PHYSICAL STUDIES OF ACTIN NUCLEATION AND CONFORMATIONAL DYNAMICS

APPROVED BY SUPERVISORY COMMITTEE

Michael Rosen

Diana Tomchick

Luke Rice

Hongtao Yu

DEDICATION

Both in high school and as an undergraduate, my academic performance vacillated unpredictably between mediocre and horrendous. The reasons for this are complicated, and I'm not sure I completely understand them, but in any case, despite the turmoil, I maintained a penchant for science and a certitude that I would pursue a scientific career. After graduating college with a meager 2.1 GPA however, it became apparent that my love for science, while certainly a necessary condition for pursuing further training, was, in the eyes of most graduate programs, wholly insufficient (I can't say I blame them for this). That I am presently writing this thesis proves that I did ultimately find my way into graduate school and earn a PhD, but I could never have done so unless, along the way, I encountered mentors who gave me chances when I hardly deserved them. The first and perhaps most pivotal example of this occurred in 2004 when I met Duane Grandgenett, a professor of molecular virology at Saint Louis University. When I met Duane, I was delivering

pizzas to earn money, was quite tired of doing so, and wasn't sure how I was going to forge a path that could lead to the career I desired. Duane agreed to hire me as a technician in his lab doing work to understand the mechanisms underlying the function of retrovirus integrases. Duane's lab provided an environment that fostered creativity, positive collaborative interactions, and independent thinking. Most importantly though, Duane gave me chance to succeed and believed in me, even when I hardly believed in myself. For this I will be forever indebted, and I hope one day that I can do for someone else what he did for me. I spent four years in Duane's lab, and in that time, was able to publish several papers and begin to gain a level of experience that began to overshadow my poor academic performance. After leaving Duane's lab, I spent three years working for Niraj Tolia at Washington University in St. Louis studying the structural mechanisms through which *Plasmodium* adhesins interact with their cognate receptors on red blood cells. Here too, I was free to learn, pursue some of my own ideas, and publish papers. With these experiences in hand, I next applied to graduate school, and ultimately joined the program at UT southwestern. I am grateful to those on the admissions committee for the UT Southwestern Graduate School of Biomedical Sciences who, like my previous mentors, overlooked my poor academic performance, and instead focused on subsequent triumphs. My seven years as a PhD student in Michael Rosen's lab at UT Southwestern represent the pinnacle of my scientific career thus far. Mike, like all of the aforementioned mentors, gave me a chance to succeed in his lab, in spite

of my college transcripts, and for this I am grateful. The manner in which Mike thinks about and approaches scientific problems resonated with me upon first meeting him. In his lab, I not only mastered new techniques and explored interesting and fundamental scientific problems, but through beginning to understand how he thinks, was able to sharpen my own scientific intuition in a way that will benefit me for the rest of my life. So, I dedicate this thesis to all of the mentors I've encountered leading up to this point, particularly those that gave me the benefit of the doubt. Additionally, I'd like to thank my thesis committee (Diana Tomchick, Hongtao Yu, Luke Rice, and Michael Rosen), and the Biophysics Faculty and students for their guidance. Additionally (and perhaps most importantly), I would like to thank my parents, Scott and Carolyn, my sister Claire, and my wife, Nicky for their continued and unyielding love and support.

PHYSICAL STUDIES OF ACTIN NUCLEATION AND CONFORMATIONAL DYNAMICS

by

Jacob Aaron Zahm

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of Molecular Biophysics

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2017

PHYSICAL STUDIES OF ACTIN NUCLEATION AND

CONFORMATIONAL DYNAMICS

Publication No._____

Jacob Aaron Zahm, PhD

The University of Texas Southwestern Medical Center at Dallas, 2017

Supervising Professor: Michael K. Rosen, PhD

Actin is a 42 kilodalton ATPase that exists ubiquitously in eukaryotic cells. Unlike other ATPases, however, actin, under suitable conditions, can polymerize, forming helical filaments. Cells, in orchestrating their myriad cellular processes, utilize actin's intrinsic capacity to polymerize, but do so in a tightly controlled fashion, such that new filaments only appear when and where the cell needs them to suit specific purposes. Such control exists at two different levels. Firstly, the stability of actin filaments is subject to "intrinsic" control arising from the state of bound nucleotide.

ATP binding favors incorporation of actin monomers into filaments. This incorporation augments actin's ATP hydrolysis activity, and the conversion of ATP to ADP in the nucleotide binding cleft considerably destabilizes filaments, facilitating the return of filament subunits to free monomers. The structural mechanism through which nucleotide conveys information throughout the actin monomer to influence polymerization behavior remains poorly understood and represents a persistent fundamental biological question. In this work I, for the first time, apply modern muti-resonance NMR methods to begin to answer these questions. In addition to the aforementioned intrinsic control, cellular actin is subject to "extrinsic" control via the action of nucleation factors. In order to form a growing filament, actin must proceed through a nucleation step in which monomers must assemble into a thermodynamically and kinetically disfavored nucleus, which ultimately proceeds to a growing filament. Nucleation factors accelerate the rate of filament formation by binding to actin monomers and arranging them into the prerequisite nucleus. In this work, I reveal the crystal structure of actin monomers in complex with the bacterially derived nucleation factor, VopL. The structure represents the first high resolution snapshot of a filament-like nucleation intermediate, and reveals general principles underlying the action of nucleation factors.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote Pichia pastoris. Clark L, **Zahm JA**, Ali R, Kukula M, Bian L, Patrie SM, Gardner KH, Rosen MK, Rosenbaum DM. *J Biomol NMR. 2015 Jul;62(3):239-45.*

The bacterial effector VopL organizes actin into filament-like structures. **Zahm JA**, Padrick SB, Chen Z, Pak CW, Yunus AA, Henry L, Tomchick DR, Chen Z, Rosen MK. *Cell. 2013 Oct 10;155(2):423-34.*

Dimerization of Plasmodium vivax DBP is induced upon receptor binding and drives recognition of DARC. Batchelor JD, **Zahm JA**, Tolia NH. *Nat Struct Mol Biol. 2011 Jul 10;18(8):908-14.*

Mechanisms of human immunodeficiency virus type 1 concerted integration related to strand transfer inhibition and drug resistance. **Zahm JA**, Bera S, Pandey KK, Vora A, Stillmock K, Hazuda D, Grandgenett DP. *Antimicrob Agents Chemother. 2008 Sep;52(9):3358-68.*

Inhibition of human immunodeficiency virus type 1 concerted integration by strand transfer inhibitors which recognize a transient structural intermediate. Pandey KK, Bera S, **Zahm J**, Vora A, Stillmock K, Hazuda D, Grandgenett DP. *J Virol. 2007 Nov; 81(22):12189-99*

Size-dependent diffusion in cycloalkanes. Bruce. A. Kowert *, Jared B. Jones, **Jacob A. Zahm** & Robert M. Turner II. *Molecular Physics Vol. 102, Iss. 13, 2004*

Size-Dependent Diffusion in the n-Alkanes. Bruce A. Kowert,*, Kurtis T. Sobush, Chantel F. Fuqua, Courtney L. Mapes, Jared B. Jones, and, and **Jacob A. Zahm**. *The Journal of Physical Chemistry A 2003 107 (24), 4790-4795*

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Cellular milieus are rife with physical forces that would likely prove lethal to a cell in the absence of suitable protective mechanisms. To maintain their physical integrity under these assaults, cells possess a rigid polymer network called the “actin cytoskeleton” that is resilient to external forces. A conflict of interest arises however, because, in addition to requiring physical rigidity, cells frequently need to exhibit significant plasticity, responding to the prevailing conditions in their environment by either moving in a directed fashion, changing shape, or deforming their membranes. The actin cytoskeleton solves these problems simultaneously because, while it is rigid, it can, subject to appropriate signals, rapidly and intricately rearrange in a manner that suits a specific purpose (Pollard 2003).

Actin is a 42 kilodalton ATPase that exists primarily and ubiquitously in eukaryotic cells, and is among the most conserved proteins in nature; actins from *Homo sapiens* and from the roundworm *Caenorhabditis elegans* share 94% sequence identity despite being separated by approximately 790 million years of evolution (Kumar S 2017). This level of conservation suggests that actin’s function is central to the maintenance of eukaryotic life.

Actin’s core biochemical function is to shuttle back-and-forth between a free, monomeric, globular (G) state, and a polymerized, filamentous (F) state, the latter of which consists of actin monomers arranged into a two-stranded helical filament**.** The

stability of actin filaments is controlled by the state of bound nucleotide. Actin binds to a single molecule of ATP, and in the unpolymerized G-state, nucleotide hydrolysis is negligible. However, once incorporated into a filament, actin rapidly hydrolyzes ATP to ADP, and this conversion significantly decreases the stability of the filament, facilitating the return from the F state to the monomeric G state. At the end of a growing filament, recently-incorporated subunits haven't yet hydrolyzed nucleotide, are bound to ATP, and are thus stably incorporated into the filament. The older subunits, having been incorporated sufficiently long to hydrolyze ATP, are bound to ADP, and are less stably incorporated. This causes filaments to disassemble on one end while assembling on the other, such that the filament recycles monomers, but exhibits no net growth. This phenomenon is called "treadmilling," and is important for actin's cellular function. The structural mechanism describing how the information about nucleotide state propagates throughout the actin molecule and ultimately controls filament stability represents a long-standing question in the actin field. In chapter one, following this introduction, I review the current literature relevant to nucleotide-driven allostery in the actin monomer.

Brunö Ferenc Straub purified actin for the first time in 1942 and discovered its role in muscle contractility (Szent-Gyorgyi 2004). In this context, polymerized actin composes an integral part of the sarcomere, wherein myosin motors pull on bundles of constant-length actin filaments to bring about muscle contraction. Nearly thirty years later scientists began to observe actin in non-muscle cells, revealing that actin's

role extends far beyond the sarcomere, into the orchestration of many cellular processes, including cell motility, endo and exocytosis, axon guidance, embryonic development, and cancer metastases, just to name a small subset.

In contrast to its apparently static role in muscle contraction the aforementioned processes require that actin transition rapidly between monomeric and filamentous states, such that new filaments form when and where they're needed, and that existing filaments return to monomers (it turns out that, even in the sarcomere, actin filaments are dynamic, despite seeming rigid and inactive (Ono 2010)),. A particularity illustrative example of such a process is found within the lamellipodium of a migrating cell. Here, directly adjacent to the cell membrane, one observes a highly branched, rapidly advancing network of actin filaments that pushes the cell membrane in the direction of motion. The older filaments in the network, which recede to the back of the advancing mesh, rapidly disassemble so that the monomers can be recycled and incorporated into new filaments directly adjacent to the moving membrane. During migration, the width of the branched actin network remains constant, somewhere between 2 and 5 μM, and as it advances against the membrane, it is subject to myosin-mediated retrograde flow which causes the entire mesh to move, as a unit, away from the cell membrane. Thus, the rate of extension at the leading edge reflects the balance between filament growth and retrograde flow.

The orchestration of actin filament dynamics in lamellipodia, and in other processes involving rapid actin rearrangements, require that the cell precisely control

when and where new actin filaments form (and where they are disassembled, although I will focus on the former here). The cell's ability to control actin polymerization relies on the fact that, in the absence of assistance, initiation of new filaments occurs slowly, relative to the rapidity with which new filaments form in cellular processes. This sluggishness arises from an intrinsic barrier that lies between the actin monomer and the actin filament. To generate a new filament, several actin monomers must arrange themselves into a kinetically and thermodynamically disfavored nucleus. The accepted pathway to this nucleus, arrived at by calculating the energetics of individual intersubunit contacts in the filament, involves two unfavorable steps to yield an actin trimer that resembles an actin filament. Once formed, the addition of subsequent monomers is energetically favorable, leading to filament growth. Cells use this barrier as a point of control, and possess an arsenal of so-called "nucleation factors" that bind to actin monomers and stabilize them in the arrangement of the requisite nucleus, and in doing so, initiate the growth of new actin filaments. The cell uses an arsenal of different nucleation factors in concert to precisely control the morphology and dynamics of the actin cytoskeleton, which, in turn, controls myriad cellular processes. How nucleation factors bind to and organize actin monomers into filament nuclei, and how these interactions specify the dynamic behavior of a given nucleation factor represent persistent questions in the actin field.

Given that cytoskeletal dynamics are a master-regulator of cellular behavior, it is unsurprising that bacterial pathogens have evolved their own set of actin regulatory

proteins that serve to alter the morphology and dynamics of the actin cytoskeleton of eukaryotic hosts in a manner that benefits their survival. “Bacterial effectors” are proteins that pathogenic bacteria inject into their eukaryotic hosts in order to alter cellular funtion in a manner that benefits survival of the pathogen. These effector proteins are functionally variable and interfere with many cellular processes, but a subset of them exert their influence on the actin cytoskeleton. Within this subset, some effectors act upon signaling pathways upstream of filament formation, while others are bona fide nucleation factors, and bring about polymerization by interacting with actin directly. Frequently these effector proteins mimic the functions of endogenous proteins, but they tend to exert effects that are extreme relative to the host counterparts; a bacterial effector that acts as a nucleation factor is likely to exhibit exaggerated potency relative to endogenous nucleation factors, for example. This feature of the effectors makes them good tools for understanding cell biology.

The work in this thesis advances our understanding of actin biophysics at two different levels. First, I develop the necessary tools and begin to address the long-standing question of how actin’s nucleotide state controls its polymerization behavior. Second, I reveal the first high-resolution structure of an actin nucleus in complex with a pathogen-derived nucleation factor. Together these studies contribute significantly to our understanding of how actin polymerization is controlled intrinsically by nucleotide state, and extrinsically, by actin nucleation factors.

CHAPTER ONE

INTRODUCTION TO ACTIN WITH A FOCUS ON ALLOSTERIC COMMUNICATION

Actin polymer dynamics

Because the subunits that compose an actin filament are asymmetric, the filament too is asymmetric, with each of the two ends being distinct from one another, one a fast-growing "barbed" end, the other a slow-growing "pointed-end." Furthermore, actin subunits change conformation upon incorporation into a filament, which implies that a free monomer can exhibit affinity for the filament end that is either higher or lower than for that of other free monomers. In the case of actin, free monomers exhibit higher affinity for filament ends than for other free monomers, leading to a scenario where elongation of existing filaments is favored over the formation (nucleation) of new filaments. Together, these features define a "polar equilibrium polymer," implying that the two ends of the filament can exhibit different growth properties. For this class of polymer, the addition of a subunit to either end of the filament ultimately results in the formation of the same interactions, but the kinetic pathways that lead to the same endpoint can differ. Thus, subunits will associate and dissociate from the respective ends at different rates, but the ratio of the rate constants for the respective ends remains conserved; i.e., the respective ends have the same "critical concentration." The critical concentration is the concentration of free

monomer in solution when the rate of subunit addition is equal to dissociation, and thus there is no net growth. When the total actin concentration exceeds the critical concentration, filament will grow until the free monomer concentration again equals the critical concentration. Actin however, because it hydrolyzes ATP in the context of the filament, is not a “polar equilibrium polymer,” but a “complex non-equilibrium polymer.” Because actin exhibits fast-growth at the barbed end and slow growth at the pointed end, during filament growth there is a “cap” of ATP-bound subunits at the barbed end, while the pointed end subunits are bound to ADP. In this scenario, the incorporation of subunits into the respective ends of the filament leads to subtly different subunit-subunit interactions, which implies that they can, in principle, exhibit different growth rates and critical concentrations, and they indeed do. Since the ends have different critical concentrations, one can imagine an actin filament bathing in a solution of ATP-bound free monomers which are present at a concentration that exceeds the barbed end critical concentration but falls below the pointed end critical concentration. This will cause the filament to assemble at its barbed end and disassemble at its pointed end, and if the system is properly tuned, the filament will recycle monomers, consume ATP, but undergo no net growth. This process, called “treadmilling,” is critically important for actins cellular function. It is through this process that lamellipodia, discussed in the introduction, are able to maintain a constant width as they drive the membrane forward.

The question of how the state of bound nucleotide controlls the stability of subunits within an actin filament represents a long-standing question in the actin field. Existing crystal structures of ATP and ADP bound actin are nearly identical to one another, and thus give little insight into this fundamental question. There is however, substantial biochemical evidence that implicate regions of the actin monomer that appear to be coupled to bound nucleotide. In this chapter I will review existing literature describing how actins nucleotide state is coupled allosterically to it's different regions.

The Actin Monomer

Actin is a 376-amino acid ATPase that falls in a superfamily with hexokinases and the nucleotide binding domain of the chaperone Hsp70. Like its relatives, actin binds to and hydrolyzes ATP. However, actin is the only member of the superfamily that exhibits polymerization behavior. This tendency to polymerize stymied early efforts to crystallize actin, but ultimately, Ken Holmes' lab solved the structure of an actin monomer bound to ADP in 1990 (Kabsch, Mannherz et al. 1990). In that study, the authors co-crystallized actin with DNaseI, which binds tightly to actin and prevents polymerization. This ended up being a common theme in subsequent actin crystallography; there are now over 50 actin structures in the PDB, and nearly all of them rely on methods to prevent filament formation. Historically this was achieved via co-crystallization with actin-binding proteins or small molecules that prevent

polymerization, or through chemical modification of residues involved in inter-subunit filament contacts. In 2004, Kathleen Trybus' group described a robust method to express actin recombinantly in insect cells (Joel, Fagnant et al. 2004), and furthermore, described a non-polymerizable mutant that her group subsequently crystallized in both ATP and ADP states (Rould, Wan et al. 2006).

A surprising feature of the existing actin structures in the PDB is that, in nearly all of them, actin is in essentially the same conformation, irrespective of its nucleotide state or what it is bound to (Dominguez and Holmes 2011). This stands in sharp contrast to other related proteins, such as the Hsp70, which undergoes dramatic conformational changes upon binding to other proteins or upon hydrolysis of bound nucleotide (Zhuravleva, Clerico et al. 2012). In light of this structural consistency, it is unclear how the actin molecule conveys information about nucleotide state in the centrally-located nucleotide binding pocket to distant parts of the molecule involved in inter-subunit filament contacts.

All of the available structures reveal that actin is a relatively flat molecule (55Å X 55Å X 35Å) comprising both alpha helical and beta sheet structural elements. There are four distinct subdomains, with 1 and 3 being related to one another, likely having arisen from a gene duplication, and subdomains 2 and 4 being insertions in 1 and 3, respectively (Dominguez and Holmes 2011) (Fig. 1-1). In all but one crystal structure, subdomain 2 is disordered, and in the case in which it was, it turned out that the ordering was likely enforced by crystal packing interactions (Rould, Wan et al. 2006).

The nucleotide binds in the center of the molecule, in a cleft formed by subdomains 1 and 3, and makes additional contacts with subdomains 2 and 4 (Fig. 1-1). In all actin structures the, the β and γ phosphates of ATP (or just the β phosphate in ADP structures) coordinate a divalent metal ion, usually Mg^{2+} or Ca^{2+}. Owing to its comparatively higher concentration in the cell, Mg^{2+} is thought to be the physiologically relevant cation, although there are circumstances involving release of calcium into the cytoplasm; the significance of this for actin dynamics and function is not known. However, as will become relevant later, the relative affinities of actin for ATP and for ADP are reversed in the presence of calcium, such that that in the presence of calcium, actin exhibits an approximately 1000-fold preference for ATP, while in the presence of magnesium, and approximately four-fold preference for ADP. This metal-driven affinity change could certainly profoundly affect the properties of the cytoskeleton. The ATP γ-phosphate directly contacts three loops, known as P1, P2, and the "sensor loop," all of which reside in subdomains 1 and 2 (Fig. 1). In comparing the crystal structures of ATP and ADP-bound actin, the only differences

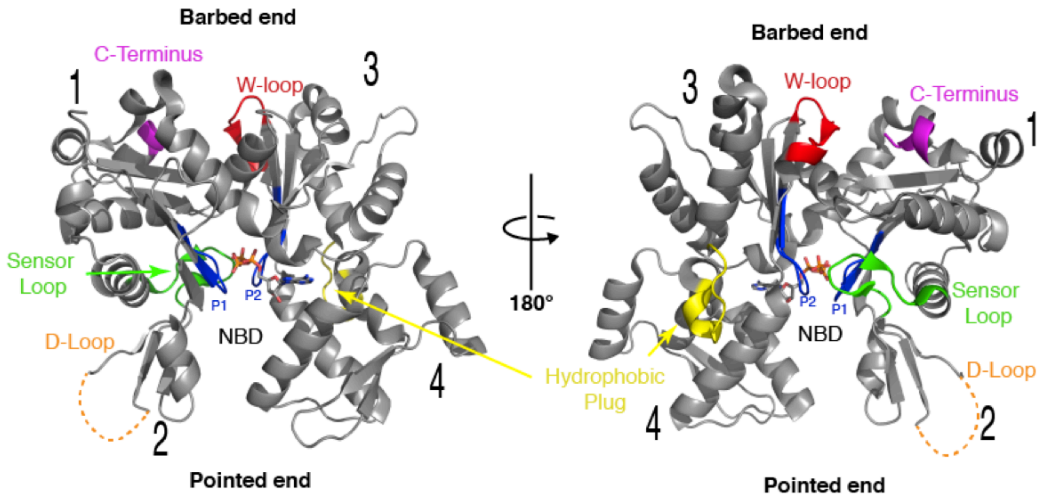


Fig. 1-1 Cartoon representation of the actin monomer. The crystal structure of the actin monomer is shown in grey (PDB code 2HF4). The identities of the four subdomains are indicated with numbers. The barbed and pointed ends are indicated. Structural elements thought to be important for filament formation appear in different colors. The nucleotide binding domain is abbreviated as “NBD”

lie in these loops (overall RMSD=0.296, sensor loop RMSD=0.372), and the small perturbations don't propagate throughout the molecule. Within the cleft formed by the junction of subdomains 1 and 3, lies an extensive hydrophobic groove that makes important contacts in the filament, and has been called an "allosteric hotspot" because it binds a multitude of regulatory factors that affect global conformation or affect ATP hydrolysis rate (Chereau, Kerff et al. 2005) (Fig. 1). For example, WH2 domains bind to the 1,3-cleft in a manner that discriminates between ATP and ADP nucleotide states, binding more tightly to the former, indicating that this region couples allosterically to the nucleotide binding cleft. Cofilin, another protein that binds to the 1,3-cleft, binds preferentially to ADP-bound actin. Contained within this "hotspot" is a small flexible loop called the "W-loop". There is evidence that the W-loop is a "nucleotide sensor," which conveys information about nucleotide state to the 1,3-cleft (Kudryashov, Grintsevich et al. 2010). As described below, there is considerable biochemical evidence for long-range communication among the W-loop, the nucleotide binding domain, and subdomain 2. Yet this allostery has yet to be described in terms of a structural mechanism.

Crystal Structures of Actin in the ATP and ADP States

Both Roberto Dominguez and Kathleen Trybus have published reports describing uncomplexed actin (crystallized in the absence of actin-binding proteins) bound to both ATP (Graceffa and Dominguez 2003, Rould, Wan et al. 2006) and ADP

(Otterbein, Graceffa et al. 2001, Rould, Wan et al. 2006), and overall these structures are all very similar to one another (Fig. 1-2) In the study from the Dominguez lab, the authors prevent polymerization by modifying Cys374 with tetramethylrhodamine, which sterically blocks the 1,3-cleft and prevents polymerization. In the Trybus study, they introduce mutations (A204E/P243K) into the pointed end, in subdomain 4, that interfere with filament formation. Also of note, in the Dominguez report, actin is bound to the non-hydrolyzable analogue AMP-PNP rather than ATP.

Comparison of the ATP and ADP-bound Dominguez structures revealed a striking difference; In the ADP state, The D-loop of subdomain 2 is ordered, forming a small helix, in contrast to their ATP-bound structure in which the D-loop remained disordered (Fig 1-2). This data was consistent with earlier biochemical data that showed that the D-loop of ADP actin is resistant to digestion by subtilisin (Schwyter, Phillips et al. 1989), consistent with the formation of a protease-resistant secondary structure. Taken together, this led to the hypothesis that nucleotide state is coupled allosterically to the conformation of the D-loop in subdomain 2. This was an attractive hypothesis, because of the critical role of the D-loop in forming the long-pitch interaction in the filament. Controlling D-loop conformation via modulation of nucleotide state had the hallmarks of an elegant mechanism. Also in corroboration of their proposed mechanism, the ADP-bound and ATP-bound actin each crystallized in the same space group, with nearly identical unit cell dimensions, increasing

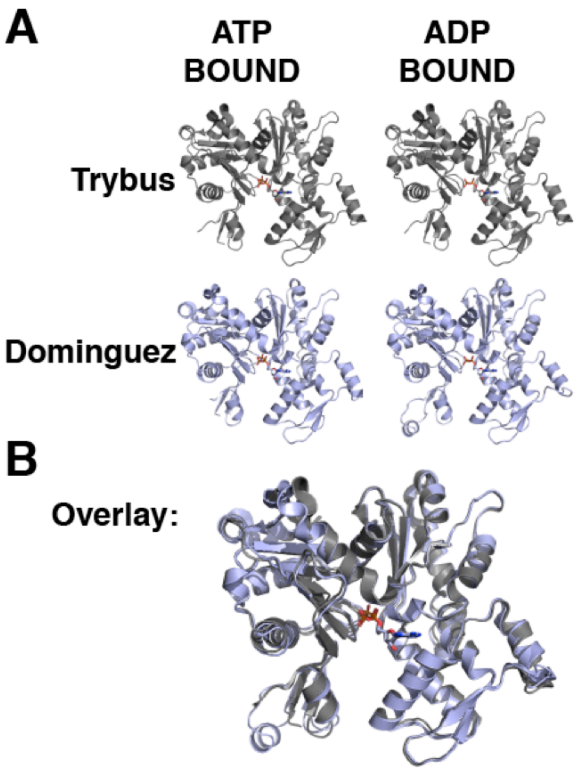


Fig. 1-2 Comparison of actin crystal structures in the ATP and ADP states.
A. Cartoon representation of the Trybus crystal structures of mutant "AP" actin bound to ATP and to ADP (PDB codes 2HF3 and 2HF4, respectively), compared to the Dominguez analogs, modified with tetramethylrhodamine (1NWK and 1J6Z). Note the ordering of the D-Loop in the ADP-bound Dominguez structure. **B.** Overlay of all four structures depicted in **A;** all of the structures are nearly identical to one another, with a maximum backbone RMSD of ~0.60.

confidence that the differences between the two structures were not the result of crystallographic artifacts. That being said, the crystallization conditions differed from one another, which can also lead to functionally spurious features that exist in one structure but not another.

Later, Trybus' group crystallized the "AP" non-polymerizable mutant actin in both ATP and ADP states. In contrast to the Dominguez studies, the D-loop remained disordered in both structures (Fig. 1-2). These crystals were not only of the same space group, but each grew in precisely the same condition, which further bolsters the credibility of interpretations involving differences between the structures. It is now the consensus in the field that the D-loop ordering in the Dominguez ADP-bound structure was likely an artifact arising from crystal packing. It does however remain a distinct possibility that the behavior of subdomain 2 is affected by nucleotide state.

Despite this controversy, there were aspects of the Trybus and Dominguez reports about which the authors agreed, and these features, while considerably subtler, are more likely authentic. Both reports observe that, in the ADP-bound structure, S14, which resides in the P1 loop, interacts directly with the ADP β-phosphate. However, in the ATP state, the γ-phosphate interferes sterically causing S14 to rotate by 130° (Fig. 1-3). This causes S14 to impinge upon the sensor loop, changing its conformation modestly. Since the sensor loop connects subdomain 2 to the ATP terminal phosphate, and lies C-terminal to the D-loop by only 17

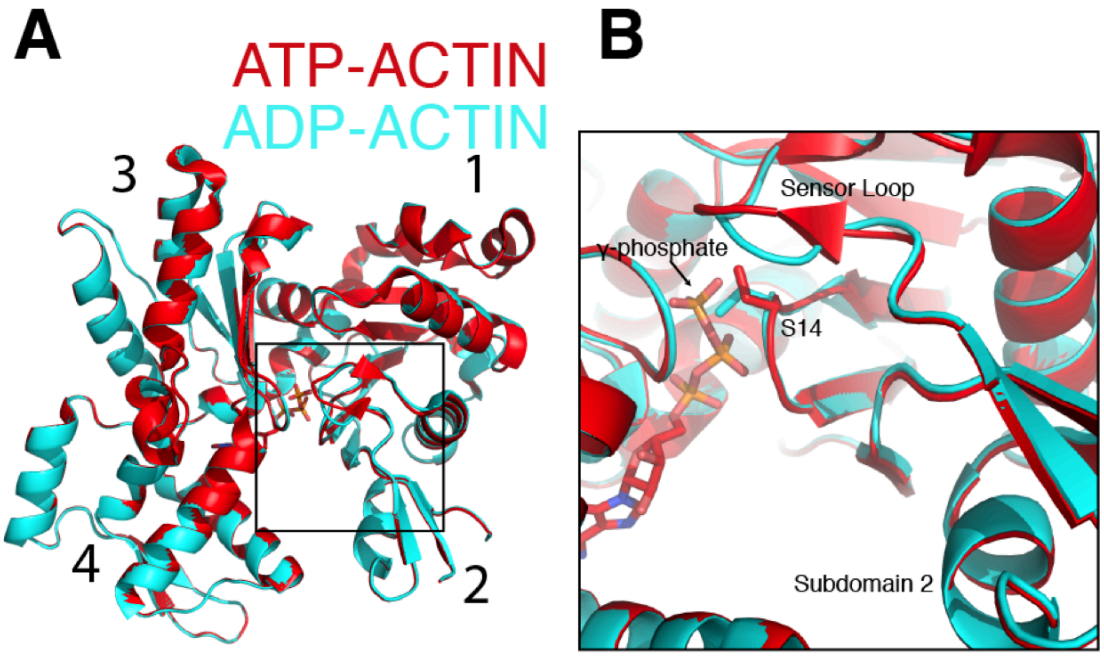


Fig. 1-3 Depiction of the major perturbation, evident in crystal structures, that occurs upon conversion of ATP to ADP in the nucleotide binding cleft.
A. ATP and ADP actin appear in cartoon representation in red and cyan, respectively. The subdomain identities are indicated. **B.** Close-up of the NBD approximately corresponding to the box in **A.** Note that, in ATP actin, S14 rotates substantially, resulting in relocation of the sensor loop, which lies both upstream and downstream from the subdomain 2 (which contains the D-loop).

residues, it is conceivable that this slight perturbation alters the behavior of the D-loop, ultimately changing the stability of filaments. As detailed in the next section, it has been shown biochemically that nucleotide state controls the behavior of the N and C-termini, the d-loop, and the W-loop in the 1,3-cleft, which in turn affect filament formation and association with actin binding proteins. However, the changes observed in the crystal structures are modest, and restricted to the vicinity of the nucleotide terminal phosphate, and thus don't explain this long-range allosteric communication.

