motor. J. Mol. Biol. 367(2):591-602.

[137] Kintses B, Yang Z, Málnási-Csizmadia A. 2008. Experimental investigation of the seesaw mechanism of the relay region that moves the myosin lever arm. J. Biol. Chem., 283(49):34121-34128.

[138] Málnási-Csizmadia A, Pearson DS, Kovács M, Woolley RJ, Geeves MA, Bagshaw CR. 2001. Kinetic resolution of a conformational transition and the ATP hydrolysis step using relaxation methods with a Dictyostelium myosin II mutant containing a single tryptophan residue. Biochemistry, 40(42):12727-12737.

[139] Bauer CB, Holden HM, Thoden JB, Smith R, Rayment I. 2000. X-ray structures

of the apo and MgATP-bound states of Dictyostelium discoideum myosin motor domain. J. Biol. Chem., 275(49): 38494-38499.

[140] Caremani M, Dantzig J, Goldman YE, Lombardi V, Linari M. 2008. Effect of inorganic phosphate on the force and number of myosin cross-bridges during the isometric contraction of permeabilized muscle fibers from rabbit psoas. Biophys. J., 95(12):5798-5808.

[141] Dantzig JA, Goldman YE, Millar NC, Lacktis J, Homsher E. 1992. Reversal of the cross-bridge force-generating transition by photogeneration of phosphate in rabbit psoas muscle fibres. J. Physiol., 451:247-278.

[142] Takács B, Billington N, Gyimesi M, Kintses B, Málnási-Csizmadia A, Knight PJ and Kovács M. 2010. Myosin complexed with ADP and blebbistatin reversibly adopts a conformation resembling the start point of the working stroke. Proc Natl Acad Sci U.S.A., 107:6799-6804.

[143] Ferenczi MA, Bershitsky SY, Koubassova N, Siththanandan V, Helsby WI, Panine P, Roessle M, Narayanan T and Tsaturyan AK. 2005. The ''roll and lock'' mechanism of force generation in muscle. Structure, 13:131-141.

[144] Root DD. 2002. A computational comparison of the atomic models of the actomyosin interface. Cell. Biochem. Biophys., 37(2):97-110.

[145] Gyimesi M, Tsaturyan AK, Kellermayer MS, Málnási-Csizmadia A. 2008. Kinetic characterization of the function of myosin loop 4 in the actin-myosin interaction. Biochemistry, 47(1):283-291.

[146] Furch M, Remmel B, Geeves MA, Manstein DJ. 2000. Stabilization of the actomyosin complex by negative charges on myosin. Biochemistry, 39(38):11602-11608.

[147] Joel PB, Trybus KM, Sweeney HL. 2001. Two conserved lysines at the 50/20-kDa junction of myosin are necessary for triggering actin activation. J. Biol. Chem., 276(5):2998-3003.

[148] Miller CJ, Wong WW, Bobkova E, Rubenstein PA, Reisler E. 1996. Mutational analysis of the role of the N terminus of actin in actomyosin interactions. Comparison with other mutant actins and implications for the cross-bridge cycle. Biochemistry, 35(51):16557-16565.

[149] Lu X, Bryant MK, Bryan KE, Rubenstein PA, Kawai M. 2005. Role of the N-terminal negative charges of actin in force generation and cross-bridge kinetics in reconstituted bovine cardiac muscle fibres. J Physiol, 564(Pt 1):65-82.

[150] Málnási-Csizmadia A, Kovács M, Woolley RJ, Botchway SW, Bagshaw CR.

2001. The dynamics of the relay loop tryptophan residue in the Dictyostelium myosin motor domain and the origin of spectroscopic signals. J. Biol. Chem., 276: 19483-19490.

[151] Takács B, O'Neall-Hennessey E, Hetényi C, Kardos J, Szent-Györgyi AG, Kovács M. 2011. Myosin cleft closure determines the energetics of the actomyosin interaction. FASEB J., 25(1):111-121.

[152] Khaitlina SY, Strzelecka-Gołaszewska H. 2002. Role of the DNase-I-binding loop in dynamic properties of actin filament. Biophys. J., 82:321-334.

[153] Sasaki N, Asukagawa H, Yasuda R, Hiratsuka T, Sutoh K. 1999. Deletion of the myopathy loop of Dictyostelium myosin II and its impact on motor functions. J. Biol. Chem., 274(53):37840-37844.

[154] Razzaq A, Schmitz S, Veigel C, Molloy JE, Geeves MA, Sparrow JC. 1999. Actin residue glu(93) is identified as an amino acid affecting myosin binding. J. Biol. Chem., 274(40):28321-28328.

[155] Johara M, Toyoshima YY, Ishijima A, Kojima H, Yanagida T, Sutoh K. 1993. Charge-reversion mutagenesis of Dictyostelium actin to map the surface recognized by myosin during ATP-driven sliding motion. Proc Natl Acad Sci U.S.A. 90(6):2127-2131.

[156] Miller CJ, Reisler E. 1995. Role of charged amino acid pairs in subdomain-1 of actin in interactions with myosin. Biochemistry, 34(8):2694-2700.

[157] Van Dijk J, Furch M, Lafont C, Manstein DJ, Chaussepied P. 1999. Functional characterization of the secondary actin binding site of myosin II. Biochemistry, 38(46):15078-15085.

[158] Geeves MA, Conibear PB. 1995. The role of three-state docking of myosin S1 with actin in force generation. Biophys. J., 68:194S-199S.

[159] Rayment I, Holden HM, Whittaker M, Yohn CB, Lorenz M, Holmes KC, Milligan RA. 1993. Structure of the actin-myosin complex and its implications for muscle contraction. Science, 261(5117):58-65.

[160] Fajer MI, Li H, Yang W, Fajer PG. 2007. Mapping electron paramagnetic resonance spin label conformations by the simulated scaling method. J. Am. Chem. Soc., 129:13840-13846.

[161] Cecchini M, Alexeev Y, Karplus M. 2010. Pi release from myosin: a simulation analysis of possible pathways. Structure, 18(4):458-470.

[162] Sasaki N, Ohkura R, Sutoh K. 2000. Insertion or deletion of a single residue in the strut sequence of Dictyostelium myosin II abolishes strong binding to actin. J. Biol. Chem., 275(49):38705-38709.

[163] Giese KC, Spudich JA. 1997. Phenotypically selected mutations in myosin's actin binding domain demonstrate intermolecular contacts important for motor function. Biochemistry, 36(28):8465-8473.

[164] Fujita-Becker S, Reubold TF, Holmes KC. 2006. The actin-binding cleft: functional characterisation of myosin II with a strut mutation. J Muscle Res Cell Motil. 27(2):115-123.

[165] Biro, N.A. and A.E. Szent-Gyorgyi. 1949. The effect of actin and physico-chemical changes on the myosin ATP-ase system, and on washed muscle. Hung. Acta. Physiol., 2(1-4): 120-133.

[166] Tehver R, Thirumalai D. 2101. Rigor to post-rigor transition in myosin V: link between the dynamics and the supporting architecture. Structure, 18(4):471-481.

[167] Wells AL, Lin AW, Chen LQ, Safer D, Cain SM, Hasson T, Carragher BO, Milligan RA, Sweeney HL. 1999. Myosin VI is an actin-based motor that moves backwards. Nature, 401(6752):505-508.