Analogy to Hsp70 and Open and Closed States

Hsp70 is a molecular chaperone that cycles, in a nucleotide dependent manner, between states of high and low affinity for substrates (Zhuravleva, Clerico et al. 2012). In this way, Hsp70 transiently binds to folding intermediates and subsequently releases them. These cycles of catch-and-release presumably change the kinetics of protein folding in a way that prevents aggregation (Fig 1-4). Hsp70 comprises an actin-like nucleotide-binding domain tethered via a C-terminal hydrophobic linker to a substrate-binding domain. The nucleotide-binding domain, like actin, consists of four subdomains, with the nucleotide binding cleft between subdomains 1 and 3. Also like actin, Hsp70 works by cycling between ATP and ADP-bound forms. In the ADP state, the Hsp70 NBD adopts an "open" conformation

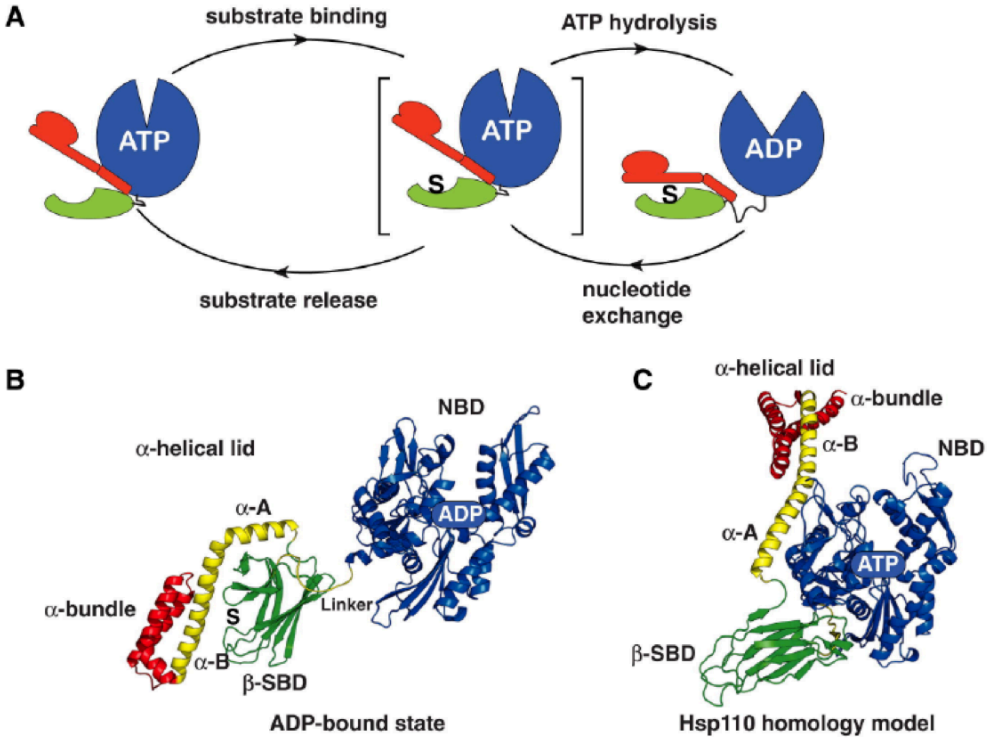


Fig. 1-4 Nucleotide-driven structural changes occurring in Hsp70.
A. Schematic representation of the "docked" and "undocked" states in Hsp70 corresponding to different nucleotide states. **B.** Detailed structural model describing the changes depicted in **A.** Notice that, when bound to ATP, the Hsp70 NBD adopts a "closed" state, with the subdomains that correspond to 2 and 4 in actin moving substantially closer together.

wherein subdomains 2 and 4 splay apart, and the SBD dissociates from the NBD, so that they exist essentially as isolated domains, tethered together through the flexible linker (Fig 1-4). In this state, the SBD binds tightly to substrate. Upon conversion to the ATP state, subdomains 2 and 4 rotate toward one another, tightening down around the nucleotide, the hydrophobic linker docks against the NBD in the 1,3-cleft (analogous to the D-loop engagement in an actin filament), and the SBD packs against the NBD and undergoes a dramatic conformational change (Fig 1-4). In this state, the affinity for substrate is reduced by up to 50-fold with rapid on-off kinetics.

The degree of structural and functional analogy between Hsp70 and actin is striking, and suggest that it is reasonable to learn lessons from one to better understand the other. The structures of actin and the Hsp70 NBD are similar to one another, and both use nucleotide hydrolysis to control the state of the 1,3-cleft which, in turn, controls the binding of peptides (the hydrophobic linker in the case of Hsp70, and the D-loop or other binding proteins in the case of actin).

As compared to actin, Hsp70 homologs diverge considerably in sequence. This variation permitted a rigorous analysis of how allosteric communication occurs in this molecule (Smock, Rivoire et al. 2010). This study uses covariance analysis of sequences of Hsp70 homologs to identify a “sector” of residues in the Hsp70 NBD important for nucleotide state-driven control of linker binding and SBD docking. Unsurprisingly these residues surround the nucleotide binding site and emanate to and line the 1,3-cleft that binds the hydrophobic linker. Lending credence to the

functional relevance of this analysis, many of the residues in this sector appeared in previous mutagenesis studies, where they were shown to be important for allostery (Smock, Rivoire et al. 2010). It is likely that one can look to this sector to identify residues in the actin molecule potentially involved in allosteric communication. It should be noted however, that the sector identified in Hsp70 lacked residues residing in subdomain 2, which, in the case of actin, is thought to be coupled to nucleotide state. This is not particularly surprising however, because the role of subdomain 2 in actin is to supply the D-loop peptide to form long-pitch inter-subunit interactions in the filament; Subdomain 2 in Hsp70 has no analogous function, and has thus likely become uncoupled from nucleotide state; accordingly, subdomain 2 in Hsp70 does not carry any sequence resembling actin's D-loop.

In its ATP and ADP-bound forms, the Hsp70 NBD adopts "closed" and "open" states, respectively (Zhuravleva, Clerico et al. 2012) (Fig 1-6). This is in contrast to actin, which generally crystallizes in a closed conformation, irrespective of nucleotide state. Crystal structures of actin bound to the nucleotide exchange factor profilin provide the only exception to this rule (Schutt, Myslik et al. 1993, Chik, Lindberg et al. 1996, Porta and Borgstahl 2012). Profilin binds tightly to actin's barbed end, making extensive contacts to subdomains 1 and 3, and in crystal structures of the profilin-actin complex, actin adopts an "open" state wherein subdomains 2 and 4 splay apart from one another, and the sensor, P1 and P2 loops move further from the nucleotide. This presumably brings about a decrease in affinity and increase in off-rate that facilitates

nucleotide exchange. There are also structures of the profilin-actin complex in which the actin adopts a closed conformation, identical to that observed in most other actin structures, which suggests that profilin-bound actin exists in equilibrium between open and closed states. Curiously, profilin only negligibly perturbs the subdomains to which it binds (1 and 3), but somehow brings about substantial reorientation of subdomains 2 and 4.

While the only crystallographic examples in which actin adopts an open state are the profilin-bound structures, there is an EM study that reveals actin in an open state within a filament (Galkin, Orlova et al. 2015). The presence of the open conformation in both the profilin-bound crystal structure and in the filament, taken together with the analogy to Hsp70, suggests that the open state exists in actin's intrinsic conformational landscape, and is likely populated to a degree controlled by nucleotide state. To test directly whether actin exists in an open-closed equilibrium in solution, Emil Reisler's group introduced cysteine residues on both sides of the nucleotide binding cleft and treated the mutants with cysteine-reactive crosslinkers of different lengths. If ADP-bound actin were to more frequently visit an open state in which the reactive cysteines are further apart, one would expect shorter crosslinking agents to work less efficiently. This is precisely what they observe, substantiating the idea that actin exists in solution in a nucleotide-dependent open-closed equilibrium (Kudryashov and Reisler 2013).

Structural elements in actin thought to couple to nucleotide state

Introduction

Actin's D-loop, C-terminus, and W-loop all exhibit varying degrees of disorder, make important inter-subunit contacts in the filament, and interact with actin-binding proteins. Furthermore, there is biochemical evidence that each of these motifs are allosterically coupled to nucleotide state, and are thus allosterically coupled to each other.

D-Loop

In the actin filament, the D-loop from one monomer interacts longitudinally with the 1,3-cleft of another monomer (Fig 2-1), forming the most substantial inter-subunit contact among those extant in a filament. Accordingly, its presence is required for filament formation. In addition to this stabilizing influence, the D-loop is subject to post-translational modifications that alter filament stability. Some cells possess a redox enzyme called MICAL that reversibly oxidizes a D-loop methionine leading to decreased filament stability (Hung, Pak et al. 2011). These modifications are important in orchestrating the cellular response to axon guidance cues during neuronal development. That the D-loop is subject to such regulation underscores its importance in controlling actin filament dynamics. The D-loop remains completely disordered in G-actin, and becomes partially ordered in F-actin, but in both cases,

interconverts among an ensemble of conformations (Galkin, Orlova et al. 2010) (Durer, Kudryashov et al. 2012). As previously discussed, ATP and ADP actin exhibit differential susceptibility to D-loop cleavage by subtilisin, suggesting that nucleotide state affects the conformation of the D-loop. Molecular dynamics simulations have led to conflicting results in this respect, some confirming the formation of the protease-resistant helix that Dominguez observed in the ADP-state (Zheng, Diraviyam et al. 2007) crystal structure, with others predicting that its existence is fleeting (Dalhaimer, Pollard et al. 2008, Splettstoesser, Noe et al. 2009). A recent study probed D-loop dynamics directly by chemically modifying a cysteine introduced into the D-loop with EPR-active nitroxide spin label (Durer, Kudryashov et al. 2012). These studies revealed that the conformational dynamics on the nanosecond timescale do not change as a function of nucleotide state. It is possible that nucleotide-state driven changes in D-loop dynamics exist in micro or millisecond regime, and thus would be unobservable by EPR. Molecular dynamics simulations suffer from the same problem, with simulations on longer timescales becoming more computationally intensive. Overall, there is evidence for allosteric coupling between the D-loop and the nucleotide-binding cleft, but it remains poorly understood.

C-terminus

Actin's C-terminus resides in subdomain 1, sits just above and adjacent to the 1,3-cleft, forms longitudinal contacts with the D-loop of an adjacent subunit in the filament,

and interacts with binding proteins that engage the 1,3-cleft (Figs 1-1 and 1-2). Analysis of B-factors across several actin crystal structures reveals that the C-terminus, like the D-loop, is quite disordered. There is also evidence that the C-terminus is involved in long-range allosteric coupling to the nucleotide binding cleft. The evidence for this arises from studies in which the authors generate actin variants lacking the last two or three residues at the C-terminus. This subtle perturbation resulted in a decrease in polymerization rate, impaired nucleotide hydrolysis, and altered the susceptibility of the D-loop to subtilisin (Mossakowska, Moraczewska et al. 1993, Strzelecka-Golaszewska, Mossakowska et al. 1995). This biochemical evidence, while suggesting allosteric connection between the C-terminus and the nucleotide-binding cleft and D loop, does not provide a quantitative mechanistic picture.

W-loop

The W-loop is part of subdomain 3 and lines the 1,3-cleft on the side opposite the C-terminus (Fig 1-1). Like the D-loop and the C-terminus, it exhibits some degree of disorder, as evidenced by high B-factors in crystal structures, and is involved in inter-subunit contacts within the filament and in interactions with binding proteins that engage the 1,3-cleft. In 1997, David Sept's group carried out molecular dynamics simulations which predicted that the W-loop adopts different conformations in the ATP, ADP-Pi, and ADP states (Zheng, Diraviyam et al. 2007). Following up on this

observation, Emil Reisler's group introduced cysteine mutations into the W-loop, labelled them with fluorescent probes, and asked whether the state of bound nucleotide affects the fluorescence properties of the attached fluorophores (Kudryashov, Grintsevich et al. 2010). They observed changes in the maximum fluorescence intensities upon switching the bound nucleotide from ATP to ADP, corroborating the predictions of the molecular dynamics study. Further evidence for allosteric coupling between the nucleotide and the W-loop is that actin binding proteins that engage the 1,3-cleft bind in a manner that discriminates between the two nucleotide states. For example, WH2 domains bind more tightly to ATP actin (Chereau, Kerff et al. 2005), while cofillin bind more tightly to ADP actin (Maciver and Weeds 1994). So, like the D-loop and the C-terminus, and the W-loop, appears to be coupled to the state of nucleotide in the nucleotide-binding cleft, but the nature of this coupling remains obscure.

Structure of the actin filament

In 1990 Holmes' group published an ~7 Å, near-atomic resolution fiber-diffraction structure of the actin filament (Holmes, Popp et al. 1990). They used the DNaseI-bound G-actin crystal structure to fit the fiber diffraction data, but the low

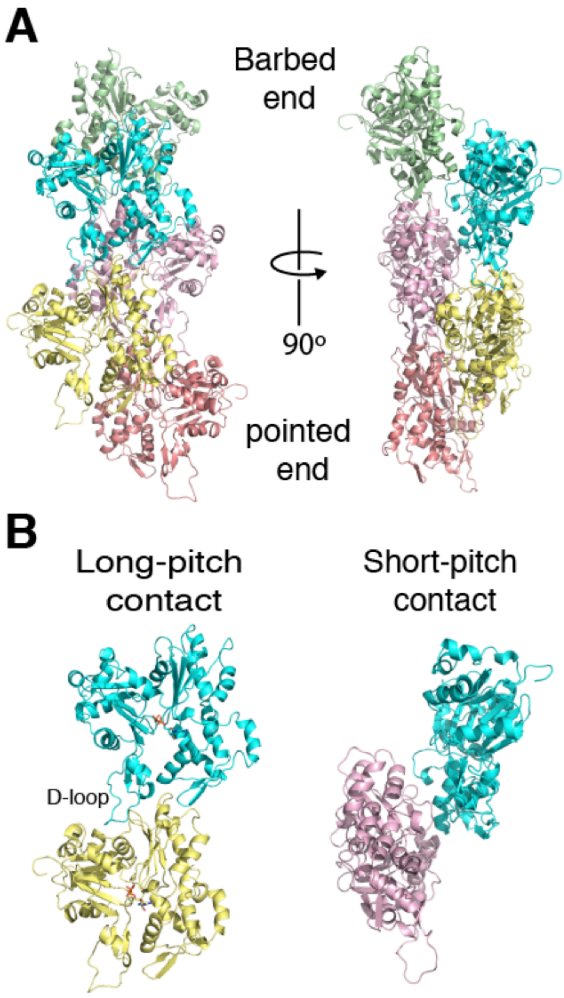


Fig 1-5 Structure of the actin filament. Model of the actin filament derived from fiber diffraction data (PDB code 2ZWH). **A.** Two orientations of a model of the actin filament comprising five subunits. **B.** Close up views of the "short-pitch" and "long-pitch" inter-subunit contacts in the filament. The ordering of the D-loop of one monomer in the 1,3-cleft of a neighboring subunit is indicated.

resolution didn't permit refinement of the actin conformation. Therefore, this structure didn't inform on the G-to-F conformational transition, but revealed that consecutive monomers in the filament are, on average, related to one another by a rotation of 166° and a translation of 28.6 Å. In 2009, Oda, et.al., taking advantage of improvements in fiber diffraction technology, published a higher resolution fiber diffraction structure with 3.3 Å radial resolution and 5.6 Å equatorial resolution (Oda, Iwasa et al. 2009). This permitted the refinement of a starting model, and for the first time, yielded a near-atomic structure of an F-actin subunit. The following year Fujii et. al., published a 6.6 Å cryo-EM structure of the filament (Fujii, Iwane et al. 2010). The fiber diffraction and cryo-EM structures were mostly mutually corroborative, but the cryo EM structure is likely to be more accurate in its details, as it involves observing actin filaments directly. In 2015 Stefan Raunser's group published a 3.7 angstrom cryo EM structure of the actin filament in complex with the filament-binding protein tropomyosin, the highest-resolution structure of the actin filament to date (von der Ecken, Muller et al. 2015). This structure is of sufficient resolution to determine side chain conformation, and it differs in its details from the earlier medium resolution cryo EM structure from Fuji and from the Oda fiber diffraction structure, but agrees more closely with the latter (RMSD = 1.7 vs. 1.9). Also in contrast to the earlier cryo EM reconstructions, which reveal that the d-loop is dynamic in the 1,3-cleft of neighboring subunits, in this study the D-loop is ordered and adopts a single conformation. One should be careful in over-interpreting the structural details with respect their relevance to a naked actin filament,

however as it's possible that the interaction with tropomyosin brings about substantial changes.

Starting with a single subunit, one can generate the canonical actin filament by iteratively applying a simple rotation of 168° about the filament axis, followed by a translation of 28.6 Å. The first application forms the inter-strand "short-pitch" interaction between the first and second actins (Fig. 1-5) the next application forms an additional "short-pitch" contact between the second and third actins and an intra-strand "long-pitch" contact between the first and third actins (Fig. 1-5) Within the resulting filament, actin makes extensive contacts with neighboring subunits. Each actin subunit interacts laterally with two subunits on the opposite strand of the helix, and longitudinally with two subunits on the same strand of the helix. The long and short-pitch interfaces bury ~990 and ~450 $Å^2$ of surface area respectively, and as a result, the former is thought to be more stable (Sept and McCammon 2001). The short-pitch interaction is mostly hydrophilic and electrostatic and comprises contacts between subdomains 2 and 4 on one monomer with subdomains 1 and 3 on the other. The longitudinal contact involves interactions between subdomains 2 and 4 on the pointed end of one monomer and subdomains 1 and 3 on the barbed end of another monomer. In all structures of actin filaments, the "D-loop" of subdomain 2, which is disordered in most crystal structures, clearly becomes ordered and packs into the hydrophobic groove of the 1,3-cleft of the longitudinally adjacent subunit (Fig. 1-5B). The long pitch contact is also mostly electrostatic and hydrophilic, with the exception

of the D-loop contact, which is hydrophobic. The D-loop adopts distinct conformations in different reports, and is thus likely to be dynamic, even in the context of a filament, although in the filament-tropomyosin cryo-EM structure, the D-loop appears to adopt a single conformation (von der Ecken, Muller et al. 2015).

Both the fiber diffraction and EM filament structures thus far discussed reveal that actin undergoes significant conformational changes upon incorporation into an actin filament. In both structures, the arrangement of secondary structure elements within the domains remains unchanged, but the domains change position relative to one another (Figs 3-6). Firstly, subdomains 1 and 2 undergo an ~17-18° rotation perpendicular to the flat plane of the molecule, such that the entire molecule is flatter and subdomains 1 and 2 are better aligned with 3 and 4 (Fig. 3B). Secondly, subdomains 1 and 2 rotate relative to 3 and 4 in the plane of the molecule, moving subdomains 2 and 4 closer to one another. This conformation is likely higher in energy than the G-actin conformation, but this cost is likely balanced by energetic gains arising from inter-subunit contacts. That the actin monomer in solution doesn't appreciably populate the F state in solution (as evidenced by crystallography),

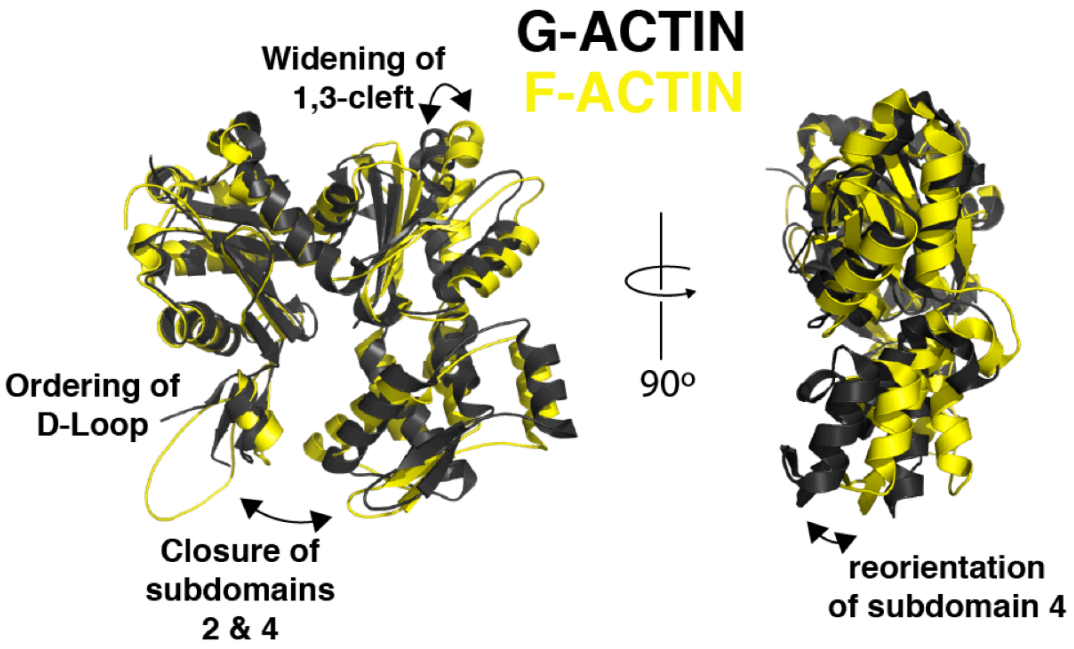


Fig. 1-6 Conformational changes that actin undergoes upon incorporation into a filament. Two views of a cartoon Representation of a model of the actin monomer, as it exists in a filament, derived from a reconstruction from electron microscopy data (PDB code 3MFP), superposed upon a g-actin crystal structure (PDB code 2HF4). The domain rotations that occur are indicated.

supports this notion. This interplay is likely important for assuring that filaments are stable but can be readily disassembled at will when necessary.

Unanswered Questions About Allostery in Actin

Taken together, the above evidence makes clear that actin's nucleotide state couples allosterically to regions of the molecule important for making inter-subunit filamentous contacts and for interacting with actin-binding proteins. Yet after over 50 years of research aimed at understanding actin biophysics, a structural mechanism that describes these phenomena remains elusive. One reason for this is the dearth of conformational diversity among structures in the PDB; with just a few exceptions, actin tends to crystallize in essentially the same conformation. Reliable observed differences between the ADP- and ATP-bound states are limited to the sensor loop, yet functional and spectroscopic differences clearly propagate to other regions. The allosteric mechanisms underlying the function of many other proteins, such as GTPases, for example, are better understood because they crystalize in multiple conformations (Lu, Jang et al. 2016). Another reason is the high level of sequence conservation among actin homologues. Often, one can gain insights into allostery by comparing homologs and finding conserved residues that co-evolve (Halabi, Rivoire et al. 2009). This is not possible for actin, although Hsp70 provides some insights into potential conformational dynamics in actin.

We have considered several models describing how nucleotide state-driven allostery might control actin's behavior. Firstly, it is possible that G-actin transiently populates a higher-energy state whose conformation more closely resembles that of a subunit within the filament, and that the degree to which this state is populated is controlled by bound nucleotide. In this model, the ATP and ADP bound actins have slightly different lowest-energy ground states, but differentially populate a common higher energy state. Such a model would explain why actin tends to crystalize in very similar conformations, regardless of nucleotide state. It is also possible that, in the ATP state (but not the ADP state), actin visits a higher energy state that doesn't resemble an F-actin subunit, but nonetheless exhibits increased affinity for a growing filament, and undergoes the G-to-F transition once associated with that filament.

Another possibility is that monomeric actin exists largely in a two-state equilibrium, whose conformations are structurally distinct and differentially populated by the ATP- and ADP-bound states. In this scenario, the ATP-bound actin would spend a minority of time in the ADP-bound conformation, and ADP-bound actin would spend a minority of time in the ATP-conformation. Here, we expect that ATP and ADP actin would have different ground states, with structural differences not reflected in current crystal structures. If this is true, it seems surprising that no one has yet captured a crystallographic snapshot of the other state. It is possible however that crystallization tends to enforce one state over the other, explaining why one tends to arise.

A related possibility is that the crystallography is, in fact, telling the whole story. Namely, that the structural perturbations driven by changes in nucleotide state are just as subtle as they appear. In this scenario, the ATP and ADP ground state conformations are nearly identical to one another. While at face value it seems less likely that a system could modulate filament dynamics with such minute structural differences, this is certainly possible thermodynamically. Given that the difference in critical concentration between ATP and ADP actin is only ~20-fold, and that a 20-fold change in dissociation constant is energetically equivalent to ~1.8 kcal/mol, equivalent to a single hydrogen bond, it begins to seem more plausible. In either of these latter two scenarios, the ATP-state would bind the filament more tightly than would the ADP-state, but the filamentous conformation would only be significantly (i.e. measurably) populated within the filament.

CHAPTER TWO

METHYL LABELING AND TROSY NMR SPECTROSCOPY OF PROTEINS EXPRESSED IN THE EUKARYOTE PICHIA PASTORIS[1]

Introduction

One of the major objectives of this thesis is to develop experimental tools to distinguish among the aforementioned models, thereby addressing long-standing questions about how nucleotide state affects actin's conformational landscape. To achieve this, I turned to NMR, a technique that is well suited to probing conformational states and for quantitatively assessing interconversion among them. In beginning to think of carrying out NMR experiments on actin, it became clear that no one had yet applied NMR to this protein.

The primary reason there are currently no reports involving NMR experiments on actin is that there are technical challenges associated with preparing an actin sample that is amenable to NMR. Firstly, at 42 kilodaltons, actin is fairly large by NMR standards, which results in unfavorable relaxation properties that cause line-broadening. There are a several strategies to overcome this problem. Firstly, growing the expression host in 100% D_2O leads to uniform deuteration of the expressed protein, which leads to longer-lived NMR signals, which narrows lines significantly.

[1] My contribution to this work consists only of experiments involving actin; development and validation of the labelling strategy are the work of Lindsay Clark and Dan Rosenbaum. I also assisted in writing the manuscript.

Secondly, by using metabolic precursors that selectively label terminal methyl groups of leucine, isoleucine, and valine with ^{13}C, one can carry out “methyl-TROSY” experiments, which, in a deuterated background, yields sharp lines, even for very large systems (Rosenzweig and Kay 2014). Both deuteration and methyl-labelling are well-established in prokaryotic expression systems, but unavailable in most eukaryotic expression systems. Unfortunately, in order to fold properly, actin depends on eukaryotic chaperones and thus requires a eukaryotic expression system. When I began to conceive of this work, insect cells were the only established system in which actin had been successfully expressed in large quantities, and insect cells are not amenable to deuteration or methyl labelling, and are thus not a good candidate.

The yeast *Pichia pastoris* is a well-established expression host amenable to uniform labelling with ^{2}H ^{13}C, and ^{15}N (Pickford and O'Leary 2004). But as I began to think of carrying out these experiments, there were no reports describing methyl labelling in this organism. Fortunately, we forged a collaboration with Dan Rosenbaum’s lab at UT southwestern, who, fortuitously, happened to be developing a system that permits labelling of isoleucine δ-methyl groups with ^{13}C, thus making it possible to express recombinant actin labeled for methyl-TROSY NMR experiments. The first part of this thesis describes this labelling strategy and its application, leading to an actin sample that is amenable to methyl-TROSY NMR experiments. This work reveals the first multi-dimensional NMR spectrum of the actin monomer, which represents a critical tool that will permit rigorous exploration of actin’s conformational

landscape, and thus represents a significant contribution to the actin field. Thie work presented in this chapter was published in 2015 (Clark, Zahm et al. 2015)

Results and Discussion

Well-resolved two-dimensional (2D) NMR spectra are essential for measuring the dynamics of backbone and sidechain moieties within proteins. Such motions are fundamental to the function and regulation of large protein complexes and membrane proteins. NMR spectra of large macromolecules suffer from poor dispersion and line broadening due to rapid transverse relaxation of nuclear magnetization and spectral crowding. To overcome this problem, proteins can be specifically labeled with ^{13}C in the methyl groups of isoleucine, leucine, and valine residues using ^{13}C α-ketoacid precursors in E. coli (Gardner and Kay 1997, Goto, Gardner et al. 1999). When paired with selective protonation in an otherwise deuterated background (Rosen, Gardner et al. 1996), this approach takes advantage of the favorable relaxation properties of 13C-methyl groups with the application of Transverse Relaxation Optimized Spectroscopy (TROSY) (Pervushin, Riek et al. 1997, Ollerenshaw, Tugarinov et al. 2003). However, these methods have remained unavailable for many eukaryotic proteins due to poor expression and folding in E. coli resulting from lack of required chaperones, lack of proper post-translational modifications, or improper membrane composition.

Several different eukaryotic hosts, including fungi (Miyazawa-Onami, Takeuchi et al. 2013), insect cells (Nygaard, Zou et al. 2013, Kofuku, Ueda et al. 2014), and mammalian cells (Werner, Richter et al. 2008), have been used to overexpress proteins for NMR. While these systems have succeeded in producing amino acid-specific and uniformly ^{15}N or ^{13}C labeled material (Hansen, Petros et al. 1992, Strauss, Bitsch et al. 2005, Chen, Cheng et al. 2006, Fan, Shi et al. 2011, Gossert, Hinniger et al. 2011), the high expense and difficulties in perdeuteration have limihotheir widespread use for larger eukaryotic proteins. The methylotrophic yeast Pichia pastoris is a well-established expression host (Cereghino and Cregg 2000) for proteins that cannot be made in E. coli - eukaryotic membrane proteins such as ATP transporters (Lee, Urbatsch et al. 2002), ion pumps (Strugatsky, Gottschalk et al. 2003) and G-protein coupled receptors (Shimamura, Shiroishi et al. 2011, Hino, Arakawa et al. 2012) have been successfully overexpressed in and purified from this organism. Genetic manipulation, transformation, and growth of P. pastoris are more rapid than for higher eukaryotes such as insect cells and mammalian cells.

Overexpression using the tightly regulated AOX1 promoter often yields milligram quantities of recombinant protein per liter of suspension culture (Cereghino and Cregg 2000). P. pastoris is also favorable for NMR studies given its ability to grow on defined minimal media, uptake isotope-containing precursors, and efficiently incorporate deuterium at non-exchangeable sites (Morgan, Kragt et al. 2000). Despite conservation of branched-chain amino acid biosynthesis pathways from E. coli **(**Fig.

2-1), site-specific methyl labeling using α-ketoacid precursors has not been reported in. P. pastoris.

We initially explored the use of ^{13}C-methyl α-ketobutyrate in P. pastoris cultures to label maltose binding protein (MBP) with ^{13}C at the δ1-methyl groups of Isoleucine (Ile) residues. MBP has well-characterized ^{1}H-^{13}C 2D NMR spectra (Gardner, Zhang et al. 1998) and is highly expressed in P. pastoris (Li, Leung et al. 2010). We collected ^{1}H-^{13}C heteronuclear single quantum coherence (HSQC) spectra on MBP that was labeled by addition of 13C-methyl α-ketobutyrate to the culture media (Fig. 2-2A). Resonances for all 22 Ile δ1-methyl groups of MBP (Gardner, Zhang et al. 1998) are observed in our spectrum (Fig2-2A, Table 2-1), while little signal is present in other regions (indicating lack of “bleed-through” of the isotope into other amino acids - see Fig. 2-3). Based on tryptic peptide mass spectra (Fig. 2-2B), we estimate the efficiency of incorporation for the α-ketobutyrate-derived ^{13}C methyl probe to be 51±7% in a protonated background. The power of TROSY to obtain spectra of high-molecular weight species can only be exploited in the context of partial or full deuteration (Gardner, Rosen et al. 1997, Wider and Wuthrich 1999, Ruschak and Kay 2010), which eliminates dipolar relaxation effects of surrounding protons on a given 13C-methyl spin system. To assess simultaneous ^{13}C methyl labeling and perdeuteration in our system, we made samples of MBP in both P. pastoris and E. coli. We quantified the level of Ile δ1 labeling in P. pastoris-derived deuterated MBP by comparing intensities to a concentration-matched E. coli sample (with assumed full incorporation

at Ile δ1 sites and deuteration), yielding a labeling efficiency of 45 ± 6% (Table 2-2). The total deuteration level of *P. pastoris*-expressed MBP was estimated at 90% through ESI-LC-MS analysis (Fig 2-4); a comparison of labeling efficiency and yields of recombinant MBP from P. pastoris vs. E. coli is shown in Table 2-3). Addition of α-ketoisovalerate led to very modest labeling of leucine δ- and valine γ-methyl groups (< 5%), suggesting that these sites could be labeled with significant future optimization.

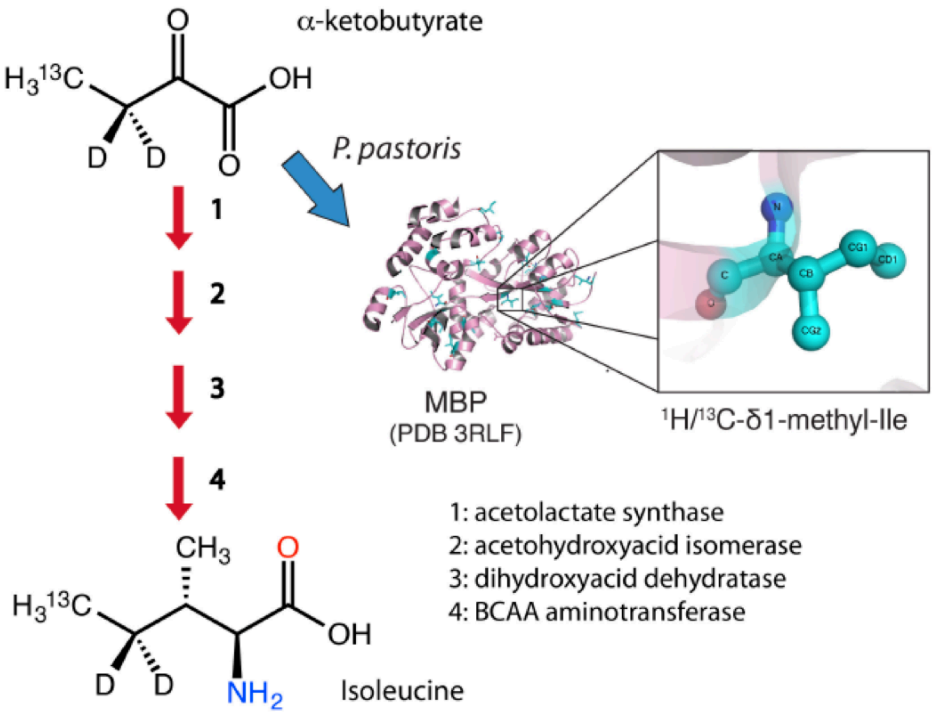


Fig. 2-1 Incorporation of 13C-methyl groups at the δ1 position of isoleucine residues in proteins expressed in *Pichia pastoris*

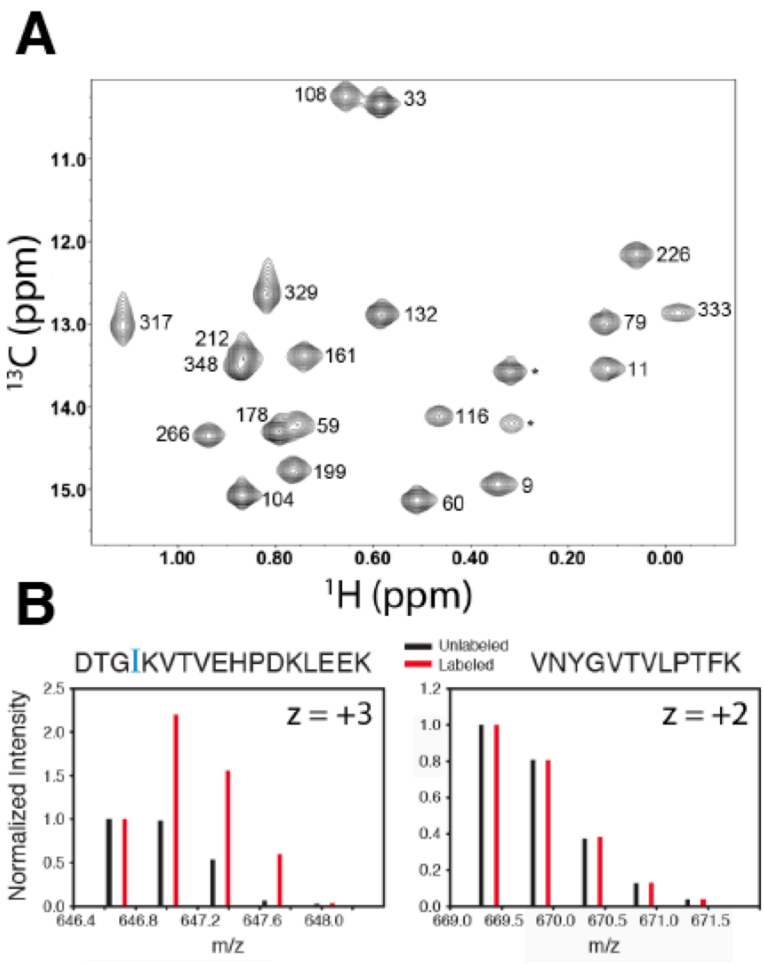


Fig. 2-2 Labeling of δ1-methyl groups of MBP expressed in *Pichia pastoris*.
A. ^{1}H-^{13}C HSQC spectrum of 225 μM MBP labeled with α-ketobutyrate. Spectrum was recorded at 25 °C on a Varian 800 MHz spectrometer. Peaks corresponding to Ile δ1-methyl groups are labeled in reference to assigned spectra (Gardner et al. 1998). Two unassigned peaks likely arising from differences in constructs are denoted with an asterisk. **B.** Mass spectra of tryptic peptides containing Ile (left) and not containing Ile (right)

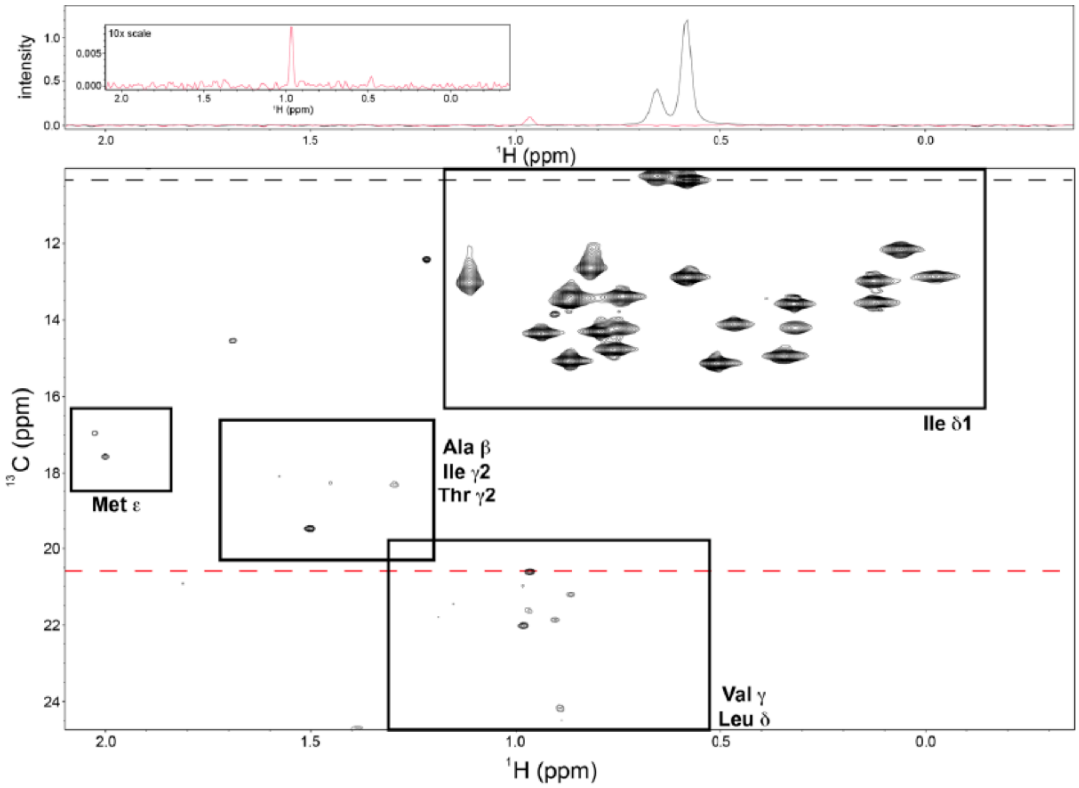


Fig. 2-3 Expanded view of the ^{1}H-^{13}C HSQC spectrum of isoleucine δ1-methyl labeled maltose binding protein shown in Fig. 2-2A. Top panel shows horizontal slices of the 2D dataset (bottom panel), taken at approximately ^{13}C = 10.3 ppm (black; Ile δ1) and 20.6 ppm (red; Val/Leu) to show representative signal-to-noise in the spectrum for the labeled Ile δ1 methyl groups versus the unlabeled (natural abundance ^{13}C) methyl groups of other amino acids. Inset of the top panel shows the ^{13}C = 20.6 ppm trace at 10× vertical scale of the surrounding panel to provide a clearer sense of signal-to-noise. Signal-to-noise measurements for all 22 Ile δ1-methyl peaks resulted in an average S/N ratio of 280.

P. pastoris (MBP res #)	1H	13C	*E. coli* (MBP res #)	1H	13C	δ 1H	δ 13C
9	0.34	14.96	9	0.325	14.504	0.02	0.46
11	0.12	13.56	11	0.108	13.1151	0.01	0.44
33	0.58	10.36	33	0.576	10.246	0.00	0.11
59	0.75	14.24	59	0.739	13.813	0.01	0.43
60	0.51	15.15	60	0.408	14.627	0.10	0.52
79	0.13	13	79	0.171	12.326	-0.04	0.67
104	0.87	15.09	104	0.856	14.654	0.01	0.44
108	0.66	10.26	108	0.65	10.048	0.01	0.21
116	0.46	14.13	116	0.462	13.724	0.00	0.41
132	0.58	12.9	132	0.564	12.401	0.02	0.50
161	0.74	13.4	161	0.761	13.139	-0.02	0.26
178	0.79	14.32	178	0.857	13.918	-0.07	0.40
199	0.76	14.79	199	0.749	14.231	0.01	0.56
212	0.87	13.44	212	0.872	13.1	0.00	0.34
226	0.06	12.17	226	0.072	11.812	-0.01	0.36
266	0.94	14.37	266	0.925	13.82	0.01	0.55
317	1.11	13.04	317	1.105	12.54	0.01	0.50
329	0.82	12.66	329	0.821	12.343	0.00	0.32
333	-0.03	12.88	333	-0.149	12.276	0.12	0.60
348	0.87	13.54	348	0.872	13.073	0.00	0.47
---	---	---	368	0.684	15.275	---	---
*	0.32	14.22	---	---	---	---	---
*	0.32	13.6	---	---	---	---	---
					Average δ:	**0.01**	**0.43**

Table 2-1 Comparison of MBP Ile δ1 chemical shifts. Chemical shifts from MBP expressed in *P. pastoris and E. coli* are listed in blue and yellow, respectively. Although the constructs of MBP used in each study vary slightly, residue numbering is kept in reference to the *E. coli* spectra for simplicity. The construct used in our study does not include Ile 368 but contains two additional peaks (notated with an asterisk) that did not readily correlate with previously published spectra and likely arises from small differences in the constructs. The ^{1}H chemical shifts between samples remain nearly identical, however the E. coli-expressed sample spectra display a nearly uniform shift upfield of 0.43ppm in the ^{13}C dimension. This shift is likely a combination of small differences in referencing and the effects of deuteration (Gardner and Kay 1997), as the E. coli-expressed sample is highly deuterated and the P. pastoris-expressed sample is protonated.

Peak number	Volume (*E. coli*)	Volume (*P. pastoris*)	Percentage	S/N (*E. coli*)	S/N (*P. pastoris*)	Percentage
1	7.93E+07	3.40E+07	42.94	1.69E+04	7.45E+03	44.15
2	5.96E+07	2.36E+07	39.64	1.27E+04	5.18E+03	40.77
3	5.07E+07	2.21E+07	43.50	1.08E+04	4.83E+03	44.73
9	1.73E+07	8.99E+06	52.06	3.68E+03	1.97E+03	53.53
12	1.23E+08	4.85E+07	39.54	2.61E+04	1.06E+04	40.66
15	4.56E+07	2.06E+07	45.13	9.72E+03	4.51E+03	46.41
17	1.25E+07	7.04E+06	56.17	2.67E+03	1.54E+03	57.76
18	5.18E+07	2.21E+07	42.63	1.10E+04	4.83E+03	43.84
19	7.13E+07	2.79E+07	39.13	1.52E+04	6.11E+03	40.24
20	4.74E+07	1.96E+07	41.26	1.01E+04	4.29E+03	42.43
		Average	44.34		Average	45.60

Table 2-2 Intensity comparison of ten well-resolved cross peaks in methyl HMQC spectra. On the basis of measured intensity in terms of peak volume and S/N, the percentage incorporations of $^{1}H/^{13}C$ at the Ile δ1-methyl positions are 44% ± 6% and 46% ± 6 %, respectively. Listed peak numbers are arbitrary.

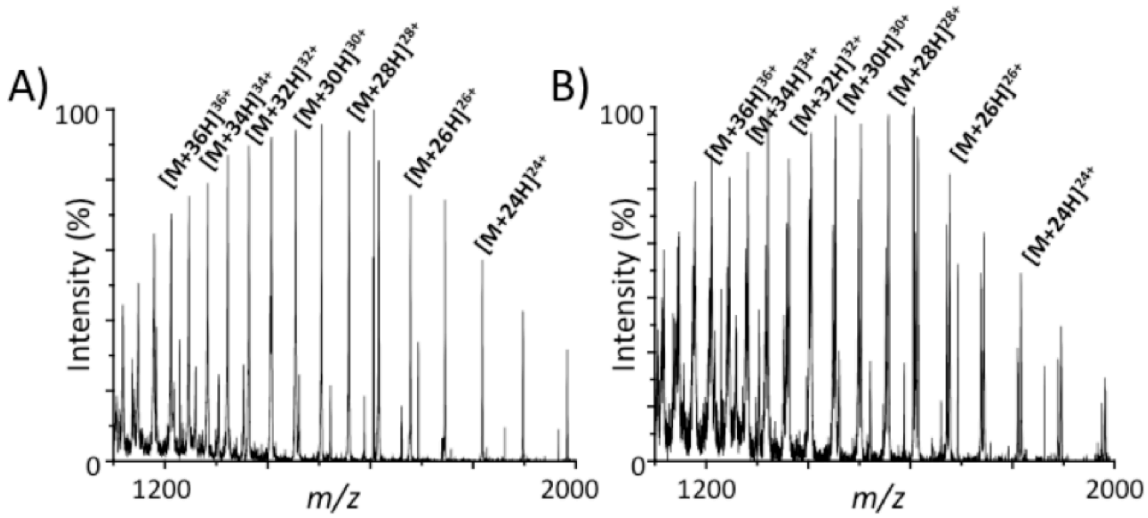


Fig. 2-4 Mass Spectra of deuterated ^{13}C-methyl-Ile-labeled MBP. **A.** ESI mass spectrum of the principal species in the LC-MS on labeled MBP purified from *E. coli*. (b) Mass spectrum from analogous sample from P. pastoris. Protonated charge states are annotated above each peak in the mass to charge spectrum (m/z). Calculated mass from the E. coli spectrum is 43,606 Da, corresponding to 91% deuterium labeling at non-exchangeable sites. Calculated mass from the P. pastoris spectrum is 43,566 Da, corresponding to 90% deuterium labeling at non-exchangeable sites.

	Protonated growth		Deuterated growth		
	Yield (mg/L)	Ile(δ1) incorporation rate (%)	Yield (mg/L)	Ile(δ1) incorporation rate (%)	Deuteration level (%)
E. coli	Not tested	Not tested	40	100	91
P. pastoris	60	51±7	10	45±6	90
K. lactis	10	Not tested	3.6	67±6	>90

Table 2-3 Comparison of MBP expressed in *E. coli* and *P. pastoris* from this manuscript and wild-type *K. lactis* from a recent study (Miyazawa-Onami et al. 2013). 100% Ile $\delta 1$ incorporation in E. coli was assumed for comparison.

The impetus for using P. pastoris for ^{13}C methyl labeling is to access proteins that are not amenable to expression and purification from E. coli – for example, the eukaryotic cytoskeletal protein actin. Actin's capacity to shuttle between monomeric and polymeric states arises from conformational dynamics that permit it to exist in distinct globular and filamentous forms (Pollard and Cooper 1986, Oda, Iwasa et al. 2009). NMR dynamics measurements would represent a significant new tool to study the biophysics of actin polymerization and interactions with regulatory molecules (Schmid, Sherman et al. 2004, Kudryashov and Reisler 2013). While the structure of actin monomer has been determined by X-ray crystallography (Otterbein, Graceffa et al. 2001, Rould, Wan et al. 2006) and actin filaments have been characterized by electron microscopy (Oda, Iwasa et al. 2009, Fujii, Iwane et al. 2010, Galkin, Orlova et al. 2015), expression of isotopically labeled actin for NMR has not been reported. Actin cannot be expressed at high levels in E. coli because of the lack of eukaryotic chaperone systems that are necessary for folding.

Biophysical characterization of actin is intrinsically difficult because actin polymerizes at concentrations above 100 nM. We therefore attempted to express a non-polymerizable Drosophila 5C actin (51.5 kDa, 94% identity to human actin) mutant in P. pastoris with mutations that impair the fast growing "barbed-end" of the filament (Zahm, Padrick et al. 2013). However, the mutant proved toxic, presumably because it interferes with the polymerization of endogenous actin. To solve this problem, we

generated a C-terminal fusion to human thymosin β4, an actin binding protein that blocks the intact, slow-growing "pointed-end" and thus ameliorates toxicity (Noguchi, Kanzaki et al. 2007). This strategy resulted in high expression levels (10 mg/L of culture) and purification to homogeneity (Fig. 2-5).

A representative HMQC spectrum of ^{13}C-δ1-methyl -Ile actin is shown in Fig 2-6. Notably, for a protein with 27 Ile residues, we observe 30 peaks in the ^{1}H-^{13}C spectrum, likely reflecting slow-exchange processes at some sites. Taking advantage of the ability to highly deuterated proteins in P. pastoris, we repeated expression of Drosophila 5C actin in cultures where cells were adapted to D_2O-containing media prior to induction, resulting in 2.5 mg/L of ^{13}C-methyl perdeuterated actin. Spectral lines in the ^{1}H-^{13}C TROSY HMQC of the deuterated sample were much narrower than lines in the HMQC spectrum of non-deuterated actin (Fig 2-6).

Future use of TROSY NMR methods to study the dynamics of high-MW mammalian protein complexes and membrane proteins will depend on the tractability of isotope incorporation. We have demonstrated efficient incorporation of ^{13}C at the Ile δ1-methyl groups of proteins expressed in P. pastoris, a robust eukaryotic expression host. In conjunction with perdeuteration, we acquired high-quality ^{1}H-^{13}C methyl TROSY spectra on Drosophila actin, which were unobtainable before. This development, along with similar approaches using other yeast systems

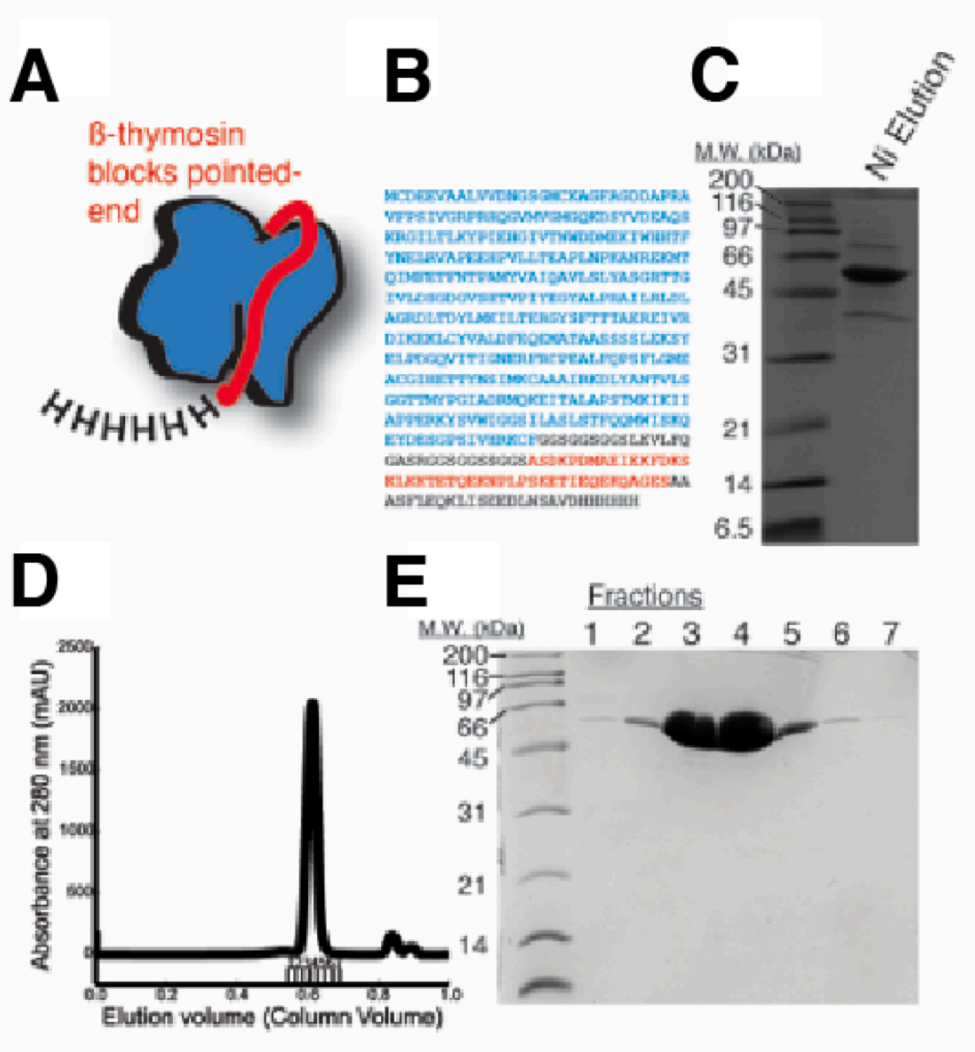


Fig. 2-5 Recombinant actin for NMR. **A.** Cartoon and **B.** sequence of actin-thymosin fusion expressed in P. pastoris. Actin is in blue, thymosin is in red, with linker or tagged regions in black. **C.** SDS-PAGE gel showing actin-thymosin (51.5kD) as the main component in the elution off of Ni-NTA. **D** Size exclusion profile of purified actin-thymosin injected on a Superdex 200 10/300 column. Fractions collected are notated. **E** SDS-PAGE gel of fractions from **D**.

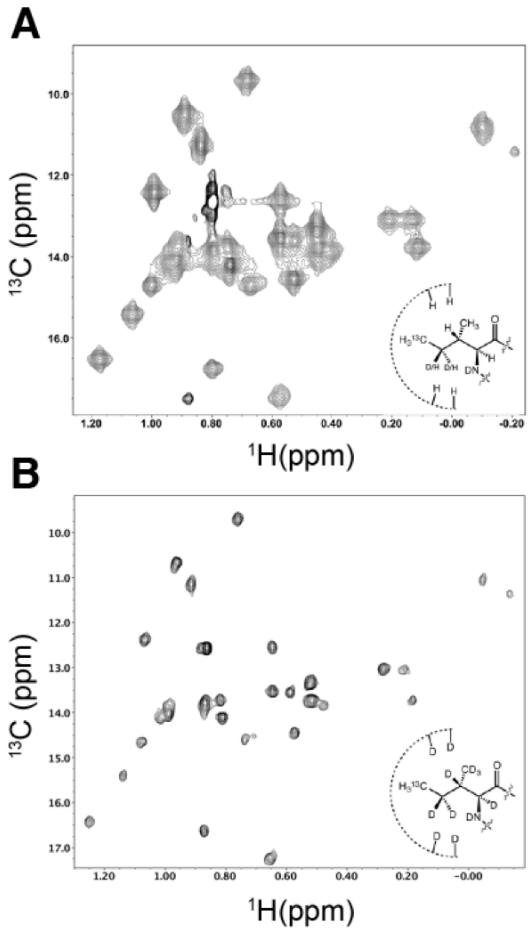


Fig. 2-6 NMR spectra of Drosophila actin labeled and overexpressed in *Pichia pastoris*. A. 1H-13C HMQC spectrum of 13C-Ile δ1-methyl-labeled actin (180 μM). **B.** TROSY-HMQC spectrum of perdeuterated, 13C-Ile δ1-methyl-labeled actin (150 μM). Spectra were recorded at 25 °C on a Varian 800 MHz spectrometer

(Miyazawa-Onami, Takeuchi et al. 2013), will allow 2D NMR spectroscopy to be applied to many previously intractable proteins.

Materials and Methods

MBP expression and purification

The gene encoding E. coli maltose binding protein with a C-terminal hexahistidine tag (41.5 kDa) was inserted into the pPICZ B vector (Invitrogen) under control of the AOX1 promoter and transformed into the KM71H strain. Expression was screened with increasing concentrations of Zeocin selection marker (100-1000 mg/mL), and a high expressing clone was chosen for future experiments. Cells were cultured at 28°C in BMG media (1% glycerol, 100 mM potassium phosphate pH 6.0, 1.34% YNB, 4 x 10-5% biotin) with shaking at 250 rpm. Once cells reached saturation (~36hrs), cells were spun down and resuspended in an equal volume of BMMH media without methanol and supplemented with DMSO (100mM potassium phosphate pH 6.0, 1.34% YNB, 4 x 10-5% biotin, 2.5% DMSO). Cells were incubated for an additional 4-5 hrs to allow for metabolism of residual glycerol. One hour prior to induction with methanol, a sterile-filtered solution of α-ketobutyric acid (methyl-^{13}C, 99%; 3,3-D2, 98%; Cambridge Isotope) was added to a final concentration of 100 mg/L. Protein expression was induced with 0.5% methanol and was maintained by additions of 0.5% methanol every 12 hours. After 36 hours, cells were harvested by

centrifugation at 5000g and frozen at -80°C. Pichia were resuspended in lysis buffer containing 200 mM sodium chloride, 20mM Tris pH 7.4, and protease inhibitors (160 µg/mL benzamidine, 2.5 µg/mL leupeptin, 1 mM PMSF, 1 µM E-64) and lysed by five passes through a microfluidizer (Microfluidics M-110P) at 25,000psi. Insoluble material was removed by centrifugation at 10,000 g, and MBP was purified through Ni-NTA and amylose affinity chromatography, followed by size exclusion chromatography using a Superdex 200 10/300 column (GE Healthcare). MBP was concentrated to 500ul (measured by UV Abs 280nm) in Amicon 30kDa MWCO concentrators. An equimolar amount of β-cyclodextrin (Acros Organics) was added to the MBP sample, and the mixture was dialyzed overnight against 20 mM HEPES pH 7.2 at 4°C.

To generate the deuterated MBP sample, cells were adapted to deuterated media as follows. 100 uL of an initial starter culture in protonated BMGH was used to inoculate 50 mL of BMGH media containing 90% D_2O (Cambridge Isotope Laboratories) and grown to an OD of 7. 100 uL of the 90% D_2O culture was used to inoculate 50 mL BMGH containing 100% D_2O, and again grown to an OD600 of 7. 100 µL of the fully deuterated culture was used to inoculate 50 mL BMGH containing 100% D_2O and 1% deuterated d8-glycerol (CIL) in place of protonated glycerol. This culture was grown to an O.D. of 7 and used in its entirety to inoculate 250mL of BMGH containing 100% D_2O and 1% d8-glycerol and grown to an O.D. of 20. This culture was spun down at 3000 g for 10 min and resuspended in 250 mL methanol-free BMMH

containing 100% D_2O. The culture continued shaking for 12 hours to ensure full metabolism of glycerol. One hour prior to induction, a sterile-filtered solution of α-ketobutyric acid (methyl-^{13}C, 99%; 3,3-D2, 98%; CIL) in D_2O was added to a final concentration of 200 mg/L. Induction was achieved by addition of d4-methanol (CIL) to a final concentration of 0.5%. Expression proceeded for a total of 36 hours at 28°C, with a supplementation of 0.5% 100% d4-methanol every 12 hrs. Cells were harvested by centrifugation and purified in a manner identical to that used for the protonated sample, except that the final gel filtration buffer was made using 100% D_2O instead of water and the final dialysis buffer prior to NMR was made with 100% D_2O.

Deuterated MBP generated in E. coli was grown as described (Gardner, Zhang et al. 1998). MBP was purified from 1L of E. coli in an identical manner to deuterated MBP expressed in P. pastoris.

Actin expression and purification

The gene encoding the polymerization-incompetent Drosophila 5C actin mutant [D287A, V288A, D289A] was cloned into pPICZ B (Invitrogen) using EcoRI and NotI restriction sites. The construct used to generate the protonated actin sample contained a C-terminally fused human β-thymosin-4 followed by a hexahistidine tag. The construct was transformed via electroporation into the GS115 strain of Pichia pastoris. Expression was screened with increasing concentrations of Zeocin selection marker (0.5-2 mg/mL) and a high expressing clone was chosen for future experiments.

A single colony was selected from a BMGH (BMG + 0.004% histidine) agar plate streaked with the high-expressing clone, and grown at 30°C to an O.D. of 5-7 in 35 mL BMGH media in a 125 mL baffled flask. The entire starter culture was used to inoculate 1L of BMGH media, and cells were grown until OD600 ~17. Cells were collected by centrifugation at 3000 g for 15 minutes and resuspended in 3L of BMMH without methanol (BMM + 0.004% histidine) and continued shaking for 12 hr to metabolize any remaining glycerol. An hour prior to induction, a sterile-filtered solution of α-ketobutyric acid (methyl-13C, 99%; 3,3-D2, 98%; Cambridge Isotope Laboratories) was added to a final concentration of 100 mg/L. Induction was achieved by addition of methanol to a final concentration of 0.5% and expression continued for 24 hours. Cells were harvested by centrifugation at 3000 g for 10 minutes and resuspended in 200 mL 50 mM Tris pH 8.0, 250 mM potassium chloride, 0.1 mM calcium chloride, 0.2 mM ATP, and 0.5 mM β-mercaptoethanol, supplemented with protease inhibitors (1 μg/mL leupeptin, 500 ng/mL pepstatin, 1 mM benzamidine, 1 μg/mL antipain, and 1 mM PMSF). Cells were lysed by four passes through a microfluidizer (Microfluidics M-110P) at 25,000psi. Insoluble material was removed by centrifugation at 50,000 g, and the supernatant was subjected to Ni-NTA affinity chromatography. The resulting material was further purified on a 4 mL Source 15Q anion exchange column. Fractions containing actin were pooled and concentrated to 1 mL and exchanged via size exclusion chromatography (Figure S5) on a Superdex 200 10/300 column (GE Healthcare) into a buffer containing 20 mM potassium

phosphate pH 7.0, 50 mM potassium chloride, 0.2 mM ATP, 1 mM TCEP, 0.1 mM calcium chloride, and 1 mM sodium azide.

To generate the deuterated actin sample, cells were grown in BMGH and were adapted to increasing concentration of D_2O in the media as follows. 100 uL of the initial starter culture was used to inoculate 35 mL of BMGH media containing 90% D_2O (Cambridge Isotope Laboratories) and grown to an OD of 5-7. 100 uL of the 90% D_2O culture was used to inoculate 35 mL BMGH containing 100% D2O, and again grown to an OD600 of 5-7. 100 μL of the fully deuterated culture was used to inoculate 35 mL BMGH containing 100% D2O and 1% deuterated d8-glycerol (Sigma) in place of protonated glycerol. This culture was grown to an O.D. of 5-7 and used in its entirety to inoculate 1L of BMGH containing 100% D_2O and 1% d8-glycerol and grown to an O.D. of 17. This culture was spun down at 3000 g for 10 min and resuspended in 30 mL methanol-free BMMH containing 100% D_2O. The resuspended culture was distributed equally among three 4L flasks, each containing 1L of methanol-free BMMH. The cultures continued shaking for 12 hours to ensure full metabolism of glycerol. One hour prior to induction, a sterile-filtered solution of α-ketobutyric acid (methyl-^{13}C, 99%; 3,3-D2, 98%; Cambridge Isotope) in D_2O was added to a final concentration of 200 mg/L. Induction was achieved by addition of d4-methanol (Sigma) to a final concentration of 0.5%. Expression proceeded for a total of 48 hours at 30°C, with a supplementation of 5 mL 100% d4-methanol at 24 hrs. Cells were

harvested by centrifugation and purified in a manner identical to that used for the protonated sample, except that the final gel filtration buffer was made using 100% D_2O instead of water.

Mass Spectrometry

Purified labeled and unlabeled MBP samples were reduced in 100 mM DTT and alkylated with iodoacetamide. Samples were trypsinized at 37 °C overnight. Digestion products were analyzed on a Shimadzu IT-TOF LC-MS system equipped with NESP-100 NANO-ESI ion source. LC was run on a Chromolith CapROD RP-18e column with a H_2O/Acetonitrile gradient containing 0.1% Formic Acid. Incorporation efficiency of 13C at the δ1 position of Isoleucine residues was quantified in the following manner: Five peptides were identified that had high signal/noise in the tryptic LC-MS data for both labeled and unlabeled MBP, and which had a unique Isoleucine present in the sequence (z = +2 for all). For each peptide, an isotopic distribution was observed, corresponding to increasing numbers of natural-abundance 13C incorporated into the peptide – M+1 has one 13C atom, M+2 has two 13C atoms, etc. We normalized the intensities of the peaks in each of these distributions by dividing by the intensity of the M+0 peak (i.e. no 13C atoms) – see Fig. 2b. The % incorporation of 13C by addition of the label was then determined by: [(Intensity M+1 labeled) – (Intensity M+1 unlabeled)]÷ (Intensity M+1 labeled). Similar values could be calculated for the M+2

peaks. Reported efficiency is the average value for the M+1 and M+2 peaks for the five peptides.

NMR Spectroscopy

All NMR data on MBP and actin were acquired on a Varian Inova 800 MHz NMR spectrometer equipped with a cryogenically-cooled $^{1}H/^{13}C/^{15}N$ probe operating at 25 °C. NMR experiments on MBP were carried out on a sample containing a 1:1 ratio of labeled MBP and β-cyclodextrin (225 μM each) in 20 mM HEPES pH 7.2, 10% D_2O. ^{1}H-^{13}C HSQC spectra were collected with spectral widths of 11990 Hz and 4026 Hz and acquisition times of 85 ms and 24 ms in the ^{1}H and ^{13}C dimensions, respectively. An inter-scan delay of 1s was employed between successive transients. Total acquisition time was approximately 2.2 hr.

Actin NMR data were acquired on samples at 180 μM (protonated, Ile δ1-methyl ^{13}C labeled) and 150 μM (deuterated, Ile δ1-methyl 1H/13C labeled, HMQC) concentrations in 90%H_2O/10% D_2O and 100% D_2O, respectively. ^{1}H–^{13}C HMQC (Griffey, Poulter et al. 1983, Ollerenshaw, Tugarinov et al. 2003) spectra were acquired with 9615.4 Hz and 4026 Hz in ^{1}H and ^{13}C dimensions, respectively. The acquisition times in directly detected ^{1}H dimension and indirectly detected ^{13}C dimension were 106 ms and 16 ms, respectively. An inter-scan delay of 1 s was employed between successive transients during data acquisition.

Data were processed using NMRPipe (Delaglio, Grzesiek et al. 1995). Data sets were zero-filled prior to Fourier transformation. The directly and indirectly detected time domain data were processed by applying a 90° phase-shifted squared sine bell or a Gaussian filter.

In order to compare the incorporation of $^{1}H/^{13}C$ at Ile δ1-methyl positions in a deuterated background, MBP was overexpressed in E.coli and P. pastoris, purified and complexed with a 1:1 molar ratio of β-cyclodextrin (Acros Organics), and dialyzed overnight against 20mM HEPES pH 7.2 prepared in 100% D_2O. NMR samples of each protein were prepared at identical concentrations (240μM). The 2D ^{1}H-^{13}C HMQC spectra on both samples were acquired at 298 K using identical acquisition parameters. The ^{1}H-^{13}C HMQC spectra were acquired with sweep widths of 9615.4 Hz and 4026 Hz in ^{1}H and ^{13}C dimensions, respectively. Acquisition times in the directly detected ^{1}H dimension and indirectly detected ^{13}C dimension were 106 ms and 16 ms, respectively. An inter-scan delay of 1 s was employed between successive transients during data acquisition.

Both datasets were processed identically using NMRPipe (Delaglio, Grzesiek et al. 1995). Ten well-resolved cross peaks in methyl HMQC spectra were identified for intensity analysis. The intensities of individual cross-peaks were measured in terms

of peak height (data not shown), peak volume and signal-to-noise ratio (S/N). The intensity analysis was performed using NMRPipe as well as using the Analysis module in CCPNMR (Vranken, Boucher et al. 2005). The incorporation of $^{1}H/^{13}C$ at δ-methyl positions in P. pastoris MBP was measured relative to E. coli MBP.

CHAPTER THREE

USING NMR TO EXPLORE NUCLEOTIDE-DEPENDENT CONFORMATIONAL DYNAMICS OF ACTIN

Introduction

One of the many applications of NMR spectroscopy is probing protein motions. A variety of NMR techniques serve this purpose, each of them useful for motions that span a limited range of motional frequency. Together, these techniques permit the quantitative assessment of molecular motions on timescales that range across twelve orders of magnitude, from picoseconds to seconds. When a protein undergoes a conformational change, it transforms to a new state with an associated energy. One can visualize the ensemble of possible states pictorially as an energy "landscape". In such a construction, the ground, lowest energy conformation exists as a global minimum. Within this minimum, there exists an ensemble of states of nearly equivalent energy, separated by low kinetic barriers. In the schematic representation, such states appear as fine structure within the well that represents the global minimum. The kinetic barriers that exist among these states are sufficiently low that thermal energy is sufficient to drive rapid interconversion among them. Such transitions occur on the picosecond to

nanosecond timescale, and arise from rapid side chain dynamics and bond rotations and vibrations.

In addition to this rapidly-interconverting ensemble of energetically equivalent states, proteins also undergo conformational transitions to states of significantly higher energy. In the energy landscape, these states appear as local minima peripheral to the ground-state global minimum. Populating such states typically requires passage through a more substantial kinetic barrier, leading to interconversion rates on the millisecond to microsecond timescale. Furthermore, because these species are of significantly higher energy, they are poorly populated, and have been called “excited states” in the NMR community (not to be confused with excited electronic states that govern chemical reactivity). Such conformational transitions are of particular relevance to biological processes, and occur in the interconversion among protein folding intermediates, conformational selection in enzyme catalysis, and in allostery (Anthis and Clore 2015). By virtue of being poorly populated, these transient excited states have defied crystallographic characterization, which, in most cases, only permits access to the highly-populated ground state.

There are several NMR techniques that permit quantitative characterization of these states. Among them, the one used most frequently is the Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiment. This experiment takes advantage of the fact that chemical exchange processes contribute to the

relaxation properties of an NMR sample in a way that one can quantify, which ultimately yields information about the excited state.

To understand how chemical exchange manifests in an NMR spectrum, it is helpful to imagine a simplified, hypothetical NMR sample that consists of a single spin in a molecule that is interconverting between two states, A and B. Owing to the disparate environments in the two states, the spin will precess at different frequencies in each of the two states, and will thus have different chemical shift values in each of the two states. One can then think of the quantity $\Delta\omega$ the difference between the chemical shift values in the two states, also a frequency, and the quantity k_{ex}, which is equal to the sum of the forward and reverse rates governing the exchange process (also a frequency). If k_{ex} far exceeds $\Delta\omega$, the system is said to be in "fast exchange," and one sees a single line in the NMR spectrum with a chemical shift value that represents the population-weighted average of the chemical shift values in the two states. I.e., if the system were arranged such that the population of state B was only 10%, one would see a single peak nearly coincident with the chemical shift value of state A, but shifted towards the chemical shift associated with state B by 10% of the total distance between them. If k_{ex} is comparable to $\Delta\omega$, the system is said to be in "intermediate exchange." In this case, the exchange contribution convolves the two frequencies in a way that the resulting signal loses its sinusoidal character, becomes chaotic, and as a result, the line becomes very broad and disappears from the spectrum.

If k_{ex} falls far below $\Delta\omega$, the system is said to be in "slow exchange." For this case, one observes two peaks in the spectrum, one corresponding to each of the respective chemical shift values from the two states in the exchange process. The respective intensities of these peaks are proportional to the degree which they are populated. Thus, it is possible that, in an exchange process, one of the two states is poorly populated to an extent that, while it is theoretically detectable in an NMR spectrum, the signal-to-noise ratio in an actual measurement precludes visualizing it. Such states, that are too poorly populated to visualize have been named "excited states." The CPMG experiment permits one to characterize these states, despite their being invisible. Later in this chapter, I will go into more detail about how the CPMG experiment works, but for now it suffices to say that it yields the following information for slow to intermediate exchange processes: The chemical shift values corresponding to the two states in the exchange process (ground and excited), the populations of those states, and the exchange rates between them.

It is our hypothesis that the control of actin polymerization by bound nucleotide may be driven by invisible exited states. Having generated actin that is amenable to NMR spectroscopy, I proceeded with experiments ultimately aimed at probing actin's millisecond to microsecond conformational landscape, with a goal of determining if and how it changes as a function of nucleotide state.

The actin C-terminus couples allosterically to each of actin's four subdomains

Having the expression and NMR labelling strategies in hand, the next goal was to assign the peaks in the spectrum. To do this, I generated a construct fused at its C-terminus to β-thymosin (Noguchi, Kanzaki et al. 2007, Clark, Zahm et al. 2015), with a TEV cleavage site between the actin and the thymosin. For this construct, cleavage with TEV protease yields non-native actin whose C-terminus remains fused to a flexible GGS linker and the N-terminal remnant of the TEV cleavage site (Fig 3-1). This construct expressed well, and I was able to readily cleave the β-thymosin peptide, purify the actin to homogeneity, and achieve concentrations suitable for NMR experiments. Expression of this construct in *Pichia pastoris,* in the presence of α-ketobutyrate yielded actin with its isoleucine residues labelled with ^{13}C at the δ-methyl positions (Clark, Zahm et al. 2015), and resulted in ^{1}H-^{13}C HSQC/HMQC correlation spectra with peaks sufficiently resolved to permit assignment of each isoleucine (Fig. 3-2). For the NMR experiments in which I assigned the peak, the actin is bound to Ca-ATP. I initially attempted to use methyl-NOESY experiments to facilitate the assignment, which, if working ideally, would have determined the distances among the isoleucine residues, allowing one to ultimately deduce the assignment based on the known structure (Xu, Liu et al. 2009). This strategy failed however, most likely because the labelling strategy only leads to ~50% incorporation of ^{13}C into isoleucine residues. This is problematic

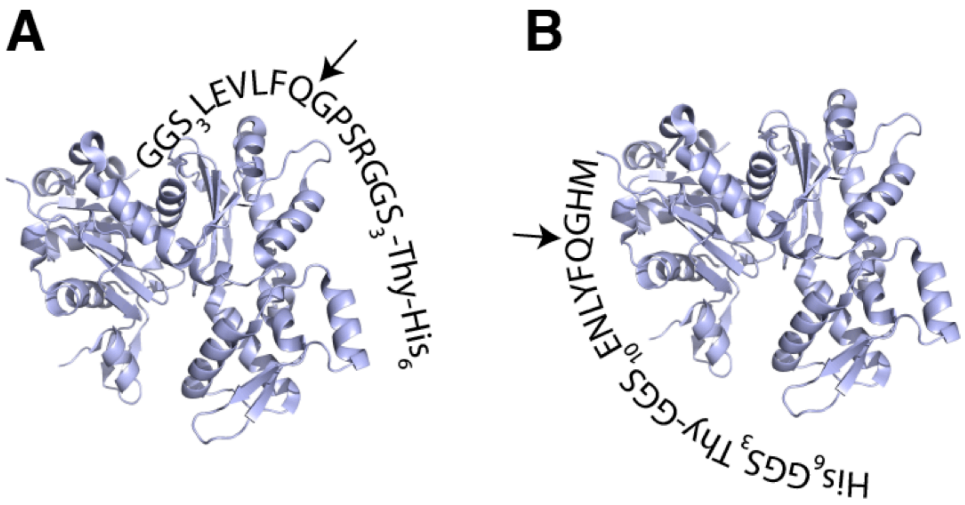


Fig. 3-1. Schematic representation of the two actin constructs used in this chapter. **A.** Schematic representation of the Prescission protease cleavable C-termimal thymosin fusion. The text emerges from the actin C-terminus. The protease cleavage site is indicated by an arrow. **B.** Schematic representation of the Tev protease cleavable N-termimal thymosin fusion. The text terminates at actin's N-terminus.

for through-space NMR experiments, because, in a protein that is labelled at a level of only 50%, the probability that two neighboring isoleucine residues are both labelled is only 25%, which results in a substantial loss of signal. My observations were consistent with this; the methyl-NOESY spectra exhibited cross peaks that were sufficiently weak so as to preclude using the technique for assignment. Ultimately, to assign the spectrum, I adopted a brute-force strategy in which I made a series of point mutants in which I substituted each of the 27 isoleucine residues with leucine and sometimes valine. For each mutant, I acquired an NMR spectrum, and observed which peak vanished, thus revealing the peak corresponding to the residue in question. Ultimately, this permitted complete assignment of the HSQC/HMQC spectra. (Fig. 3-2) To determine whether this particular construct was suitable to explore the role of conformational dynamics in controlling polymerization behavior, it was necessary to confirm that this construct, in the absence of the polymerization-blocking mutations, polymerized readily. Of particular concern was the presence of non-native sequences at the C-terminus; previous work implicates actin's C-terminus as being allosterically coupled to nucleotide state (Mossakowska, Moraczewska et al. 1993, Strzelecka-Golaszewska, Mossakowska et al. 1995), and chemical modification of the C-terminus results in material that is refractory to polymerization (Otterbein, Graceffa et al. 2001). Upon polymerizing the material and pelleting the resulting filaments by ultracentrifugation, I observed that the amount of material remaining in the

supernatant increased with increasing actin concentration (Fig 3-3). In material that exhibits a critical concentration, one expects to observe that the amount of free actin remaining in the supernatant to be equal to the critical concentration, regardless of the total amount of material in the reaction. Thus, the concentration dependence of the supernatant material suggests that the protein assembles without a sharp critical concentration. One explanation for the aberrant behavior of this construct is that it forms unstable filaments that on average are much shorter than canonical filaments. This behavior is reminiscent of that of actin actin from the parasite *Toxoplasma gondii*, which also forms short, unstable filaments that are necessary for the parasite to invade host cells (Skillman, Diraviyam et al. 2011). The unusual behavior of *Toxo* actin is due to "isodesmic" polymerization, which arises when the nucleation and elongation steps share the same rate constants, leading to a much shorter average filament length, and the absence of a sharp critical concentration (Skillman, Ma et al. 2013). Giving rise to this phenomenon in *Toxo* actin, are mutations that destabilize contacts among subunits within the filament. While we didn't rigorously characterize the actin with the C-terminal cleavage remnant, some of its properties appear to be consistent with isodesmic polymerization. Visualization of these filaments by negative stain electron microscopy revealed (admittedly, without rigorous quantification) that the filaments indeed appeared to be shorter than, and morphologically distinct from

those that arise from polymerization of rabbit skeletal actin, and appear to form non-filamentous aggregates (Fig. 3-4).

To solve this problem, I generated another construct, also fused to β-thymosin, but, in this case, to the N-terminus rather than the C-terminus (Fig. 3-1). The C-terminus of this construct is that of native actin. Owing to the directionality of the TEV cleavage sequence, treatment of the N-terminal β-thymosin fusion with TEV (Fig. 3-1) leaves behind a nearly native N-terminus with only an additional glycine-histidine-methionine peptide preceding the actin sequence. The presence of non-native sequences at actin's N-terminus does not tend to interfere significantly with polymerization behavior (Verkhusha, Shavlovsky et al. 2003). Following TEV cleavage and separation from β-thymosin chromatographically using anion exchange chromatography, followed by treatment with 1M NaCl to remove the thymosin, followed by a Ni-NTA affinity chromatography to capture the liberated thymosin. The resulting material, lacking the polymerization-blocking mutations, polymerized much more readily and completely (Fig. 3-4) than the previous construct. Furthermore, this material yielded HSQC/HMQC spectra that are sufficiently similar to those of the C-terminally fused construct, that it was unnecessary to reassign the spectrum (Fig 3-6A). Both of these samples are were bound to Ca-ATP. Interestingly, chemical shift differences that arise from removing the C-terminal prescission protease cleavage remnant

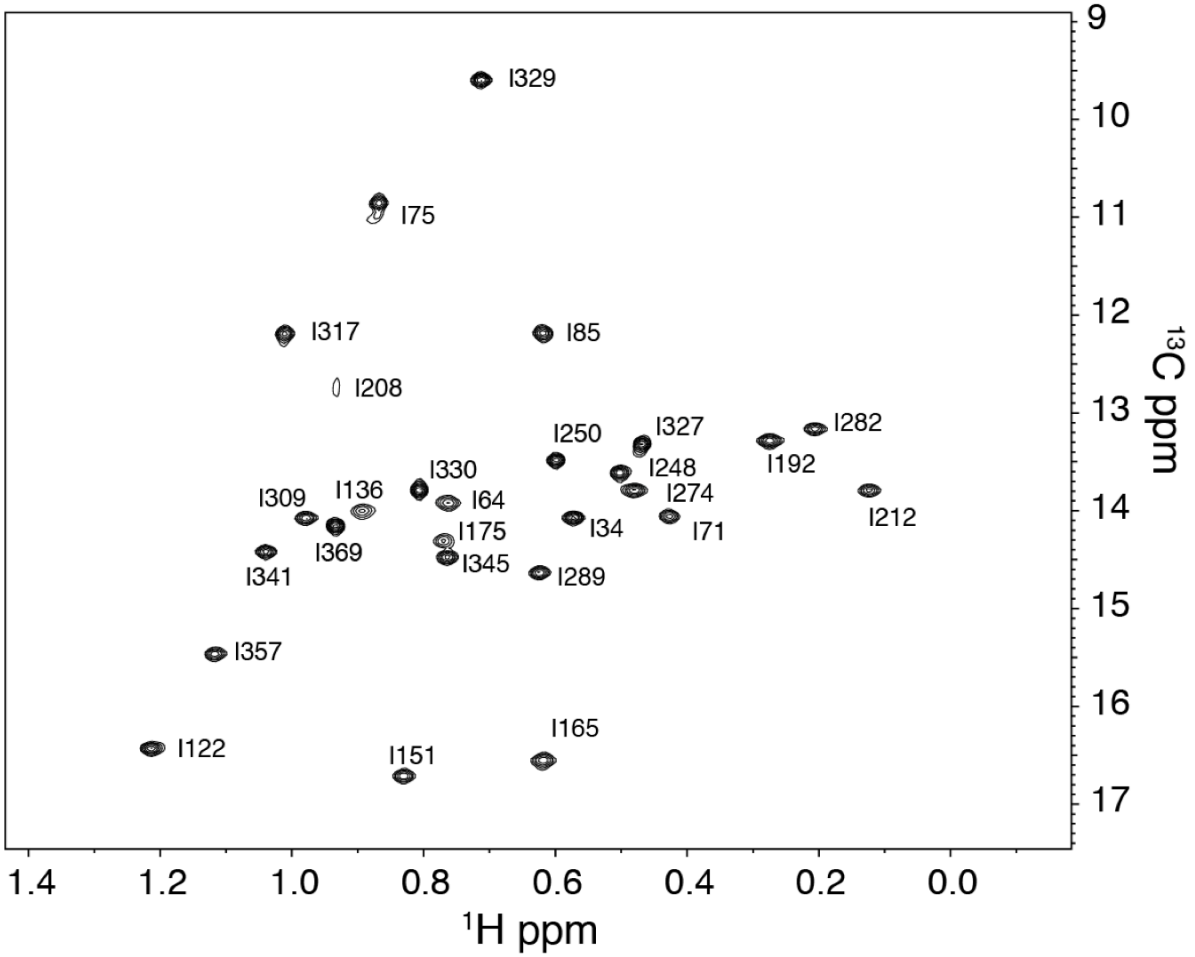


Fig. 3-2 Assigned spectrum of the actin construct containing a C-terminal cleavage remnant from the removal of β-thymosin with prescission protease. Peak assignments, arrived at through serial mutagenesis, are indicated. There are 27 peaks present in the spectrum corresponding the 27 isoleucine residues in the construct.

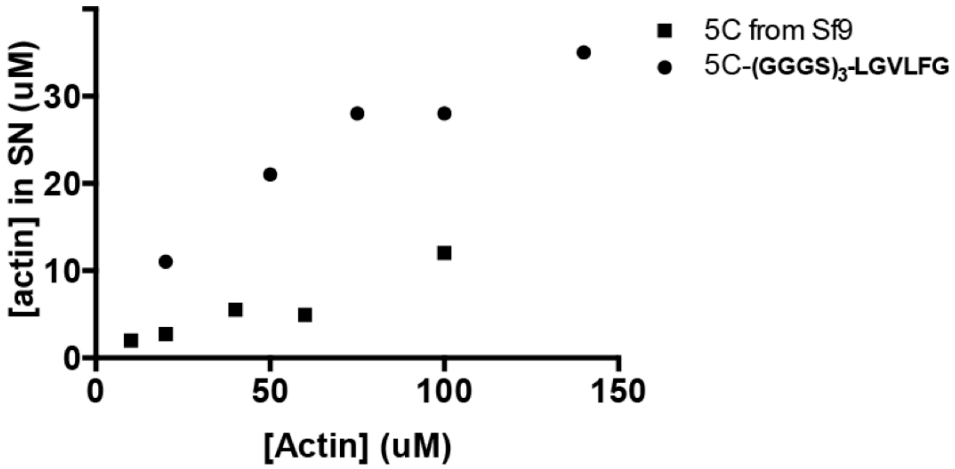


Fig. 3-3 Comparison of the polymerization behavior of 5C actin with a C-terminal cleavage remnant to 5C actin with a native sequence purified from SF9 cells. On the x-axis is the starting concentration of the solution of actin that was ultimately polymerized. On the Y axis is the concentration of actin remaining in the supernatant following overnight polymerization and ultracentrifugation. The C-terminally modified 5C appears as circles, the native actin from SF9 appears as squares.

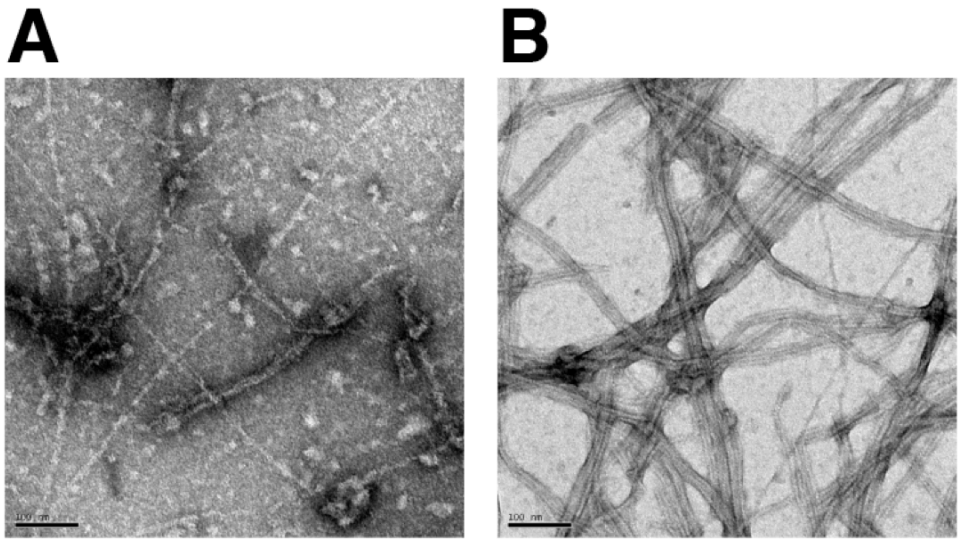


Fig. 3-4. Negative stain EM images of filaments of 5C actin with the C-terminal cleavage remnant to rabbit skeletal muscle actin filaments. A. Negative stain EM image of 5C actin with the C-terminal β-thymosin cleavage remnant. **B.** Negative stain EM image of rabbit skeletal muscle actin.

arise not only in peaks corresponding to residues near the C-terminus, but extend to isoleucine residues in each of actin's four subdomains. This observation supports biochemical data that implicate the C-terminus as being allosterically coupled to distant regions of the molecule (Mossakowska, Moraczewska et al. 1993, Strzelecka-Golaszewska, Mossakowska et al. 1995), and potentially provide an alternate explanation for why C-terminal modification with tetramethylrhodamine interferes so severely with polymerization (Otterbein, Graceffa et al. 2001), beyond merely occluding the subdomain 1,3 cleft. Further supporting this idea, a subset of the residues exhibiting significant chemical shift changes lie in regions of the molecule thought to sense nucleotide state and be in involved in filament formation (Fig 3-5B). These residues include I64 in the D-loop, I165 and 175 in the in the W-loop, I136 in the 1-3-cleft, and I357 in the C-terminus (Fig. 3-6B). Taken together, these data suggest that chemical modification of the C-terminus does more than simply sterically block D-loop engagement; it likely alters actin's conformational landscape in a way that may also affect the G-to-F transition. The involvement of both I64 and I208 on the pointed end of the molecule might indicate that perturbation of the C-terminus results in altered interdomain motions between subdomains 2 and 4.

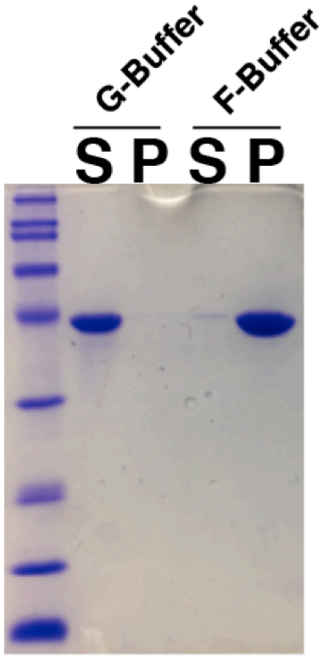


Fig. 3-5 The Construct lacking the C-terminal cleavage remnant polymerizes efficiently. 10 μM actin was exposed to F-buffer and polymerized for four hours, after which the filaments were pelleted by ultracentrifugation in a TLA.120 rotor at 55,000 RPM.

Actin undergoes global conformational changes between the ATP- and ADP-bound states

Preparation of ATP and ADP-bound actin samples with the same divalent cation

The relative binding affinities of actin for ATP and ADP depend strongly on the identity of the bound divalent cation (Sheterline, Clayton et al. 1995). In the presence of Ca^{2+}, Actin exhibits an ~100-fold preference for ATP over ADP, with K_d values of 0.12 nM and 12 nM, respectively. In the presence of Mg^{2+}, the rank order reverses, with actin exhibiting a 4-fold preference for ADP over ATP, with affinities of 0.3 nM and 1.2 nM, respectively. To make meaningful comparisons between ATP and ADP-bound actin samples, the solution conditions should be identical, with both containing the same divalent cation. Because of the metal-dependent differences in affinity, achieving this requirement presented more of a challenge than one might have expected. All of the above actin samples on which we acquired NMR spectra were bound to Ca-ATP. The established protocol for converting actin's nucleotide state from ATP to ADP is well-established and relies on the reversal of affinity that occurs in the presence of Mg^{2+}, coupled with the action of hexokinase, which converts any ATP free in solution to ADP. Given that this procedure produces Mg- ADP actin, we attempted to generate Mg-ATP actin

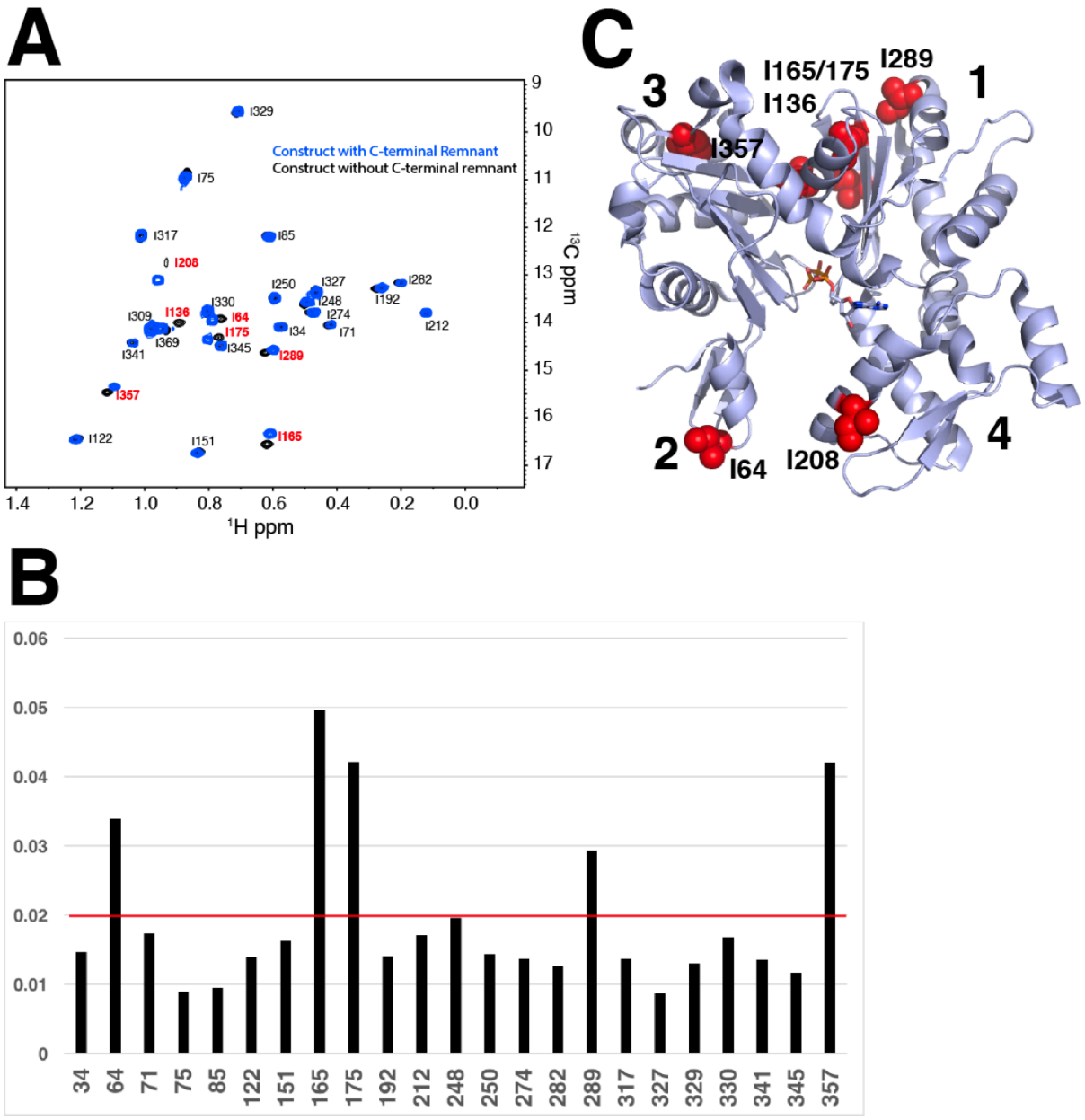


Fig. 3-6 Perturbation of the actin C-terminus causes global chemical shift changes likely reflecting structural differences. A. Overlaid spectra of Ca-ATP actin which contains the C-terminal cleavage remnant (blue), and Ca-ATP actin with a native C-terminus (black). Residues that correspond to peaks that differ substantially from one construct to the other are indicated with bold, red labels. **B.** Chemical shift differences observed between corresponding peaks in arising from the two different constructs, calculated as described previously (Williamson 2014). Peaks above the cutoff of 0.02 (indicated by a red line) were considered significant.

C. residues exhibiting chemical shift differences above the threshold from **B** mapped onto the structure

samples in order to make the direct comparison. In doing so, we found that Mg-ATP actinslowly hydrolyzes ATP, at a rate that is substantial relative to the timeframe of the NMR experiments. This is consistent with previously published data for AP actin (Rould, Wan et al. 2006). To circumvent this problem, we attempted to load actin with the hydrolysis-resistant ATP analog AMP-PNP. But upon doing so, we found that the resulting material hydrolyzed the AMP-PNP over time, precipitated, and underwent proteolysis. So, absent the possibility of comparing ATP and ADP bound samples in the presence of Mg^{2+}, I attempted to generate samples bound to Ca^{2+}. Ca^{2+}-bound ATP actin remains stable for weeks at 4 °C) and was suitable for the experiments I was proposing. However, generating Ca^{2+}-bound ADP samples proved to be a formidable task.

As mentioned previously, the procedure for switching nucleotide relies on the reversal of relative binding affinities that occurs upon changing the bound cation from Ca^{2+}, which favors ATP, to Mg^{2+}, which favors ADP. Thus, the starting and final buffer conditions each favor the most stable ligand pair, Ca-ATP at the start and Mg-ADP at the finish. But this would not be the case if generating Ca-ADP-actin. My initial strategy was to convert Ca^{2+} ATP actin to Mg^{2+} ADP actin using the established protocol, and through extensive dialysis or size-exclusion chromatography, replace the divalent metal to yield Ca^{2+}-ADP actin. Whenever I attempted to do this, however, *even when I knew I was starting with ADP-actin*, by the end of the procedure the system had reverted to ATP-actin according to both

NMR and chromatographic assays. After weeks of struggling (and hypothesizing about potential ATP synthesis by actin!), it turned out that this phenomenon was due to the presence of trace amounts of ATP present in the commercial ADP preparations. This ATP was back-loading onto the protein during dialysis or chromatography in the presence of calcium, due to the 100-fold higher affinity of Ca^{2+}-actin for ATP over ADP.

To overcome this hurdle, I attempted to use hexokinase to remove traces of contaminating ATP from the commercial ADP stocks. While this enabled me to generate Ca-ADP-actin, I ran into a new problem. Over the course of a few days, the NMR spectra of Ca-ADP-actin began to show peaks resembling those of Ca-ATP-actin. This precluded the possibility of acquiring meaningful CPMG data on these samples. Here, a likely explanation for the apparent back conversion is that the ADP degrades slowly at room temperature, which produces free phosphate. In the presence of Ca^{2+}, free phosphate likely occupies the γ-phosphate binding site available in ADP-actin, yielding ADP-Pi actin, which is isoelectronic with ATP-actin, and has biochemical properties closely resembling ATP-actin, and is likely to yield a spectrum nearly identical to that of Ca^{2+}-ATP actin. The solution to this problem ended up being to substantially reduce the concentration of ADP in the final dialysis step, wherein Mg^{2+} ADP actin converts to Ca^{2+} ADP actin, from 200 μM to 200 nM. Under these conditions, the resulting NMR sample remains stable

for many days at room temperature, without significant back conversion to ATP or ADP-Pi, permitting the acquisition of CPMG data.

NMR spectra of actin in the ATP and ADP states

With stable Ca-ATP-actin and Ca-ADP-actin samples in hand, Rustam and I proceeded to NMR analyses. Comparing the HMQC spectra of ^{2}H-actin bound to Ca^{2+}-ATP and that bound to Ca^{2+}-ADP reveals chemical shift changes in peaks corresponding to isoleucine residues throughout the molecule, indicating that information about nucleotide state propagates globally (Fig. 3-7). Chemical shift perturbations were quantified as previously described (Williamson 2014), and a threshold was set such that chemical shift changes above 0.02 ppm were considered significant. The residues experiencing the largest chemical shift changes (I71 and I75) cluster near the ATP terminal phosphate, within subdomain 2 near the D-loop, and within the 1,3-cleft (Fig. 3-7). While of significantly smaller magnitude, but still above the threshold of significance (Fig 3-7) there is another subset of chemical shift differences in isoleucine residues throughout subdomains 1 and 3, within the 1,3-cleft, and propagating to the D-loop. Interestingly residues in subdomain four fall below the threshold, suggesting that it couple less to nucleotide state than the other three domains. It is compelling that the chemical shift perturbations propagate from the nucleotide binding cleft, to the D-loop, and to the W-loop and throughout the 1,3-cleft, suggesting that the state of bound

nucleotide is allosterically coupled to these regions of the molecule important for filament formation.

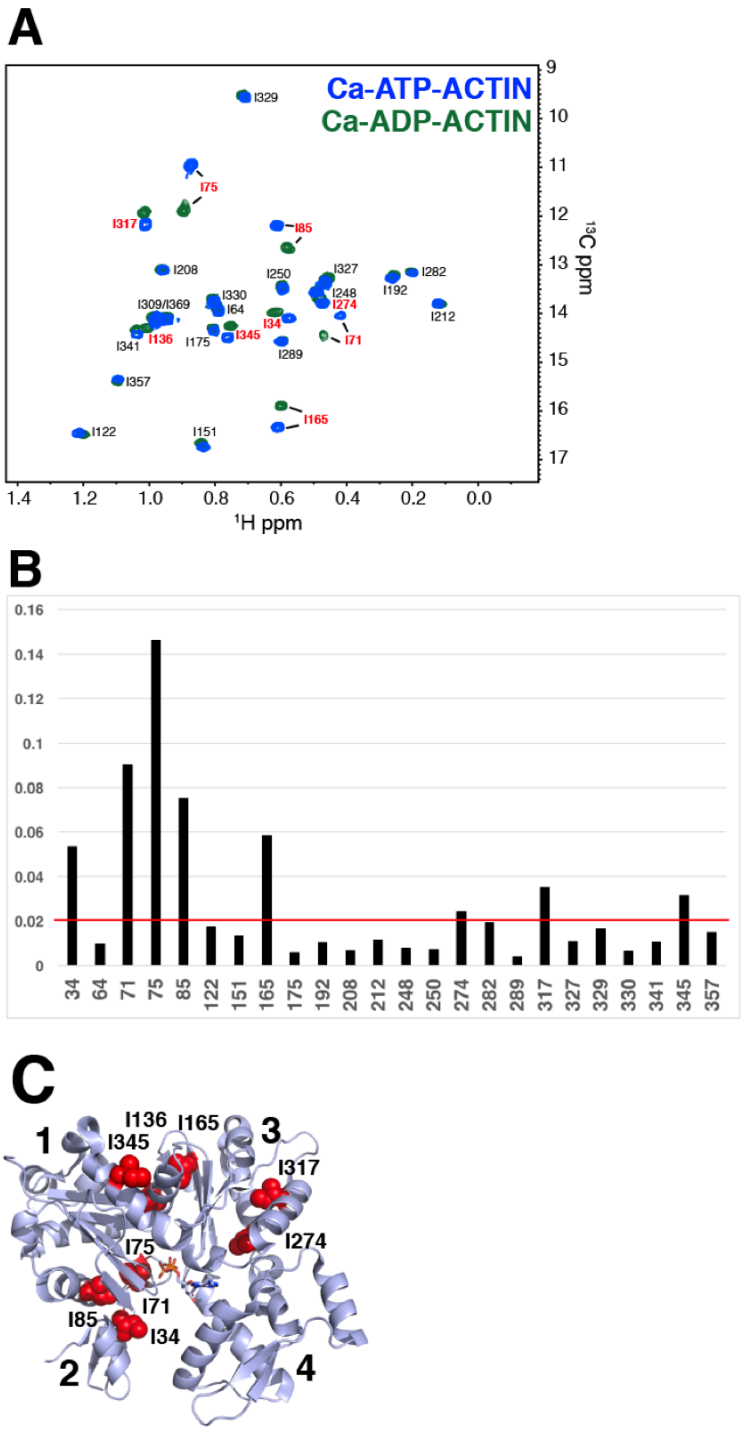


Fig. 3-7 The state of bound nucleotide dictates actins global conformational state. A. HMQC spectra of actin vound to Ca-ATP (shown in black) overlaid with

that of actin bound to ADP (shown in green), both in identical solution conditions. **B.** Labels in bold, red typeface indicate residues that exhibit significant changes in chemical shift upon conversion from ATP to ADP. **B.** Chemical shift differences observed between corresponding peaks arising from the two different nucleotide states, calculated as described previously (Williamson 2014). The red line appears at a value of 0.02. Chemical shift changes above this cutoff were deemed significant and appear in **C.**

CPMG relaxation dispersion experiments to probe dynamics of the ATP and ADP states

The CPMG relaxation dispersion experiment yields information about chemical exchange processes occurring in proteins with interconversion rates on the microsecond to millisecond time scale. It can provide information about conformational states that are poorly-populated and difficult to detect using other methods. The method relies of the fact that chemical exchange processes contribute significantly to T2 relaxation, and thus affect the intensity and linewidths of peaks in an NMR spectrum. T2 relaxation can be thought of as a "dephasing" of magnetization in the transverse plane, and one can imagine the dissipating signal as arising from a series of vectors rotating in the transverse plane, each with a slightly different frequency. This results in the vectors "fanning out," bringing about an ever-increasing dissipation in the NMR signal. Using a 180° pulse about either the x or y axes, one can rotate the "fanned out" signal in a manner that the faster-moving vectors fall to the back of the fan, and the slower-moving vectors move to being in the lead. Now, as the vectors continue rotate, the fast ones will catch up with the slow ones, resulting in refocusing of the NMR signal. This manipulation is known as a "spin echo," or a "Hahn echo," the latter name paying homage to its discoverer. If all vectors rotate at the same rate before and after the 180° pulse, they will refocus exactly together. But if exchange processes are present in a protein, they cause the fanning vectors to interconvert as they fan out,

leading to a configuration that refocuses less efficiently upon application of a spin echo. However, if one applies a train of spin echos, and does so at a frequency (ν_{CPMG}) that far exceeds that of the exchange process in question, the vectors will refocus before they have time to interconvert, and the overall signal will refocus more efficiently. Intuitively then, it is possible to quantify the exchange contribution to T2 relaxation by applying spin echo trains of increasing frequency, and measuring the corresponding increase in signal intensity as the pulsing rate meets and then exceeds the chemical exchange rate. In an actual CPMG experiment, for each peak in the NMR spectrum, one plots the relaxation rate (which is inversely related to the signal intensity) as a function of CPMG pulse frequency, and if exchange processes are present, one observes a decaying series of points. For simple exchange processes involving 2-state equilibria, such as one might expect to see in cases where inter-domain movements lead to discrete "open" and "closed" states, one expects that the CPMG curves, each one corresponding to a peak in the spectrum, are amenable to a global fit with a common set of rate and population parameters:

$$R_2 = f(\nu_{CPMG}, k_{AB}, k_{BA}, \Delta\omega_{AB}, R_2^0)$$

where ν_{CPMG} is the frequency of the pulse train, k_{AB} and k_{BA} are the forward and reverse rate constants in the exchange process, $\Delta\omega_{AB}$ is the chemical shift difference between the two states, and R_2^0 is the intrinsic T2 relaxation rate (in the

absence of exchange contributions). The actual equation one fits to derive these parameters is the "Carver and Richards equation":

$$R_2(\nu_{CPMG}) = \frac{1}{2}\{R_{2A} + R_{2\ddot{v}} + k_a + k_b - 2\nu_{CPMG} cosh^{-1}[D_+ \cosh(\eta_+) - D_- cos(\eta_-)]\}$$

where

$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\psi + 2\Delta\omega^2}{(\omega^2 + \xi^2)^{1/2}}\right]$$

$$\eta_\pm = \frac{\sqrt{2}}{4\upsilon_{CPMG}}\left[\pm\psi + (\omega^2 + \xi^2)^{1/2}\right]^{1/2}$$

$$\psi = (R_{2A} - R_{2B} + k_A - k_B)^2 - \Delta\omega^2 + 4k_A k_B$$

$$\xi = 2\Delta\omega(R_{2A} - R_{2B} - k_A + k_B)$$

By fitting the CPMG data globally at two field strengths, one arrives at the following quantitative information about the exchange process in question: 1. The difference in chemical shift between the two states, $\Delta\omega_{AB}$, 2. The forward and reverse exchange rate between the two states k_{AB}, k_{BA}, which then permit calculation of the populations of state A and state B.

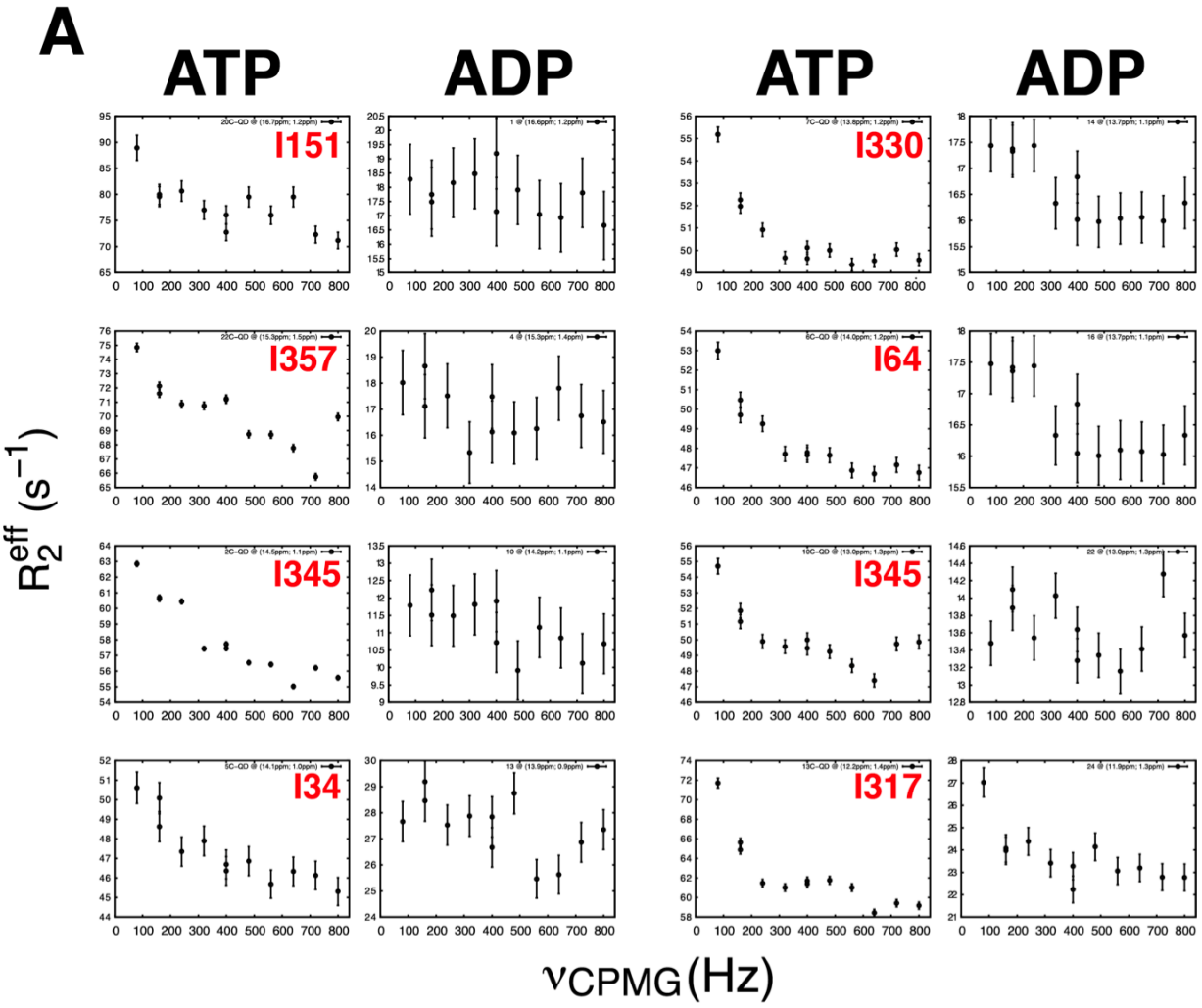
One of the ultimate goals of this work is to apply CPMG experiments to the actin monomer in each of its two nucleotide states. Doing so might reveal that the two nucleotide states exhibit drastic differences in conformational landscape that drive their differential propensity to incorporate into filaments. For several years now, my colleague Rustam and I have struggled to produce actin samples that are bound to the same divalent cation in each of the two nucleotide states, and remain

stable long enough to collect CPMG data. Only in the past month have we overcome these hurdles, and for the first time we have CPMG data of actin bound to Ca-ATP and to Ca-ADP at 600 MHz field strength. The one caveat, at this point, is that the Ca-ADP bound sample used fully deuterated material, and the Ca-ATP sample, fully protonated material. Since the intrinsic T2 relaxation properties (i.e. the non-exchange component of relaxation) of these samples are quite different, the CPMG curves exist on different absolute scales. But since the exchange processes are likely to be similar, one expects that the relaxation dispersion—the relaxation component that changes with pulsing rate—will be unaffected by the presence or absence of deuteration. Thus, assuming that deuteration does not affect the exchange process, a reasonable comparison can be made between the two datasets. The data are preliminary, and, in the near future, we will acquire data on a deuterated Ca-ATP sample and be able to make more exact comparisons. Once our 800 MHz spectrometer with cold probe is back on line (the probe has been offline for several months during renovation of the NMR facility), we will also collect data at a second field, enabling proper fitting of the data and extraction of the exchange parameters.

All this being said, there are aspects of the preliminary data that are compelling. It appears that there are significant differences in the relaxation properties of actin samples bound to different nucleotides. There are residues throughout the actin molecule that, in the ADP state, exhibit little relaxation

dispersion, but, in the ATP state, exhibit significant relaxation (Fig. 3-8). There are also residues that exhibit no relaxation dispersion, indicating that they are not dynamic on μs/ms timescales (data not shown). There is one case in which a residue, I71, exhibited exaggerated relaxation dispersion behavior in the ADP state, relative to the ATP state. This is somewhat unsurprising, as I71 lies in a loop directly adjacent to the ATP terminal phosphate; when ADP is present in the nucleotide-binding cleft there is considerably more space available in which this loop can move. Curiously though, I75, which lies in this same loop, did not exhibit the same behavior, and appears to be more dynamic in the ATP state (Fig 3-7B). Globally, the data suggest that ATP-bound actin is more dynamic on μs-ms timescales than ADP-bound actin. Although we do not yet have enough information to attribute significance to this observation, the possibility exists that the exaggerated dynamics of the ATP state reflect the existence of a poorly-populated excited state that exhibits an augmented propensity to participate in filament formation. As we acquire the necessary data and fit the data globally, we will hopefully be able to characterize this state. The residues that exhibit differential relaxation behavior in different nucleotide states lie in regions of the molecule implicated to be important for inter-subunit contacts within the filament, as described in chapter 1. The residues lie near the D-loop in subdomain 2, and throughout subdomains 1 and 3, lining the 1,3-cleft… The residues in subdomains 1 and 3 contact the C-terminus (I345), and the W-loop (I151), both of which have

been implicated in nucleotide-state driven allostery. Also, I208 in subdomain 4 undergoes nucleotide-dependent changes in relaxation behavior. Because this residue lies there the interface of subdomains 2 and 4, it is possible that this residue is sensitive the an open-closed equilibrium; as discussed in chapter 1, there is tentative evidence that open and closed states are important for filament formation and controlled by nucleotide state. Perhaps these residues (and thus the whole molecule, undergo coordinated motions leading to the population of an excited state that exhibits increased affinity for growing filaments as compared to that of the ground state. To know whether this is true will require global fitting of the data, but at this point it remains a compelling hypothesis.


A
ATP
ADP
ATP
ADP
I151
I330
I357
I64
I345
I345
I34
I317
R_2^{eff} (s^{-1})
ν_{CPMG}(Hz)

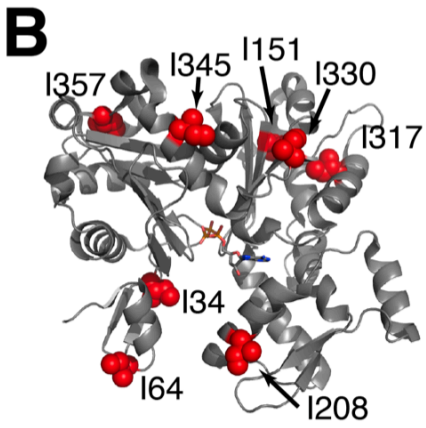

B
I357
I345
I151
I330
I317
I34
I64
I208

Fig. 3-8 In the ATP-bound state, relative to the ADP state, actin exhibits augmented conformational dynamics. A. Relaxation dispersion curves of residues that, when actin is bound to ATP rather than ADP, exhibit enhanced exchange contribution to T2 relaxation, as evidenced by the characteristic decrease in relaxation rate as a function of increasing CPMG pulse frequency. The identities of the residues appear as red text. In the first and second columns are the relaxation dispersion curves of the indicated residues in the ADP and ATP states, respectively. **B.** The residues from **A** mapped onto the actin crystal structure.

Conclusion

Looking ahead, the question remains of how we might characterize this excited state state and examine its potential functions. The only structural information one can gain from a CPMG experiment is chemical shift values, which on their own, do not yield much information about the structure of the excited state. One possibility, if we are lucky, is that the conformation of the excited state closely resembles the conformation of a subunit in the filament. If this is the case, we should be able to use solid-state NMR to arrive at chemical shift values for each of the isoleucine residues in F-actin. We are currently collaborating with Mark Pfuhl's solid-state NMR group at University College London to obtain these F-actin chemical shifts. If the F chemical shifts closely match those of the ATP-bound G excited state, it would lead to an elegant mechanism in which ATP actin visits an excited state conformation resembling that in the actin filament, thus explaining its greater tendency to polymerize.

Another potential hypothesis is that actin exists in a simple two-state equilibrium, and binding of ATP favors one state and ADP the other. In this model, when bound to ATP, the excited state would be ADP-like, and when bound to ADP, the excited state would be ATP-like. Our chemical shift data do indeed indicate that ATP and ADP bound actins lie in different ground states, which supports this hypothesis. But there are other aspects of our NMR data that preclude the

reciprocal excited state relationship. Upon conversion from ATP to ADP, the chemical shift changes that occur in the HMQC spectra are largely in the carbon dimension. If the aforementioned hypothesis were true, the excited state of the ATP state would closely match the ground state of the ADP state, and vice versa. The following observations suggest that this likely is not the case: When we attempt to collect CPMG data using "single-quantum" experiments, we see no relaxation dispersion. The single-quantum experiment only assesses exchange contributions that arise from ^{13}C. If we collect "multiple-quantum" CPMG data, which convolves exchange contributions from both carbon and proton, we observe significant relaxation dispersion. Since the chemical shift differences between ATP and ADP actin lie in the carbon dimension, and because we see no dispersion using single-quantum CPMG experiments, it is unlikely the ADP ground resembles the ATP excited state. Apparently, the excited state chemical shift changes will exist dominantly in the proton dimension.

The work described in this chapter applies modern, multi-resonance NMR techniques for the first time to the actin monomer, and begins to answer long-standing questions about how actin's nucleotide state controls its conformational landscape. We find that perturbing the actin C-terminus brings about chemical shift changes throughout the molecule, including the D-loop, suggesting that this element is coupled to other parts of the molecule that make important inter-subunit contacts in the filament. We find that switching actins nucleotide state also brings

about substantial chemical shift changes, again in residues throughout the molecule shown previously to make inter-subunit contacts in the filament. Finally, we acquire preliminary CPMG relaxation dispersion data on actin in both the ATP and ADP states. The data reveal that, in the ATP state, actin exhibits enhanced relaxation dispersion, particularly in residues surrounding the nucleotide terminal phosphate, those in subdomain 2, near the D-loop, and those lining the 1,3-cleft. This observation suggests that, in the ATP state, these elements of the molecule undergo chemical exchange that is absent or greatly diminished in the ADP state. It remains to be seen whether the motions among these residues are coordinated and whether they are relevant to ATP actins higher propensity to incorporate into an actin filament. Given that they cluster in regions that form inter-subunit contacts and have been shown to be coupled allosterically coupled to nucleotide state, the hypothesis is compelling.

Materials and methods

Actin expression and purification

For the Prescission protease-cleavable C-terminal β-thymosin fusion, 5C actin (residues 1-375), followed by a C-terminal element consisting of $(GGS)_3$-LEVLFQP-$(GGS)_3$-β-thymosin-His_6 was cloned into pICZb using EcoRI and NotI restriction sites. The TEV-protease cleavable, N-terminal β-thymosin fusion consisted of a N-terminal element containing 6xHis-$(GGS)_3$-β-thymosin-$(GGS)_{10}$-

ENLYFQG followed by 5C actin (residues 1-375). For each of these constructs versions with and without the polymerization-blocking mutations were generated. Aside from the use of different proteases, the expression and purification protocols for the two constructs are otherwise identical. Constructs were introduced into *Pichia pastoris* GS115 via electroporation, and the cells plated on increasing concentrations of Zeocin to select for high-expressing clones, as described in the manual provided with the Invitrogen *Pichia* expression kit. High-expressing clones were stored as glycerol stocks that were used repeatedly for subsequent growths. To express the material, an ~60 mL starter culture was grown with vigorous shaking in BMGH (pH 6.0) media over ~36 hours at 30°C to a point of saturation. This starter culture was subsequently used to inoculate 6 X 1L of BMGH media in 4L baffled flasks, and these were grown over the course of ~36 hours to saturation. At this point, the cells were spun down at 3000g and resuspended in 6 X 1L of BMMH media containing no methanol, and containing 100 mg/L methyl-^{13}C 3,3-D_2-α-ketobutyric acid (Cambridge Isotope Laboratories). Cells were incubated with vigorous shaking at 30° for 1 hour, at which time methanol was added to the culture to a final concentration of 0.5% (v/v) to induce expression. Expression was allowed to proceed for 36 hours, at which time the cells were harvested by centrifugation. For deuterated cultures, the expression protocol is similar, except in generating the starter culture, it is necessary to adapt the cells to increasing concentrations of D_2O, as described in Chapter 3. Once the cells adapted to and

grew in 100% D_2O, the larger culture was inoculated. The subsequent steps are identical, except all media contains D_2O in the place of H_2O, and induction is achieved using fully deuterated methanol. Cell pellets were resuspended in 200 mL 50 mM TRIS pH 8.0, 250 mM NaCl, 0.1 mM $CaCl_2$, 0.2 mM ATP and a protease inhibitor cocktail, and stored at -80°C. Cells were thawed and lysed by passing four times through a microfluidizer (Microfluidics M-110P) at 25,000 psi. The crude lysate was next centrifuged at 20,000 rpm in a Sorvall SS-34 rotor. Actin was purified from the cell lysate using Ni-NTA affinity chromatography; the details of which have been described previously (Zahm, Padrick et al. 2013). Following affinity chromatography, the material was treated with either Prescission or TEV protease, depending on the construct, and incubated overnight at 4°C. Following protease cleavage, the material was subjected to anion exchange chromatography on a 4 mL SOURCE 15Q column. This chromatography step removes the majority of the thymosin from the sample, but usually a small amount remains. To remove the remaining contaminating thymosin, the concentration of NaCl in the pooled fractions was increased to 1M, leading to dissociation of 6-His-tagged thymosin from actin. To separate the thymosin from the actin, the eluate was next passed over ~1 mL Ni-NTA agarose equilibrated with 20 mM Tris pH 8.0, 1M NaCl, 0.1 mM $CaCl_2$, and 0.2 mM ATP. The eluate from this step was next subjected to size-exclusion chromatography using a superdex 200 column (GE healthcare) equilibrated with the desired buffer. For the C-terminally tagged construct, the final

gel filtrations step was carried out in 10 mM Hepes pH 7.0, 100 mM KCl, 0.2 mM ATP, 0.1 mM $CaCl_2$, 1 mM TCEP, and 1 mM NaN_3 in 100% D_2O. To generate the Ca-ADP bound N-terminally tagged NMR sample, the final gel filtration step was carried out in 10 mM Hepes pH 7.0, 100 mM KCl, 0.2 mM ATP, 0.1 mM $CaCl_2$, 1 mM TCEP, and 1 mM NaN_3. Fractions containing actin were pooled, and concentrated to ~250 μL. This concentrated sample was treated with 3 $MgCl_2$, 2 mM Dextrose, and 40 U/mL hexokinase, and incubated on ice for 7 hours, resulting in nucleotide exchange. Prior to the subsequent dialysis step, a 100 mM ADP stock (Sigma) was treated with 1 mM MgCl2, 2mM Dextrose, 40 U/ml hexokinase to remove any contaminating ATP. This hexokinase was next used to prepare 500 mL of dialysis buffer containing 10 mM HEPES pH 7.0, 100 mM KCl, 1mM NaN3, 0.2 μM ADP, 20 mM CaCl2, 1mM TCEP, in 100% D2O. The sample, now bound to Ca-ADP, was subjected to NMR experiments.

CHAPTER FOUR

INTRODUCTION TO ACTIN NUCLEATION FACTORS

GENERAL INTRODUCTION

The nucleation of new actin filaments is an intrinsically slow process, with a solution of purified actin requiring up to an hour to polymerize to steady state. This sharply contrasts with what happens in the context of cellular processes, where polymerization occurs on timescales of seconds, with filaments appearing in precisely the correct place, at a specific point in time. Furthermore, in cells actin networks take on diverse morphologies, some comprising a dense meshwork of short, highly branched filaments, while others consist of parallel bundles of long filaments (Campellone and Welch 2010). Actin's intrinsically slow polymerization plays a central role in enabling the cell to exert such exquisite spatial, temporal, and morphological control over actin filament networks.

In the cellular environment, actin rarely, if ever, polymerizes of its own volition; the initiation of a new filament requires the action of nucleation factors. These nucleation factors are generally also subject to control, becoming activated in response to upstream signals. Nucleation factors act to overcome the intrinsic kinetic and thermodynamic barrier that slows generation of new actin filaments. A common theme that has developed through a variety of structural studies is that nucleation factors achieve this activity by arranging actin monomers into a small

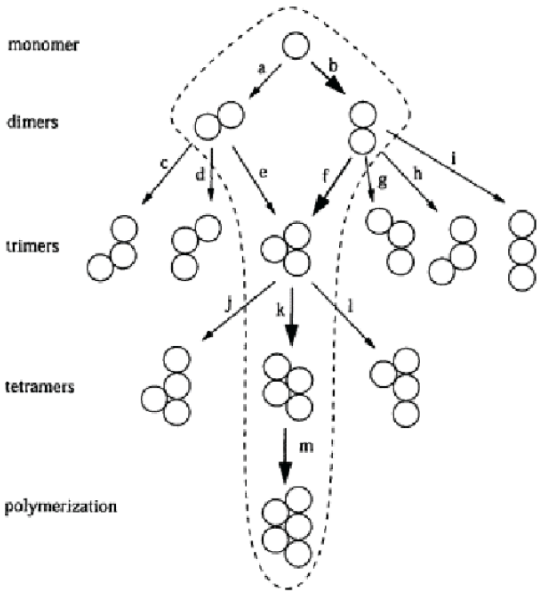
filament-like nucleus that can subsequently undergo rapid elongation. The work in this part of the thesis describes a high-resolution crystal structure of an actin filament nucleus in complex with an actin nucleation factor. Not only was this the first high-resolution snapshot of a nucleation intermediate, but the interpretations provide a general framework through which to understand the action of many nucleation factors.

THE NATURE OF THE ACTIN NUCLEUS

Early kinetic analysis demonstrated that actin polymerization involves a slow lag phase phase and a subsequent rapid growth phase (Wegner and Engel 1975). One explanation for such behavior is that the process is involves an unfavorable nucleation step followed by favorable elongation phase. In such a process, the nucleus consists of a kinetically and thermodynamically unfavorable arrangement of protomers, that, once formed, becomes stable and can readily accept additional monomers. This work from Wegner and Engel also predicted, purely by modeling the kinetics of polymerization, that the minimum number of promoters in an actin nucleus is three. Later, Sept and McCammon came to the same conclusion by carrying out more sophisticated simulations in which they predict the rates and dissociation constants for all of the possible pathways leading to an actin nucleus (Sept and McCammon 2001), and find that their model accurately describes polymerization data. They find that the dominant pathway involves the initial

formation of a long-pitch dimer, followed by the addition of a subsequent promoter, to form a small, filament-like trimer that has both short and long pitch interactions, and can readily accept monomers. They predict that the K_d values of the first, second, and third steps in the pathway are 4.6 M, 0.6 mM, and 0.14 μM, respectively. These predictions corroborate the earlier studies, and establish a likely kinetic pathway through which actin nucleation proceeds. My work, described in this part of the thesis, reveals the structure of an actin nucleus and vindicates these predictions regarding the stoichiometry of nucleation.

A



B

Reaction	ΔA (Å^2)	ΔG_{elec}	ΔG_b	k_- (s^{-1})	k_+ ($\mu M^{-1}\ s^{-1}$)	K_d
(a)	1621	3.54	6.34	5.11×10^{10}	1.23	4.2×10^4 M
(b)	2045	2.74	0.91	1.63×10^8	35.7	4.6 M
(c)	2045	2.65	0.82	4.57×10^7	23.5	1.9 M
(d)	2045	2.93	1.10	7.31×10^7	18.5	4.0 M
(e)	3389	6.58	−9.84	0.80	11.8	0.068 μM
(f)	2948	7.27	−4.42	1.30×10^3	2.18	0.60 mM
(g)	1621	3.64	6.44	5.69×10^{11}	0.99	5.7×10^5 M
(h)	1621	3.68	6.48	6.09×10^{11}	0.77	7.9×10^5 M
(i)	2045	2.85	1.02	3.01×10^8	54.3	5.5 M
(j)	2045	2.87	1.04	1.28×10^8	22.4	5.7 M
(k)	3381	7.00	−9.42	1.51	11.1	0.14 μM
(l)	2045	2.67	0.84	9.37×10^7	22.9	4.1 M
(m)	3381	6.93	−9.49	1.41	11.6	0.12 μM

Fig. 4-1. Calculation of the most favorable nucleation pathway (adapted from Sept and McCammon, 2001). A. Possible pathways to an actin filament nucleus by adding monomers in short-pitch and long-pitch arrangements. The most favorable pathway is circumscribed by a dotted line. Among the two alternatives first steps in the pathway to the trimeric nucleus, the most favorable is marked with a bold "**b**". **B.** The calculated rates and energies for the pathways shown in **A.** ΔA represents buried surface area, ΔG_{elec} the electrostatic component of binding free energy, ΔG_b the total binding free energy, k_- and k_+ the on and off rate constants, respectively.

ACTIN NUCLEATION FACTORS

The Arp2/3 complex

The Arp2/3 complex is a nucleation factor that exists as a constitutive seven-subunit complex, and is nearly as ubiquitous in eukaryotes as is actin (Goley and Welch 2006). In contrast to other nucleation factors that generate stand-alone filaments (see below), the Arp2/3 complex binds to an existing (mother) filament and subsequently nucleates a new (daughter) filament that branches from the side of the mother filament at a characteristic angle of 70° (Mullins and Pollard 1999). The Arp2/3 complex stabilizes and remains associated with the mother filament and the pointed end of the daughter filament following nucleation. Each Arp2/3 nucleation event gives rise to a stable filament branch that can then serve as a substrate for another Arp2/3 mediated nucleation event. Repetition of this process gives rise to the branched filament networks that compose cortical actin networks and lamellipodia, and is involved in many other cellular processes that rely on branched actin networks (Blanchoin, Boujemaa-Paterski et al. 2014).

The Arp2/3 complex is a hetero-heptamer composed of subunits Arp2, Arp3, ArpC1, ArpC2, ArpC3, ArpC4, ArpC5. Knockout studies have shown that the function of Arp2/3 complex requires the presence of all seven subunits (Winter, Choe et al. 1999, Gournier, Goley et al. 2001, Rogers, Wiedemann et al. 2003)

Analysis of the sequences of Arp2 and Arp3 reveal that they are closely related to actin, a feature that is centrally important to the Arp2/3 complex mechanism of action (Nolen, Littlefield et al. 2004).

In the absence of intervention, the Arp2/3 complex is inactive, and in order to nucleate a new filament, it must undergo activation through interaction with a nucleation promoting factor (NPF) (Mullins and Pollard 1999). The canonical NPF is the Wiskott Aldrich Syndrome Protein (WASP). WASP consists of distinct regions: An N-terminal EVH1 domain that binds to the proline-rich domain WIP (WASP interacting protein), a GTPase binding domain (GBD), and an unstructured VCA domain (Kim, Kakalis et al. 2000). The acronym "VCA" describes three distinct regions: an actin-binding **V**-region (**V**erprolin-like; also called a WASP homology 2 domain, or WH2), a **C**entral hydrophobic domain, and the C-terminal **A**cidic region. The WASP V region contains a WH2 domain that binds tightly to the 1,3-cleft of actin with nanomolar affinity, and is responsible for recruiting the initial actin monomers that participate in the nucleation event (Chereau, Kerff et al. 2005, Padrick, Doolittle et al. 2011). The C-domain likely engages Arp2 and Arp3 in their 1,3-clefts in a manner analogous to a WH2 engaging an actin monomer (Boczkowska, Rebowski et al. 2008, Ti, Jurgenson et al. 2011). The A region engages multiple sites on the Arp2/3 complex, providing most of the binding energy between WASP and the assembly (ref Blanchoin, JCB, 2001). Together with mother filament binding, these interactions recruit the first two actin monomers

to the Arp2/3 complex and drive a conformational rearrangement to the active state (see below). WASP generally exists in an autoinhibited state, with the C region of the VCA sequestered by the GBD. Upon interaction with an activated GTPase and other upstream signals, WASP undergoes an allosteric transition that releases the VCA (Kim, Kakalis et al. 2000). Upon this "relief of autoinhibition," the V region of the VCA binds to actin monomers, and thereafter, the two actin-bound VCA domains engage the Arp2/3 complex at two distinct sites. Ultimately, nucleation of a daughter filament requires the association of a 2:2:1 Actin:WASP-VCA:Arp2/3 complex with a mother filament.

The first crystal structure of the Arp2/3 complex revealed the configuration of the seven subunits in the inactive state (Robinson, Turbedsky et al. 2001). As expected, Arp2 and Arp3 adopt folds that closely resemble actin. In the structure, Arp2 and Arp3 sit adjacent to one another, but remain splayed apart. The authors hypothesized that, upon interaction with a VCA domain, the Arp2/3 complex undergoes a conformational change that brings Arp2 and Arp3 into an arrangement that closely resembles the barbed end of an actin filament, which creates a template for addition of new actin monomers. Subsequent work used cryo electron tomography to examine the structure of the branch point that joins mother and daughter filaments (Rouiller, Xu et al. 2008). These studies confirmed the proposed conformational change; in the reconstruction, Arp2 and Arp3 are

aligned to form a short-pitch dimer that closely resembles the filament short-pitch dimer.

Two studies from Shae Padrick revealed that dimerizing VCA domains brings about synergistic enhancement of Arp2/3 activation, indicating that Arp2/3 activation requires not one, but two VCA binding events (Padrick, Cheng et al. 2008, Padrick, Doolittle et al. 2011). This leads to an elegant structural model describing how Arp2/3 complex nucleates filaments: Each of the WH2 domains of the VCA binds to an actin monomer, and the adjacent C-regions bind in an analogous manner to the misaligned Arp2 and Arp3 subunits, tethering together the entities that will ultimately form the first four subunits in the nucleated filament. It is important to note though, that when engaged with a WH2 domain, an actin monomer cannot form a bona fide filament because the WH2 occupies the 1,3-cleft, which engages the D-loop in a filament. So, in order to transition to a filament, the V and the C regions would necessarily have to disengage from actin and Arp2 and Arp3, respectively. The A regions of the two VCAs appear to bind to two distinct patches on the Arp2/3 complex, one on Arp3 (Padrick, Doolittle et al. 2011, Ti, Jurgenson et al. 2011), the other on either on Arp2 or ArpC1 (still unresolved) (Balcer, Daugherty-Clarke et al. 2010, Padrick, Doolittle et al. 2011, Boczkowska, Rebowski et al. 2014). There remains controversy in the field regarding the sequence of events leading to activation of the Arp2/3 complex. There are reports that purport that VCA bind to the Arp2/3 complex is sufficient to bring about the

realignment of Arp2 and Arp3 into an active conformation. One study from Matt Welch's group uses FRET probes, placed in Arp2 and Arp3 to show that CA peptides are sufficient to bring about conformational changes in the complex that correlate with activity (Goley and Welch 2006). Another study shows, using negative stain and cryo EM at ~2 nm resolution, that binding of WASP family proteins brings about repositioning of Arp2 and Arp3 into a configuration that more closely resembles a filament short-pitch dimer (Xu, Rouiller et al. 2012). Another study uses negative stain EM to show that Arp2/3 complex can cap the ends of existing filaments (Volkmann, Page et al. 2014), and that in this context, the Arp2 and Arp3 subunits closely resemble those in the branch reconstruction, suggesting that there is more than one pathway to a state in which Arp2 and Arp3 adopt a flament-like short-pitch arrangement. In interpreting the results of this study, it is important to consider the fact that nucleating a new actin filament is quite different than binding to pre-existing filament pointed ends, and the results are thus not likely relevant to the canonical Arp2/3 mediated nucleation pathway. It is unlikely that the conformational changes revealed in any of these aformentioned studies represent fully activate Arp2/3 complex, as they don't take into account that interaction with a mother filament is required for nucleation. In support of this notion, there is a study in which the authors crosslink the Arp2 and Arp3 subunits together in the presence of bound VCA and find that the resulting material exhibit enhanced activity relative to VCA-bound Arp2/3 in the absence of crosslinker

(Rodnick-Smith, Luan et al. 2016). This suggests that VCA-binding biases the Arp2/3 configuration towards an active state, but that additional interactions are necessary to induce the fully active state.

Two of papers from Jeff Gelles' lab use single-molecule fluorescence microscopy to examine to examine the order of events in the Arp2/3 nucleation pathway. In one study, they find that the Arp2/3 complex, in the absence of VCA, interacts transiently with pre-existing filaments, and that addition of VCA accelerates Arp2/3 binding to pre-existing filaments by ~2 fold; independent of this effect, VCA increased the likelihood that a given Arp2/3 complex bound to a mother filament would nucleate a daughter filament (Smith, Daugherty-Clarke et al. 2013). In a subsequent study, the authors use three-color single molecule fluorescence microscopy to visualize the Arp2/3 compex, actin filaments, and the VCA (Smith, Padrick et al. 2013). They observe that nucleation occurs among Arp2/3 complex bound to both VCA and mother filament and that departure of the VCA precedes filament formation. Perturbations that enhance the off-rate of the VCA stimulate branch formation. This establishes the departure of the VCA as the licensing event for daughter filament nucleation. This calls into question the idea that VCA peptides alone are sufficient to drive the Arp2/3 complex into a nucleation competent configuration, as it is the departure of VCA that initates nucleation. These data make sense structurally: The idea that the WH2 domains and C-regions of WASP block the 1,3 clefts of both actin and the Arp subunits, repectively,

both of which make intersubinit contacts in the daughter filament, the formation of a new branch of a new branch should require that the two VCA domains dissociate from the mother filament Arp2/3 assembly.

An interesting feature of the Arp2/3 complex is that it remains stably associated with the filament it nucleates. Accordingly, in its activated state, the Arp2 and Arp3 subunits rearrange to closely resemble a filament. This contrasts with other nucleation and elongation factors that are processive, carrying out cycles of nucleation or repeatedly assisting in recruiting and placing monomers in a growing filament. For these latter nucleators, binding stably to a filament would arrest the required processivity, thus rendering them non-functional. Such nucleators have been hypothesized to bind to imperfect filament-like structures, but exhibit reduced affinity for the bona fide filament, allowing them to rapidly dissociate from the entities they nucleate. One example of such nucleators are the formin proteins, which can both nucleate actin filaments and assist in their elongation.

Formins

Formins are another class of nucleation factor that operate by a mechanism distinct from that of the Arp2/3 complex. Phillip Leder's lab discovered formins in 1990 using genetic screens, finding them to be important for the "formation" limbs and organs in mice (Mass, Zeller et al. 1990, Woychik, Maas et al. 1990), hence

their name. Formins are ubiquitous in eukaryotes, with many organisms possessing multiple isoforms (fifteen in *homo sapiens*, three in *saccharomyces cerevisiae*) (Chesarone, DuPage et al. 2010). The discovery that Arp2/3 complex nucleates branched filament networks left open the question of how cells nucleate the long, cable-like unbranched actin filaments present in cells (Evangelista, Pruyne et al. 2002, Pruyne, Evangelista et al. 2002). In 2002 Evangelista et.al., answered this question, discovering that the formin proteins nucleate actin filaments *do novo* (i.e. without a required mother filament), and following nucleation, remain associated with the growing barbed end of the filament, a phenomenon termed "processive capping." This continued association protects the growing filament from capping protein and, in the presence of profilin, dramatically increases the filament's elongation rate (Zigmond, Evangelista et al. 2003). The action of formins is important in many cellular processes, including endosomal dynamics (Gasman, Kalaidzidis et al. 2003), formation of filopodia (Pellegrin and Mellor 2005) and lamellipodia (Yang, Czech et al. 2007), and, in yeast, the formation of actin cables and the cytokinetic actin ring (Sagot, Klee et al. 2002).

The defining feature of formins is the presence of an N-terminal FH1 domain and a C-terminal FH2 domain (Chesarone, DuPage et al. 2010). The FH1 domain is unstructured and contains tracts of prolines that are involved in recruiting profilin-bound actin (Kursula, Kursula et al. 2008). The FH2 domain forms a dimer (Xu,

Moseley et al. 2004), and is sufficient for both nucleation and processive capping activities.

Most formins possess additional N-terminal regulatory sequences that function to maintain the formin in an autoinhibited state until it interacts with the appropriate activated GTPase, resulting in relief of autoinhibition. Autoinhibition arises from the interaction a 'diaphanous autoinhibitory domain' (DAD), which lies C-terminal to the FH2 domain, with an element in the N-terminal regulatory domain called the 'Diaphanous inhibitory domain' (DID) (Alberts 2001, Li and Higgs 2005, Otomo, Tomchick et al. 2010). This interaction is thought to prevent nucleation and barbed-end association by sterically blocking the FH2-actin interaction. Activated GTPases bind the DID competitively, displacing the DAD, leading to loss of the intramolecular contacts that give rise to autoinhibition (Otomo, Otomo et al. 2005).

In 2005, Otomo et.al., solved the crystal structure of the formin Bni1p FH2 domain in complex with actin monomers, leading to models of both nucleation and processive capping In the asymmetric unit, a single FH2 domain bound to an actin monomer (Otomo, Tomchick et al. 2005). Although somewhat uninformative at first glance, examination of the crystallographic symmetry reveals more information. Due to a two-fold crystallographic screw axis, one observes pseudo actin filaments in the crystal, decorated with FH2 domains, in which successive actin monomers are related to one another by 180° (as opposed to the 166° in the

canonical filament). While this may at first seem too deviate too far from a real filament to be of any relevance, there are several reports that establish that the actin filament tolerates significant conformational and rotational deviations from ideality (Schmid, Sherman et al. 2004, Galkin, Orlova et al. 2010, Galkin, Orlova et al. 2015, Ngo, Kodera et al. 2015). The FH2 actin interactions along the screw axis lead to an interesting structural model describing both nucleation and processive capping: Each FH2 dimer holds two actin monomers in a pseudo short-pitch arrangement, thereby generating s short-pitch-like dimer that can serve as a template for incoming monomers. The crystallographic symmetry is such that each FH2 dimer contacts three monomers, with successive actins related to one another by 180°. This leads to a compelling model for processive capping: The non-ideal actin trimer, while stabilized by the FH2 domain, experiences strain due to its tendency to adopt an ideal filamentous arrangement in the context of the barbed end. The authors posit that when it does transition from the non-ideal to the ideal arrangement, there is relief of this strain. The FH2 dimer, now being bound to an ideal filament, undergoes a conformational transition that allows it to accept an incoming monomer and, once again, bind to its preferred non-ideal filament. Repeated cycles of this lead to processive capping. In this model, one subunit of the FH2 dimer advances to a position where it can accept an incoming monomer and the cycle repeats. Tom Pollard's lab, in attempting to explain why formin dissociation increases as a function of increasing filament elongation rate, propose

a related model wherein the addition of an incoming monomer precedes translocation of the formin. Subsequently, when the barbed end relaxes to a canonical filament configuration, the FH2 domain adopts an open state and translocates. In this model, an increase in elongation rate increases the time the FH2 domain spends in an open state, facilitating its dissociation (Paul and Pollard 2009). Thinking back to the Arp2/3 complex, in which Arp2 and Arp3 form a perfect short-pitch filament dimer and remain tightly associated with the filaments they nucleate, one begins to envision a trend where deviations from ideality (and consequent strain energy) can potentially underlie the dynamic behavior (or lack thereof) of nucleation factors.

In another interesting parallel to Arp2/3 complex, there are several examples of formins that possess WH2 motifs that bind to actin monomers and are important for nucleation. These include INF2 (Chhabra and Higgs 2006), FMNL3 (Heimsath and Higgs 2012), and mDia1 (Gould, Maiti et al. 2011). The idea is, in a manner analogous to the delivery of actin to Arp2/3 by the WASP WH2 domains, the formin-associated WH2 motifs recruit actin monomers and deliver them to the FH2 domain, assisting in the formation of an actin nucleus that can proceed to an elongating, processively-capped filament. These observations, together hint at the idea that WH2 domains play an accessory role in actin nucleation, generally acting in concert with other proteins, such as Arp2 and Arp3 in the case of the Arp2/3 complex, or an FH2 domain in the case of formins.

WH2-based nucleators

WH2 motifs are pervasive among proteins that regulate the actin cytoskeleton, appearing in the context of Arp2/3 complex in the VCA domains of NPFs, and in some formins as C-terminal elements that deliver actin monomers to the FH2 domain (Paunola, Mattila et al. 2002, Dominguez 2016). WH2 domains are short ~17 amino acid long motifs that bind to actin through a short amphipathic helix, followed by a variable linker region, followed by a characteristic 'LKKT' motif (Dominguez 2007). A number of crystal structures of actin bound to WH2 domains reveal that the amphipathic helix occupies actin's 1,3-cleft, and the LKKT motif binds to the side of the molecule, contacting subdomain 1. These short sequences bind actin with low-nanomolar affinity (Chereau, Kerff et al. 2005). In addition to their integral role in nucleation machinery, WH2 domains are also present in β-thymosin, a protein that sequesters actin monomers, preventing them from participating in filament assembly (Hertzog, van Heijenoort et al. 2004). In this case however, the motif comprises 35 amino acids, with a region C-terminal to the LKKT motif that extends along the face of the actin monomer, terminating in a short helix that packs tightly between subdomains 2 and 4, effectively blocking the pointed end (Irobi, Aguda et al. 2004). In β-thymosin and in all other WH2-containing proteins, the purpose of the WH2 domain is to bind tightly to actin monomers, and its simplicity lends it versatility.

An interesting feature of WH2 domains is that when bound to actin, they occupy a region of actin's 1,3-cleft that, in the context of the filament, engages the D-loop of a neighboring subunit (Oda, Iwasa et al. 2009, Fujii, Iwane et al. 2010, Zahm, Padrick et al. 2013). Thus, proper WH2-binding and filament formation are mutually exclusive, leading to somewhat of a paradox regarding the role of WH2 domains in assembling actin filaments. The solution to this conundrum is evident in Arp2/3-mediated nucleation, wherein the WH2 domains of a WASP family NPF deliver actin monomers to Arp2 and Arp3, and prior to or immediately upon nucleation, they dissociate, permitting the D-loop engagement among subunits in the nascent filament (Smith, Padrick et al. 2013). This ends up being a feature common to all WH2s that participate in nucleation (Dominguez 2016).

In contrast to the WH2s that cooperate with Arp2/3 complex and with the formin FH2 domain, tandem arrays of WH2 motifs exist in other proteins that were, upon their discovery, not thought to cooperate with other proteins, but to be sufficient for nucleation activity . These proteins, called "WH2-based nucleators," because they exhibit some degree of nucleation activity on their own, were thought to be nucleation factors in their own right. These proteins are called SPIRE, Cobl, the bacterial effectors VopL and VopF, and Sca2. In the discussion of each of these proteins it will come apparent that, in a manner analogous to WH2 participation in both Arp2/3 and formin-mediated nucleation, the WH2 domain

serves an accessory recruiting function and is part of a more elaborate nucleation scheme that involves other proteins.

Spire

Spire contains a tandem array of four WH2 domains, and was the first discovered among the WH2-based nucleators, SPIRE was initially identified in a genetic study as being important for *Drosophila* embryo development (Manseau and Schupbach 1989). Subsequent characterization revealed that SPIRE has the capacity to nucleate filaments on its own, albeit with potency far less than for activated Arp2/3 complex (Quinlan, Heuser et al. 2005). This report goes on to show that Spire, by virtue of its tandem actin-binding domains, stabilizes linear arrays of actin monomers. This led to an initial hypothesis that the SPIRE WH2 arrays stabilize four monomers into a single-stranded filament, which subsequently accepts additional monomers to form both long and short pitch contacts, yielding a boa fide filament seed that can elongate. In support of this hypothesis, the authors observed that SPIRE remains associated with the pointed end of the filaments it nucleates. This led the authors to posit that SPIRE represents a new class of nucleation factor whose activity depends solely on tandem arrays of WH2 motifs.

The report that originally identified SPIRE also identifies another gene that encodes a formin, called *Cappuccino*, whose disruption results in a phenotype that exactly mirrors the SPIRE-null phenotype. Following the initial characterization of SPIRE, the same authors discovered that the SPIRE protein interacts directly with the formin *Cappuccino* both in vivo and in vitro. This interaction brings about a decrease in the processive capping activity *Cappuccino*, but enhances the nucleation activity of SPIRE. That dimerizing the WH2 arrays results in enhanced nucleation activity, in retrospect, is unsurprising, given that the productive nucleation pathway involves the formation of both short and long pitch oligomers. In support of the idea that the cooperation between SPIRE and *cappuccino* represents a general theme, it later became apparent that there is an analogous interaction between the mammalian SPIRE orthologs Spir1 and Spir2 with the formins FMNL1 and FMNL2. These observations suggest that SPIRE does not represent a new class of actin nucleator, because, just as is the case with NPFs and formin-associated WH2s, the role of SPIRE WH2 cooperates with additional nucleation machinery.

Cobl

Cobl is another WH2-based nucleator that contains a tandem array of three WH2 domains, and was originally identified as being important for neuronal morphogenesis, and for a variety of other cellular processes. While similar to

SPIRE, in that it possesses a WH2 array, Cobl differs from spire in that the linker between its second and third WH2 domains is long (65 amino acids) relative to the linker between the first and second (17 amino acids). Spire WH2 domains are equally spaced with linkers consisting of 14 amino acids. It was discovered that the cobl WH2 array is sufficient for nucleation activity in vitro, but that this activity required the long linker between the second and third WH2 domains. This led to the hypothesis that the first and second WH2 domains stabilize a long-pitch dimer, and the third WH2 domain, by virtue of the longer linker, can deliver the third actin to generate a trimeric nucleus. A subsequent study however reported that the WH2 array is insufficient for nucleation, and requires the presence of a tract of lysine residues N-terminal to the WH2 array. In the absence of the lysine stretch, the WH2 array sequestered actin monomers, but didn't nucleate filaments. It remains unclear how the lysine motif promotes the nucleation activity of the Cobl WH2 array, although it could involve a redistribution of charge that potentiates the capacity of the first WH2-actin complex to recruit additional actin monomers. Although the biochemistry is still in its infancy, it appears that the activity of Cobl in cells depends on the presence of several other cofactors that dimerize, so a direct interaction would thus dimerize the Cobl WH2 arrays, likely enhancing their nucleation activity (Dominguez 2016). So like Spire, it appears that Cobl requires the presence of accessory proteins to realize its full nucleation potential.

VopL/VopF

Vibrio outer proteins L and F (VopL and VopF, VopL/F hereafter) are type-III secreted effector proteins from the intestinal pathogens *Vibrio Parahaemolyticus Vibrio* and *Vibrio Cholorae,* respectively. Genetic analysis revealed that these proteins share similar domain structures (57% sequence identity), each containing a tandem array of three WH2 domains and three proline-rich motifs. The proteins were thus hypothesized to be actin-regulatory molecules. Indeed, ectopic expression of VopL in Hela cells results in a dramatic reorganization of the actin cytoskeleton into characteristic stress fibers. Furthermore, addition of VopL to purified actin brings about tremendous stimulation of actin nucleation, with a potency unprecedented for other nucleation factors, with concentrations of 5 nM vopL in solution being equivalent to 50 nM WASP-activated Arp2/3. Similar cellular and biochemical data were described for VopF. This led to the conclusion that VopL/F are WH2-based nucleators.

Analysis of the VopL domain structure revealed that, C-terminal to the WH2 array, there is a domain, which is predicted to be folded, that bears no significant homology to other known proteins, named the “VopL-C-terminal domain” (VCD). Upon the discovery that VopL nucleates actin filaments, it was assumed that its nucleation activity arises from the known actin regulatory sequences—the WH2 motifs—and that the VCD likely plays an accessory role, if any at all. Surprisingly, biochemical analysis revealed that this is not the case. On its own, the WH2 array

does not nucleate actin filaments. The isolated VCD however, does exhibit nucleation activity, albeit with significantly less potency than the full-length protein. Only when both WH2 motifs and VCD are fused together in a single chain does the molecule nucleate actin filaments with high potency. Taken together, these data suggest that the core nucleation activity resides in the VCD, and that the presence of the WH2 array confers significant enhancement of this core activity.

The crystal structure of the VCD revealed that the molecule is a U-shaped dimer stabilized by a coiled-coil, consisting of a rigid “base,” and two folded “arms” connected by flexible linkers to either side of the base. The distance between the tip of the arms and the top of the base is similar to the stagger between consecutive monomers in the actin filament. Furthermore, the tips of the arms and the top of the base possess patches of positively charged surface residues that, upon mutation, result in significant impairment in VCD-mediated nucleation. Taken together this led to the hypothesis that the VCD binds directly to actin monomers, templating a filament-like short pitch dimer that can accept additional monomers to yield an elongating actin filament.

One of the goals of my thesis work was to determine the mechanism of VCD-based nucleation, and to understand how the presence of WH2 domains brings about enhancement of the activity intrinsic to the VCD. To achieve this, I solved the crystal structure of the VopL VCD in complex with actin, which is detailed in Chapter 5. In brief, the structure reveals the VCD bound to three actin

monomers in a spatial configuration that approximates that of three consecutive monomers in an actin filament. In this arrangement, the actin monomers could readily accommodate the WH2 domains in their 1,3-clefts, which suggests that the WH2 arrays serve a recruiting function, binding to actin monomers and delivering them to the VCD. The structure represents the first high-resolution snapshot of an actin nucleation intermediate, provides mechanistic insight into the action of other nucleation factors, and has helped to transform our understanding of how WH2 arrays function in systems that nucleate actin filaments.

At the start of my thesis work, the WH2-based nucleators were recently discovered, and the field was still trying to reconcile their functional importance in vivo with their poor nucleation activity in vitro. My work and the work of others resulted in the understanding that WH2 domains, on their own, are generally poor nucleators, but bind actin with high affinity, and are thus well adapted to serving a recruiting function, delivering actin to a low affinity organizing entity. This turns out to be the case with the WH2 domains in the NPFs that activate the Arp2/3 complex, the WH2 domains C-terminal to the FH2 domains of some formins, and the WH2 domains of WH2-based nucleators, which, in all cases discovered, seem to require accessory factors to potentiate their activity.

Another feature of the VopL structure that is relevant to other nucleation factors is that the configuration of the thee monomers in the crystal structure deviates somewhat from that of consecutive monomers in the canonical actin

filament: The rotational and translational positions of the actin monomers differ from the filament, and the D-loops in the structure remain disordered. Furthermore, in full-full length VopL, the 1,3-clefts would engage WH2 motifs, the latter of which would have to dissociate prior to D-loop engagement and transition to a bona fide filament. The complete or partial disengagement of WH2 motifs, coupled with the rotational and translational repositioning of monomers that accompanies transition to a filament would likely result in a decrease in affinity, resulting in dissociation of VopL. This feature is well suited to the biochemical activity of VopL, which is to carry out repeated nucleation events. The Arp2/3 complex, on the other hand, remains stably associated with the filaments it nucleates, anchoring them to the mother filament. Accordingly, upon activation by an NPF, Arp2 and Arp3 realign into a configuration that, to a much higher degree, matches that extant in the canonical filament. Formins, in a manner related to VopL, also exhibit dynamic behavior, likely stabilizing imperfect filament like structures that, upon transition to a filament, drive the FH2 domains processively along the end of the growing filament.

In hindsight, VopL was a good model system to study general features of actin nucleation; the VopL-actin crystal structure encapsulates many of the features which are generally applicable to nucleation mechanisms: First, the structure deviates from ideality, explaining how nucleation factors can tune their dynamic behaviors to suit specific purposes. Second, the structure suggests that

WH2 arrays can synergize with other domains, giving rise to potent nucleation activity. Lastly, the structure represents the first example of a high-resolution snapshot of a nucleation intermediate, and substantiates the long-standing hypothesis that nucleation factors function by arranging filament-like nuclei

CHAPTER FIVE

The Bacterial Effector VopL Organizes Actin into Filament-like Structures[3]

Abstract

VopL is an effector protein from Vibrio parahaemolyticus that nucleates actin filaments. VopL consists of a VopL C-terminal domain (VCD) and an array of three WASP homology 2 (WH2) motifs. Here, we report the crystal structure of the VCD dimer bound to actin. The VCD organizes three actin monomers in a spatial arrangement close to that found in the canonical actin filament. In this arrangement, WH2 motifs can be modeled into the binding site of each actin without steric clashes. The data suggest a mechanism of nucleation wherein VopL creates filament-like structures, organized by the VCD with monomers delivered by the WH2 array, that can template addition of new subunits. Similarities with

[3] This section is an exact copy of a paper, written by Mike Rosen, Shae Padrick and myself, published in Cell in Oct. 2013.

Arp2/3 complex and formin proteins suggest that organization of monomers into filament-like structures is a general and central feature of actin nucleation.

Introduction

Many important cellular processes, including cell motility, vesicle trafficking, and cell division, depend upon precise spatial and temporal control of actin polymerization (Pollard 2007, Dominguez 2009, Pollard and Cooper 2009, Campellone and Welch 2010). Actin can polymerize on its own but does so slowly, primarily due to kinetic barriers that hinder spontaneous nucleation (Pollard and Cooper 1986, Sept and McCammon 2001). Cellular actin nucleation factors accelerate filament formation by catalyzing nucleation in response to upstream regulatory signals. Their actions afford precise spatial and temporal control over actin filament dynamics in vivo (Pollard 2007, Padrick and Rosen 2010) (Padrick and Rosen, 2010, Pollard, 2007). The Arp2/3 complex, formin proteins, and Wiskott-Aldrich syndrome protein (WASP) homology domain 2 (WH2)-based nucleators are ubiquitous eukaryotic actin nucleation factors (Campellone and Welch, 2010). The structural mechanisms by which these systems mediate filament assembly are incompletely understood.

The Arp2/3 complex is a protein assembly comprising seven subunits that contains two actin-related proteins (Arp2 and Arp3), which are structurally similar to actin (Machesky, Cole et al. 1994, Kelleher, Atkinson et al. 1995). The VCA

region of proteins in the WASP family acts in concert with existing actin filaments to activate the Arp2/3 complex; the net result is nucleation of a new filament from the side of an existing one (Pollard 2007). During Arp2/3 activation, the WH2 regions from two VCAs bind to and deliver actin monomers to Arp2 and Arp3 (Padrick, Cheng et al. 2008, Padrick, Doolittle et al. 2011, Ti and Pollard 2011). Crystal structures of inactive Arp2/3 complex and EM analyses of the active form have shown that nucleation also involves substantial reorganization of the two Arp subunits to an arrangement that resembles successive “short-pitch” monomers in an actin filament (Robinson, Turbedsky et al. 2001, Xu, Moseley et al. 2004, Rodal, Sokolova et al. 2005, Nolen and Pollard 2007, Rouiller, Xu et al. 2008, Padrick, Doolittle et al. 2011). Nucleation thus appears to be based on an arrangement of the Arp subunits and recruited actins that mirrors the “barbed end” (or rapidly growing end) of the polarized actin filament, which readily incorporates additional monomers.

Formin proteins also act by recruiting and organizing actin monomers. These proteins nucleate filaments through a conserved formin homology 2 (FH2) domain, which tracks processively with the growing barbed end of the nascent polymer (Paul and Pollard 2009). Crystal structures of formin-actin complexes indicate that the FH2 domain arranges monomers in a conformation that resembles a strained actin filament, leading to models of both nucleation and processive elongation (Otomo, Otomo et al. 2005, Paul and Pollard 2009,

Thompson, Heimsath et al. 2013) (Otomo et al., 2005, Paul and Pollard, 2009, Thompson et al., 2013). In some formins, the FH2 domain acts in concert with sequence motifs proximal to or overlapping with an adjacent regulatory element (the DAD motif). These sequences, which appear to be related to the WH2 motif, can accelerate nucleation and are thought to deliver actin to the FH2 domain (Chhabra and Higgs 2006, Gould, Maiti et al. 2011, Heimsath and Higgs 2012).

The WH2-based nucleation factors are defined by arrays of WH2 motifs. Well-studied examples include cordon-bleu (cobl), leiomodin (lmod), and SPIRE (Qualmann and Kessels 2009) (Qualmann and Kessels, 2009). Members of this class vary in the number of WH2 motifs they possess, how these WH2 motifs are positioned relative to one another, and in nucleation potency. In some members (e.g., cobl), WH2 motifs are positioned in a manner that permits stabilization of a short-pitch actin-actin contact, which may be important for efficient nucleation (Qualmann and Kessels 2009, Carlier, Husson et al. 2011). In SPIRE, the arrangement of WH2 domains is more consistent with stabilization of longitudinal actin-actin contacts instead of the short-pitch actin dimer needed to produce a barbed end. This is consistent with EM analyses showing structures resembling a short, single actin strand in the presence of SPIRE as opposed to the pair of strands that compose an actin filament (Quinlan, Heuser et al. 2005). In isolation, the SPIRE WH2 array exhibits relatively weak nucleation activity. But an interaction with the dimeric formin, Cappuccino, brings together two SPIRE WH2 arrays,

greatly enhancing activity (Quinlan, Hilgert et al. 2007, Vizcarra, Kreutz et al. 2011). Thus, although different WH2-based nucleation mechanisms are possible, highest potency appears to involve stabilization of both strands of the nascent filament.

Vibrio parahaemolyticus is a gastrointestinal pathogen and is a cause of food-borne illness worldwide (Yeung and Boor 2004). Transmission occurs primarily through the consumption of raw or undercooked shellfish harvested from contaminated marine waters and results in diarrheal disease that is usually self-limiting (Yeung and Boor, 2004). Vibrio parahaemolyticus, like many bacterial pathogens, hijacks eukaryotic cytoskeletal processes through injection of effector proteins into host cells (Haglund and Welch 2011, Alto and Orth 2012). One of the effector proteins that Vibrio parahaemolyticus injects is the actin nucleation factor VopL. VopL injection causes a substantial reorganization of the host cytoskeleton (Liverman, Cheng et al. 2007), leading to the formation of characteristic actin stress fibers. VopL has two distinct domains: an N-terminal array of three WH2 motifs and a unique VopL C-terminal domain (VCD). Although the VCD is sufficient for nucleation activity, its potency is greatly enhanced by inclusion of the tandem WH2 arrays (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011).

In order to understand the activity of VopL and to gain general insights into structural mechanisms of actin filament nucleation, we determined the crystal structure of the VopL VCD in complex with actin. In this structure, the VCD dimer

binds to three actin monomers that show striking similarity to three consecutive monomers in an actin filament. Modeling shows that, in this arrangement, each actin monomer is accessible to a WH2 domain. The structure and complementary biochemical data lead to a model in which the VCD functions as a low-affinity "organizer," evolved to arrange actins in a filament-like configuration, and the WH2 arrays, although being poor organizers, bind actin monomers with high affinity and deliver them to the VCD. Together, the two elements enable full-length VopL to potently template new actin filaments. The division of organization and delivery appears to be a general feature of actin nucleation factors.

Results

A New Nonpolymerizable Actin Mutant

To understand the mechanism of VopL-mediated nucleation, we determined the crystal structure of the VopL VCD in complex with actin. Crystallographic studies involving actin require some means of preventing polymerization. To achieve this, we introduced three mutations into the barbed end of Drosophila melanogaster 5C actin (D287A, V288A, and D289A). These mutations are located in the interface between actin monomers in a longitudinal contact (along the filament axis) (Fig. 5-1A). Pelleting assays show that the

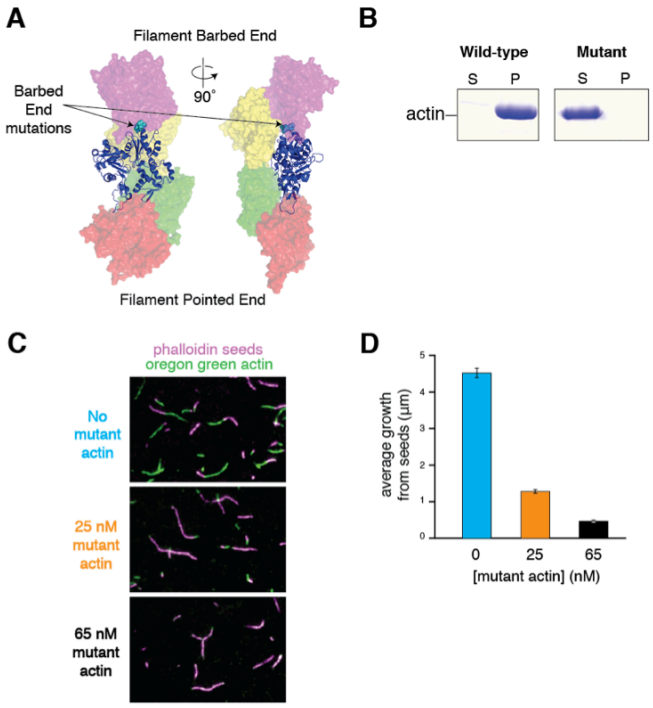


Fig. 5-1. Development of a Barbed-End-Blocked Nonpolymerizable Actin
A. A single actin monomer in an ideal filament (Oda et al., 2009) is shown in ribbon representation, with the four flanking actins shown as transparent surfaces. The locations of the side chains mutated to prevent polymerization (cyan spheres) are shown. **B.** Actin pelleting assays comparing mutant Drosophila 5C actin to wild-type rabbit muscle actin. 9 mM samples of rabbit muscle and barbed-end-blocked actin were allowed to polymerize in 50 mM KCl for 16 hr. Samples were separated by centrifugation into a soluble pool and pellet pool (containing fil- aments) and analyzed by SDS-PAGE. **C.** Mutant actin exhibits capping activity at filament barbed ends. Filament seeds (stained with Alexa-546 phalloidin) were mixed in G-buffer with 0.5 mM unlabeled rabbit actin, 0.15 mM rabbit actin labeled with Oregon green maleimide, and either 0 mM, 0.025 mM, or 0.065 mM mutant actin. Polymerization was initiated by addition of 103 KMEI buffer and allowed to proceed for 10 min, at which time individual filaments were imaged by TIRF microscopy.
D. Average length of growth from the barbed ends of phalloidin seeds in the presence of increasing concentrations of mutant actin. Values shown are average filament lengths growing from the barbed end of phalloidin stabilized seeds, with 423, 275, and 70 seeds analyzed at 0, 25, and 65 nM mutant actin, respectively. Error bars represent the SE in filament length. Average value for the 65 nM mutant actin sample is likely an overestimate as many seeds did not grow visibly at either end and hence were rejected as barbed and pointed ends could not be distinguished from one another.

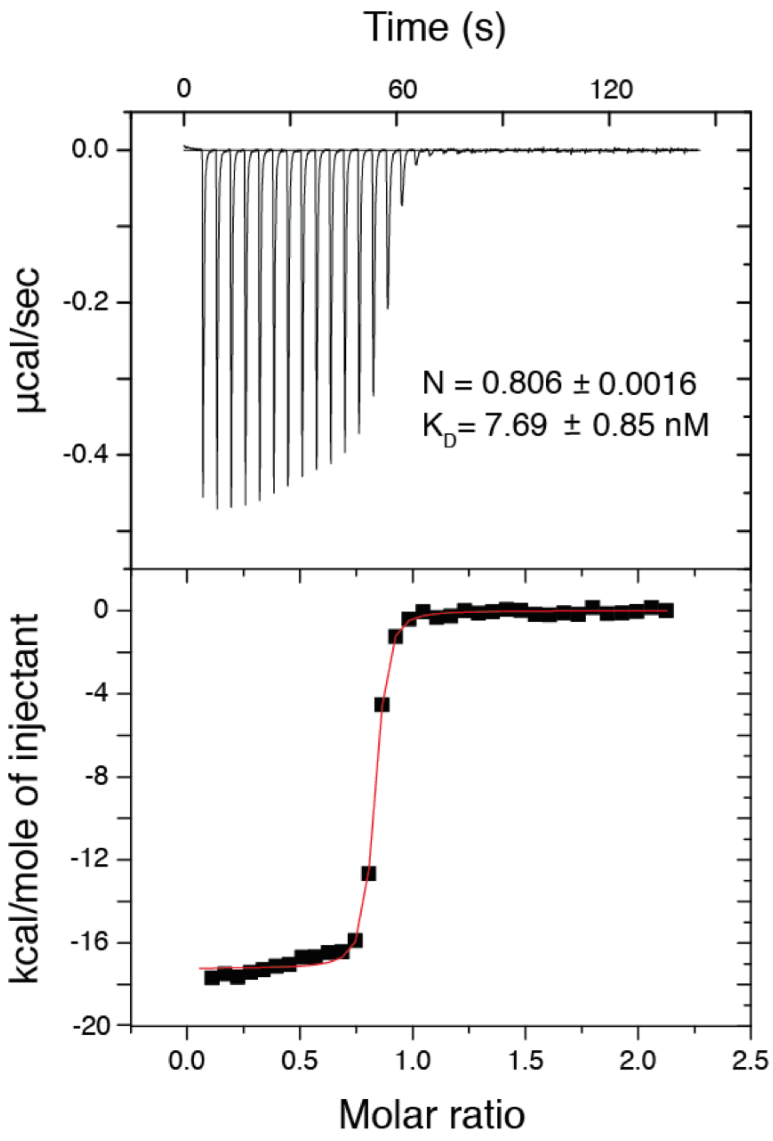


Figure 5-2. Assessment of WH2 Binding to DVD Actin by Isothermal Titration Calorimetry

Titration performed in KMEI buffer supplemented with 0.2 mM ATP. Upper and lower panels show raw and integrated heats, respectively. Red line in the lower panel represents the best fit to a one-site binding model. The resulting stoichiometry and Kd values appear in the top panel. The departure in the stoichiometry from the expected value of 1 is likely due to uncertainty in the concentration of the WH2 peptide.

D287A/V288A/D289A mutant actin does not form filaments under conditions that induce polymerization of wild-type actin (Fig. 5-1B). Although unable to polymerize, the mutant retains properties associated with wild-type actin; it binds VopL WH2 motifs with nanomolar affinity (Fig. 5-2) and is able to interact with the fast-growing barbed

Structure of the VCD-Actin Complex

Co-crystallization of the mutant actin and VCD yielded crystals that contained both components and diffracted to 2.75 Å, enabling structure determination by molecular replacement (Figs. 5-3 and 5-4B, and table 5-1; PBD ID 4M63). In this structure, each asymmetric unit contains a VCD dimer and three actin monomers. The actin monomers adopt a spatial arrangement that resembles that of three successive subunits in an actin filament (Fig.5-2 and 5-4B). Given the many possible arrangements actin might adopt in an arbitrary crystal, it is likely that this resemblance is functionally significant. As in free VopL, each VCD monomer consists of three structural units: arm, base, and carboxy-terminal helix (Figs. 5-5 and 5-6) (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011). Contacts between the base units form an elongated platform, stabilized by the C-terminal helices, which form a coiled-coil. The arms emerge from the ends of this platform and are directed opposite to the coiled coil. In both the free protein and

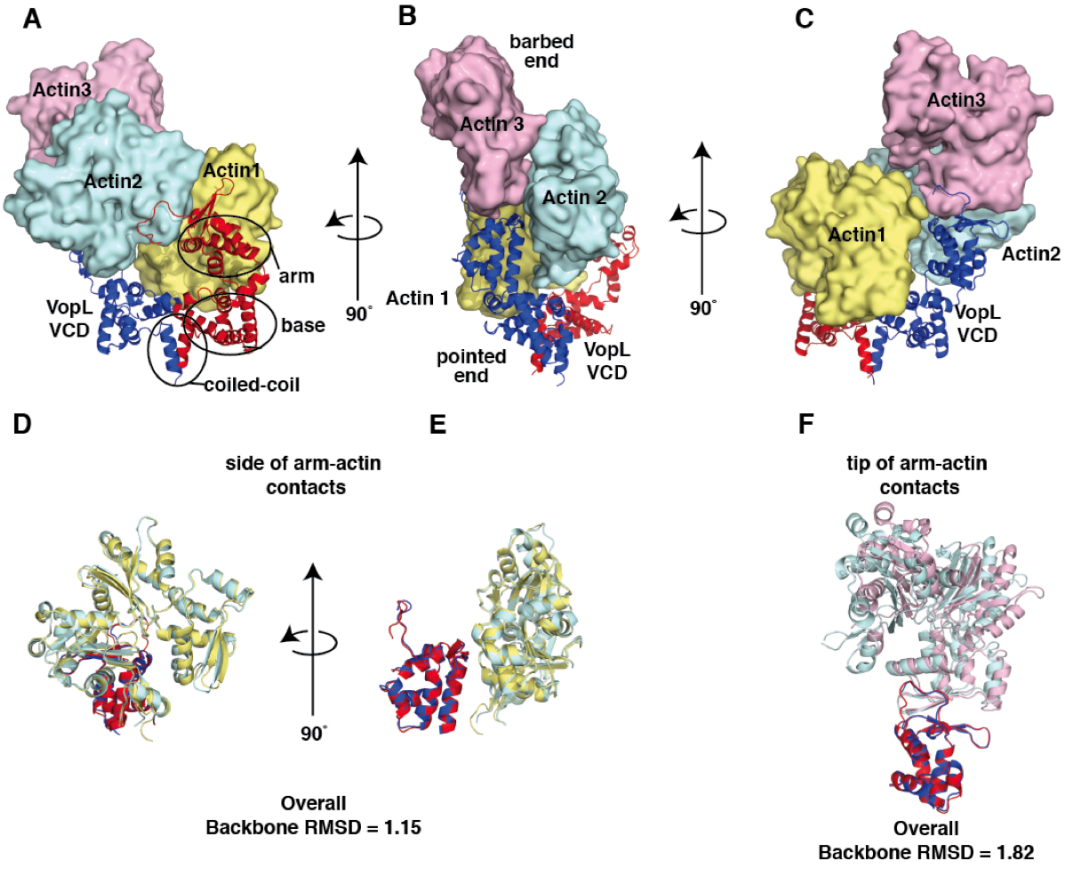


Fig. 5-3. Overall Structure of the VopL VCD in Complex with an Actin Trimer A-C. VopL VCD dimer is shown as a ribbon with one VopL monomer (chain A) red and the other (chain B) blue. Actin monomers 1, 2, and 3 are shown in surface representation and are colored yellow, cyan, and pink, respectively. **B** indicates barbed and pointed ends of the actin trimer. **D-E.** Side of arm contacts. Alignment of arm B + actin 2 (oriented as in [A]) with arm A + actin 1. **F.** Tip of arm contacts. Alignment of arm B and actin 3 (oriented as in **C**) with arm A + actin 2. The alignments were made using backbone atoms of the VCD arms only. Overall backbone RMSD is the result of aligning atoms from both actin and VCD arms See also Figure 5-4 and Table 5-1.

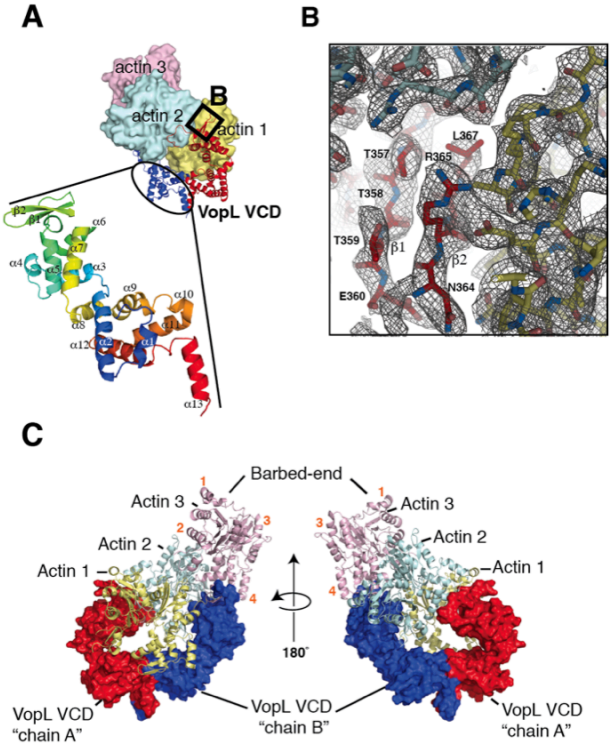


Fig 5-4. Additional Views of the Crystal Structure

A. Overall structure of the VCD-actin complex colored as in Figure 2. Expanded region shows a single VopL VCD monomer labeled with secondary structure elements. α helices and β strands are labeled in order of appearance proceeding from the N terminus to the C terminus. In the magnification, the VCD backbone is colored with N-terminal residues in blue, C-terminal residues in red and intermediate residues spanning the rainbow according to sequence position. **B.** Representative example of 2Fo-Fc electron density contoured at 1s and the refined model, in the region inside the box drawn in panel A. In this region, the VCD chain A β1 β2 hairpin interacts with actins 1 and 2, forming side contacts with the former, and tip contacts with the latter. The colors of carbon atoms of VCD chain A, chain B, actin 1, and actin 2 correspond to the colors of the corresponding representations depicted in panel A. Nitrogen and oxygen atoms are colored red and blue, respectively. VopL secondary structure elements and residues are labeled. **C.** Alternative view of the VCD-actin complex depicting VopL in surface representation and actin in ribbon representation. VopL monomers within a dimer appear as surfaces colored either blue or red. Actin monomers appear in ribbon representation and are colored yellow, cyan, or pink. The VopL chain identities and the identities of the actin monomers are indicated. The location of the apparent actin barbed end is indicated and actin subdomains 1-4 are indicated by orange numbers.

Data collection	
Energy (eV)	12,682.5
Resolution range (Å)	44.6 – 2.75 (2.80 – 2.75)
Unique reflections	49,988 (1,399)
Multiplicity	3.6 (3.3)
Data completeness (%)	98.1 (99.2)
R_{merge} (%)[a]	6.6 (74.0)
I/σ(I)	19.7 (1.5)
Wilson B-value (Å^2)	75.4
Refinement statistics	
Resolution range (Å)	44.6 – 2.75 (2.80 – 2.75)
No. of reflections R_{work}/R_{free}	49,937/2,484 (1,839/98)
Data completeness (%)	98.1 (69.0)
Atoms (non-H protein/ATP/Ca^{2+})	11,864/93/3
R_{work} (%)	21.9 (38.4)
R_{free} (%)	27.2 (40.5)
RMSD bond length (Å)	0.004
RMSD bond angle (°)	0.66
Mean B-value (Å^2) (protein/ATP/Ca^{2+})	93.4/63.0/77.8
Ramachandran plot (%) (favored/additional/disallowed)[b]	96.4/3.4/0.2
Maximum likelihood coordinate error	0.46
Missing residues, by chain	A: 335-338, 475-484; B: 332-338, 476-484; C: 1-5, 44-49, 325-326, 375; D: 1-4, 41-49, 375; E: 1-5, 39-54, 66-68, 375.

Table 5-1. Data collection and refinement statistics [a]$R_{merge} = 100\Sigma_h\Sigma_i\left|I_{h,\ i} - \langle I_h\rangle\right|/\Sigma_h\Sigma_i I_{h,\ i}$ where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.
[b]As defined by the validation suite MolProbity (Davis, Leaver-Fay et al. 2007).
Data for the outermost shell are given in parentheses. The resolution limit stated for the structure is consistent with the resolution of the data used for structure refinement. The data completeness for the resolution shells was 99+% for all shells except for the final one (from 2.8 – 2.75 Å, the completeness was 69%, see table included below), and was 98.1% overall. Inclusion of this final shell of data improved both the appearance of the electron density map and the final refinement statistics, especially the Ramachandran plot. Inclusion of data with an average I/σ(I) = 1.5, even with a merging R-factor upwards of 99%, is considered acceptable if there is sufficient information content (Karplus and Diederichs).

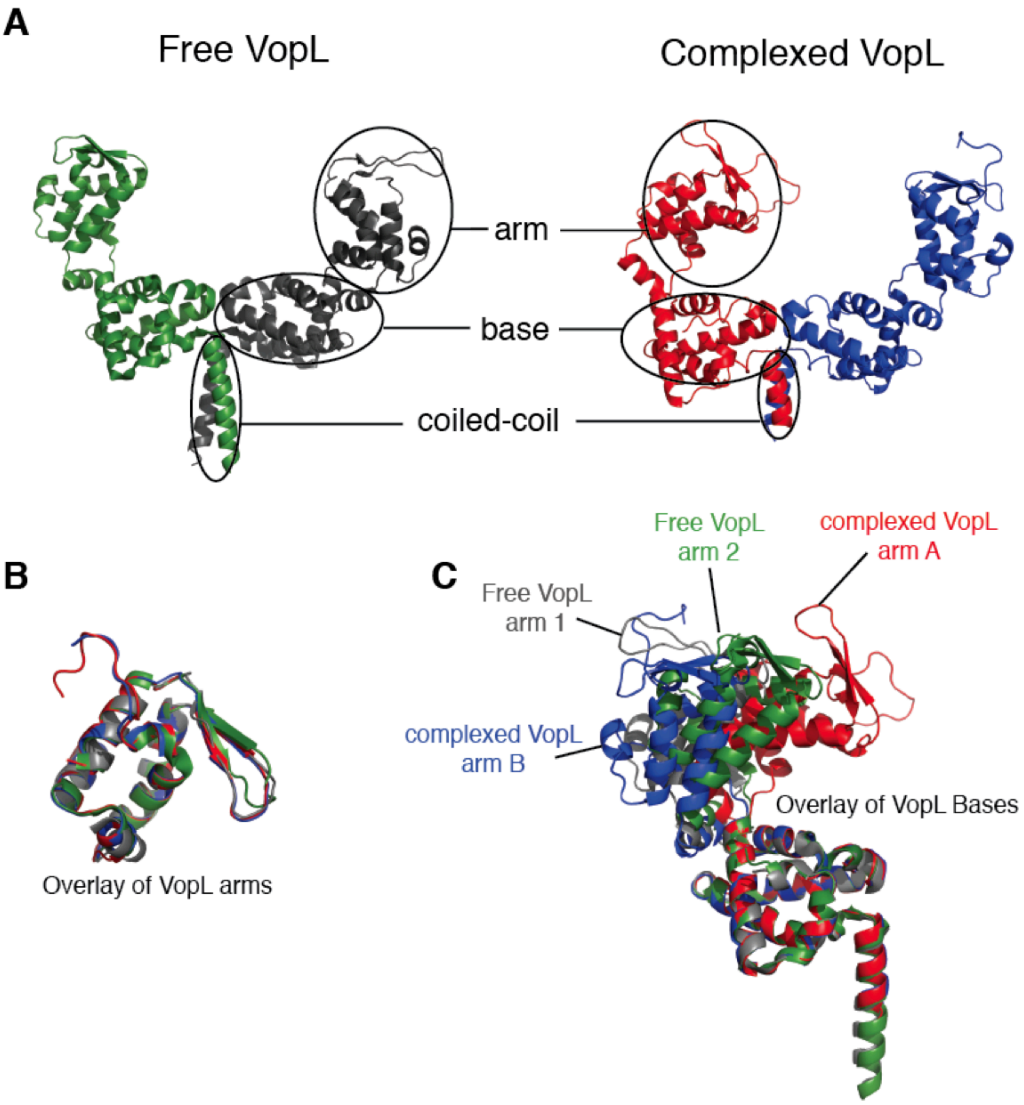


Fig. 5-5. VCD Arms in Different Structure Are Related by Rigid Body Rotations

A. Overview of free VopL (Yu, Cheng et al. 2011) and VopL in the actin-bound structure. The arm (residues 283–382), base (residues 245–278 and 395– 456), and coiled-coil (residues 462–474) domains are indicated. **B.** Alignments of the free and complexed VopL arms reveal a high degree of similarity. **C.** Backbone alignment of the four VopL monomers, restricting the alignment to the base, reveals that each of the four VopL arms adopts a distinct orientation relative to the base. See also Figure 5-6.

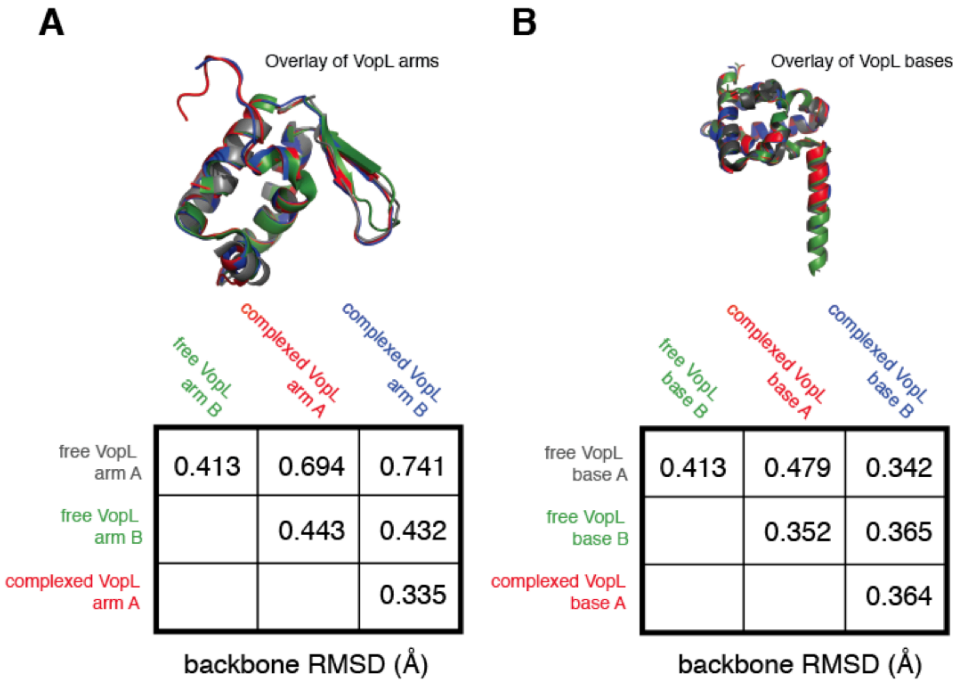


Fig5-6. Alignments of the Free and Complexed VCD Arms and Alignments of the Free and Complexed VCD Bases. A. Alignments of the free and complexed VopL arms. The table indicates the RMSD values for all possible pairwise backbone alignments of the of the four VopL arms. The alignment was restricted to residues 283-322, 340-355, and 369-382, which are ordered in both arms in both structures. **B.** Alignments of the free and complexed VopL bases. The table indicates the RMSD values of all possible backbone alignments of the four VopL base subunits. The entire VopL base was used in the alignments.

in complex with actin, the VCD dimer is asymmetric; all four arms (from the two structures) exhibit distinct orientations relative to the base platform, owing to rigid body rotations about the flexible arm-base linkers (Figs.5-6B and 5-6C).

The actins are assembled onto this structure with their pointed ends directed toward the VCD (Fig. 5-2B). Actin 1 binds to the face of the VCD dimer, making extensive contacts to both base units as well as both arms. Actin 2 does not contact the base. Rather, it sits with the pointed end of its subdomain 4 engaged with the tip of one arm. It also contacts the side of the opposite arm through the face of subdomain 2. Actin 3 sits analogously on the tip of the opposite arm, again through contacts of the pointed end of its subdomain 4. This arrangement is distinct from those in the two previously hypothesized models based on the free VCD structure (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011).

In the VCD-actin complex, the VCD arms have undergone rigid body rotations relative to the arms in the free VCD structure (Figs. 5-5A 5-5C). This allows each arm to bind to two actins, and the resulting arm-actin assemblies are pseudosymmetric: the arm A (red) contacts to actins 1 and 2 are analogous to those of arm B (blue) with actins 2 and 3 (Fig. 5-2). The similarity of these arm-actin assemblies exists even at the level of detailed contacts. We classify the arm-actin contacts into two groups—actin-“tip of arm” contacts and actin-“side of arm” contacts (Figures 2D–2F). Structural alignments of the tip of arm contacts show

that the spatial arrangement of actin 2 (cyan) relative to the VCD arm A (red) mirrors that of actin 3 (pink) relative to the VCD arm B (blue), with an overall backbone RMSD of 1.82 Å (Fig. 5-2F). The deviations derive from a small rigid body rotation of the actins, relative to the arms, apparent in an alignment based on the arms alone. Similarly, structural alignments of the side of arm contacts show that the spatial arrangement of actin 1 (light yellow) relative to VCD arm A (red) mirrors that of actin 2 (cyan) relative to VCD arm B (blue) (Figures 2D and 2E). Alignment of these elements yields a backbone RMSD of 1.15 Å. In this case, deviations are distributed throughout the structures, as evidenced by restriction of the alignments to the respective arms.

Actin Adopts a Filament-like Arrangement in Complex with the VCD

Strikingly, the two actins bound to each arm closely resemble the short-pitch dimer present in an actin filament (Fig. 5-7). In an idealized filament, successive monomers are related by a 166.4° rotation and 27.6 Å translation along the rotation axis (Fujii et al., 2010, Holmes et al., 1990, Oda et al., 2009) (Fig. 5-7E). In the VCD complex, actins 1 and 2 are related by a 154.9° rotation and 27.0 Å translation (Fig. 5-7E). Superposition of the actin1/2 dimer on a canonical actin filament model (Oda, Iwasa et al. 2009) gives a backbone RMSD of 3.33 Å. Actins 2 and 3 are similarly related by a rotation of 167.0° and translation of 29.2 Å and superimpose on the canonical filament with backbone RMSD of 2.57 Å. These

relationships suggest that the fundamental role of the VCD arm is to organize two actins into a short-pitch dimer.

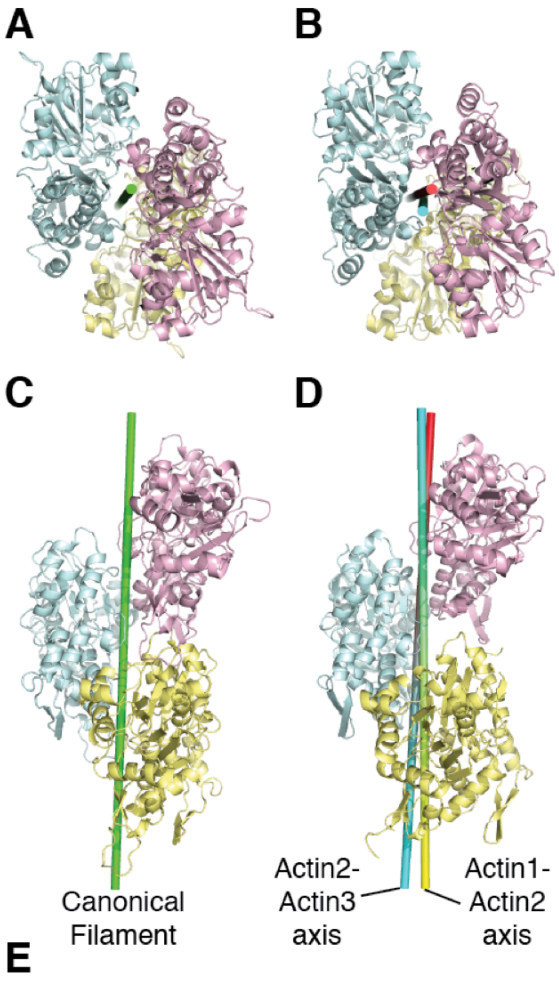
The VCD base separates the two arms, allowing them to share actin 2. This leads to an overall organization that bears close similarity to three successive monomers in a filament Fig. 5-7. The two respective short-pitch actin pairs are each associated with a rotation axis. Because the arrangement of the actins is close to that in a filament, these axes are nearly coincident (angle between axes is 5.1°, distance of closest approach between axes is 2.4 Å) (Fig. 5-7). Further, this arrangement allows the "hydrophobic plug" of actin 2 (residues 264–270) to contact the interface between actins 1 and 3, as it does in the canonical filament model (Oda, Iwasa et al. 2009, Fujii, Iwane et al. 2010) (Fig. 5-8). Thus, the inherent symmetry of the VCD, coupled with the flexibility of the arm-base attachments, enforces a filament-like arrangement. This arrangement likely lies at the core of the VopL nucleation mechanism.

Mutagenic Validation of VCD-Actin Contacts

In a previous study, mutation of charged VCD surface patches identified several residues important for nucleation activity (Table 5-2). With the VCD-actin structure in hand, it is now apparent that these mutations disrupt VCD-actin contacts and thus validate the structure. The details of additional actin-VCD contacts that were not subject to mutagenesis appear in Figure S5.

Two types of mutations were introduced into the VCD, those in the arm and those in the base; the former will be discussed first. In our structure, the tip of the VCD chain A arm contacts subdomain 4 of actin 2, and, in a nearly identical fashion, the tip of the VCD chain B arm contacts subdomain 4 of actin 3. The VCD residues that compose the “tip of arm” contacts lie in a loop spanning residues Y322–A328, in the α6 helix (I341–T350), and in the first β strand (R354–D360) (Figs. 5-10A-C). Two triple mutants affecting patches at the tip of the arm were K323E/R347E/R354E and D326G/V327G/P333G. Mutation of the first patch disrupts a hydrogen bond between VCD K323 and the side chain of actin S239 and backbone hydrogen bonds between the side chains of VCD R347 and R354 and actin G245 and V247, respectively. Mutation of the second patch disrupts a hydrogen bond between VCD D326 and actin S239 and a hydrophobic interaction between VCD V327 and actin L216.

In contrast to the VopL arms, which lie in nearly identical binding sites, the base of VopL chain A and the base of VopL chain B interact differentially with actin 1, making contacts to its subdomains 2 and 4, respectively. Contacts between actin and the VopL chain A base involve the N terminus of the VCD (L249–E251), the C terminus of helix α2 (E262–A279), the C terminus of helix α9 (K399–E408), and α10 (L409-K421) (Figures 5A and 5D). The contacts between



E

Relationship between successive monomers

Axis	Rotation	Translation
Canonical Filament	166.4°	27.6 Å
Actin1-Actin2	154.9°	27.0 Å
Actin2-Actin3	167.0°	29.2 Å

FIG. 5-7 Actins in the VopL Complex Resemble the Canonical Actin Filament. A and C. Canonical actin filament (Oda, Iwasa et al. 2009). **B and D.** Actins from the VopL complex colored as in Fig. 5-2. Axes relating pairs of actin monomers are shown as cylinders. **E.** Rotations and translations associated with the depicted axes.

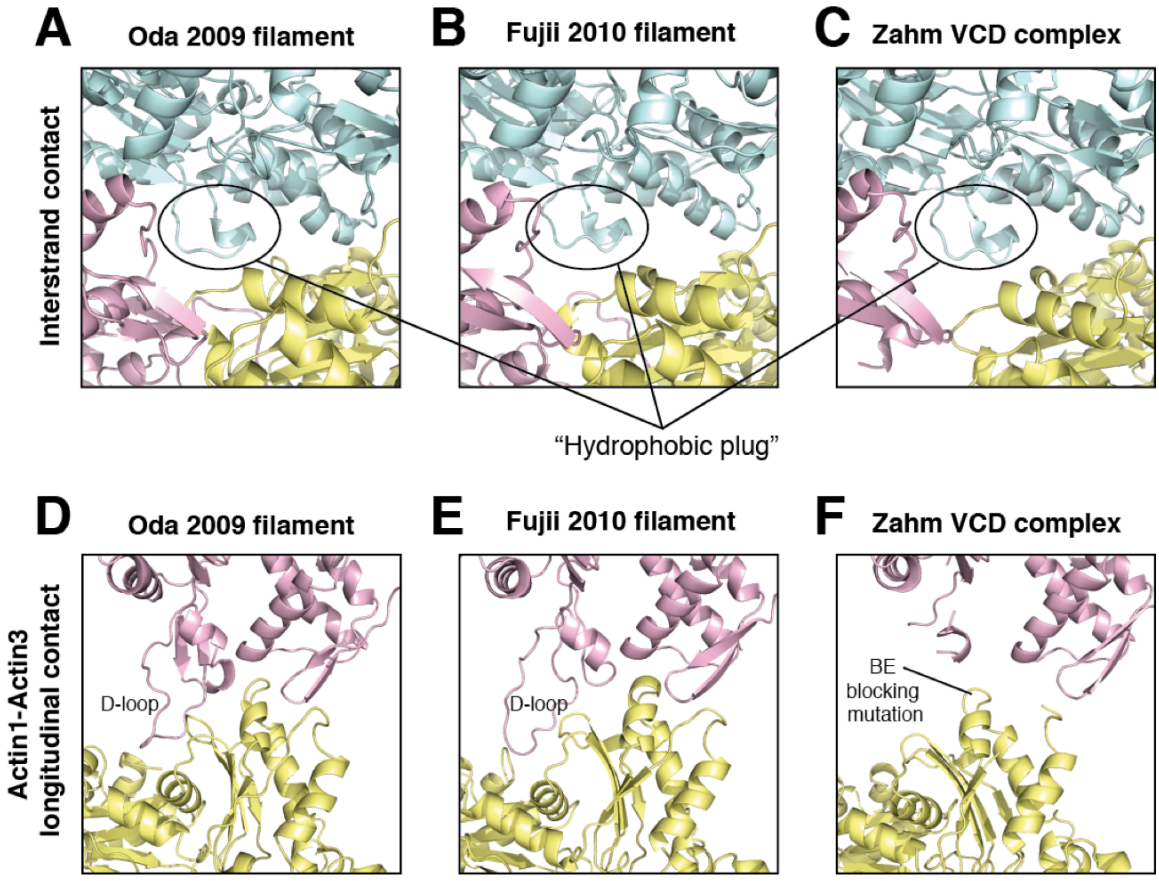


Fig. 5-8. Actin-Actin Contacts in the VCD-Actin Structure Compared to Those in Actin Filament Models, A and D, a filament model derived from oriented fiber diffraction (Oda et al.). **B and E.** a filament model derived from cryo-electron microscopy (Fuji et al.). **C and F.** The VopL-actin complex presented in this paper. Panels A, B, and C, show the organization of actins at the interstrand interface, highlighting the ‘‘hydrophobic plug.’’ This element contacts the opposing paired strand at the longitudinal interface between two actins. In the three models, this loop occupies a similar location. Panels **D**, **E**, and **F**, show the longitudinal actin-actin contact made in the three models. Filament models are characterized by the ordering of the D-loop and its insertion into a cleft in the opposing actin. In the VopL complex, the D-loop is not ordered or inserted into the yellow (lower) actin cleft, and the yellow actin has been rotated away from the proper orientation by several degrees. This interface includes the DVD mutations blocking the filament barbed end.

VopL Construct	$t_{1/2}$ (s) at 50 nM
Actin control	1352 +/- 51
Wild type	41 +/- 10
Arm mutants	
K323E / R347E / R354E	300 +/- 12
D326G / V327G / P333G	132 +/- 6
Base mutants	
K421A / Y425A / R428D	112 +/- 17
Y425A / R428D	98 +- 5
E408K / D413K / E417A	64 +/- 6

Table 5-2. Actin Polymerization Activity of VopL Mutants.

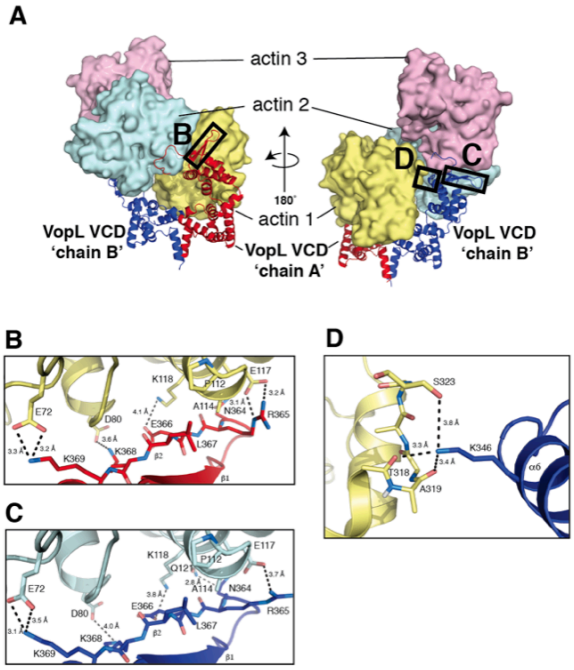


Fig5-9. Details of the Actin Side of Arm Contacts and an Additional Contact between VCD Chain B and Actin 1. Residue numbers and secondary structural elements are indicated. Contacts within 4 A° are shown as dotted lines. Probable hydrogen bonds have distances indicated. **A.** Entire VCD-actin complex. The VopL chain identities, and the actins numbers are labeled. Boxes indicate the specific regions that appear in the close-up views in panels **B** and **C**. **B.** Interface between VCD chain A arm and actin 1. **C.** Interface between VCD chain B arm and actin 2. The details of the side-of-arm contacts between the VCD chain A arm and actin 1 are analogous to those involving the VCD chain B arm and actin 2. The interfaces are predominately electrostatic and polar in nature, with actin residues Q121, E117, K118, D80 and E72 in subdomain 2 forming contacts with VopL residues in the second b strand: N364, R365, E366, K368 and K369, respectively (**B** and **C**). An additional small hydrophobic patch is composed of actin residues P112 and A114 in subdomain 2, and VopL residue L367, also in the second strand. **D.** Interface between VopL chain B arm and actin 1. Depicts an interaction between VCD chain B and actin 1 that has no analogous counterpart in VCD chain A. This is a polar interaction between K346 in the a6 helix of VCD chain B and the backbones of T318 and A319 of actin 1, as well as with the side chain of actin residue S323.

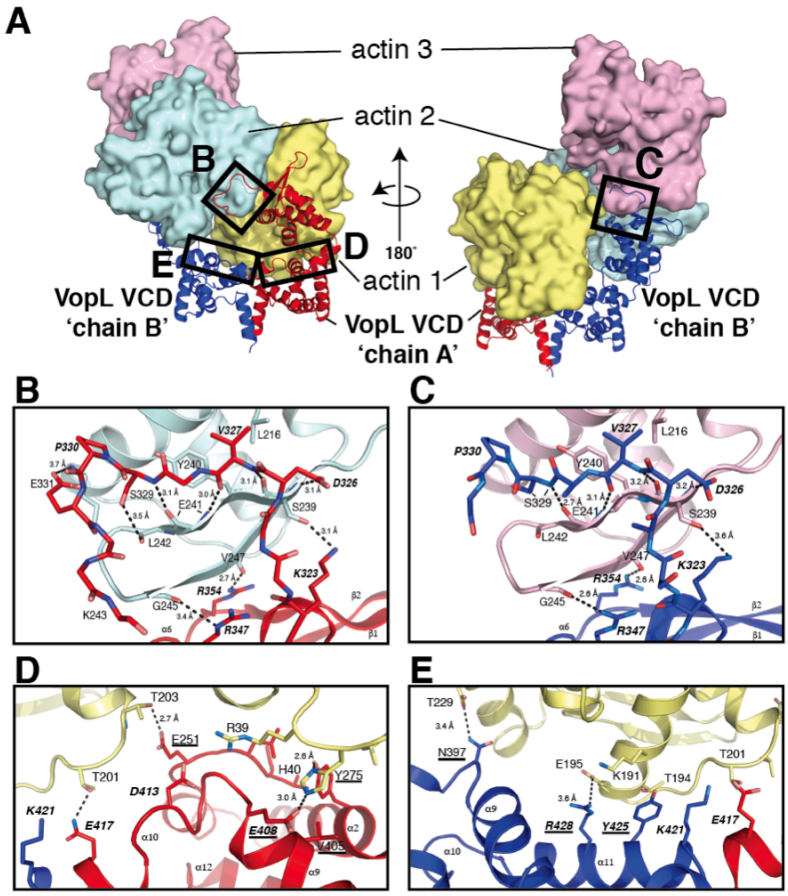


Fig. 5-10. The Details of the Tip of Arm Actin Contacts and the Contacts of the VCD Base to Actin 1. A. Entire VCD-actin complex. The VopL chain identities and the actins are labeled. Boxes indicate the specific regions that appear in the close-up views in **B–E**. **B** Interface between the tip of VCD chain A arm and actin 2. **C.** Interface between the tip of VCD chain B arm and actin 3. **D.** The interface between VopL chain A and actin 1. **E.** The interface between VopL chain B and actin 1. In all panels, the coloring scheme matches that in Fig. 5-2. In stick presentations (**B** and **C**), carbon atoms match the ribbon color; nitrogen and oxygen atoms are colored blue and red, respectively. Residue numbers and secondary structural elements are indicated. Contacts within 4 A˚ are shown as dotted lines. Probable hydrogen bonds have distances indicated. Residues labeled with bold italics were mutated in the VCD homodimer to validate the structure (Table S2). Underlined labels indicate residues that were mutated in the VopL heterodimers to selectively impair either one or both of the symmetric binding sites for actin 1. Residues appearing in bold, underlined italics were mutated in both cases. See also Figure S5 and Table S2.

chain B of the VCD base and actin are less extensive than those of chain A and occur mainly through helix α11 (residues 421–433) and the α8–α9 loop (residues 395–399) (Figures 5A and 5E). The remaining mutations lie in the interface between the VCD base and actin 1 and target two surfaces patches. Mutation of the first patch (E408K/D413K/E417A) disrupts the following contacts between VCD chain A and actin subdomains 2 and 4: a hydrogen bond between VCD E408 and actin H40, a long-range electrostatic contact between VCD D413 and actin R39, and a hydrogen bond between VCD E417 and actin T201. Mutation of the second patch (Y425A/R428D and K421A/Y425A/R428D) disrupts contacts between subdomain 4 of actin and VCD chain B. VCD Y425 makes hydrophobic contacts with the hydrophobic portions of actin T194 and K191, and VCD R428 forms a hydrogen bond with actin T201. Mutation of K421 likely disrupts a salt bridge between the VCD dimer subunits and is not directly involved with actin binding. All of the mutations discussed above disrupt contacts between the VCD and actin. Accordingly, these mutations impair VopL nucleation activity and thus establish that the VCD-actin contacts observed in the crystal are functionally important.

VCD Heterodimers Support the Asymmetric Engagement of Actin

Although the mutagenesis studies described above confirm that the contacts in the crystal structure are important for VopL nucleation activity, we sought to develop a strategy that addresses more directly the asymmetry of the structure and its role

in nucleation. The VCD dimer, by virtue of its 2-fold symmetry, can bind the actin trimer in two equivalent orientations. In any given nucleation event, VCD symmetry is broken through contacts of actin 1 to one face of the dimer. Thus, our structure predicts that mutagenic disruption of both faces of the dimer should strongly impair activity, whereas disruption of only one face should leave an equivalent binding site intact on the other face and thereby confer less impairment. Mutagenic disruption of only one face requires the introduction of different sets of mutations into the respective VCD monomers within a dimer. To achieve this, we created heterodimeric VCD proteins by replacing the C-terminal coiled coil with an engineered coiled-coil heterodimer (O'Shea, Lumb et al. 1993). Actin 1 contacts two patches of residues in the base platform, one on each VCD monomer (N397, Y425, and R428 on monomer A and E251, Y275, V405, and D413 on monomer B) (Fig. 5-10A and 5-10E). Mutating both patches in both monomers (E251A, Y275A, V405A, and E408A in one patch and N397E, Y425A, and R428E in the other patch) does not disrupt folding (Fig, 5-11) but does decrease activity relative to the wild-type heterodimer, as both faces of the VCD dimer are impaired (Figures (Figs. 5-11A-D and 5-12A-B). Similarly, mutating both patches on one monomer also decreases activity appreciably, again because both faces are affected. In contrast, and as predicted, when only one face of the dimer is disrupted by mutating the first patch on one monomer and

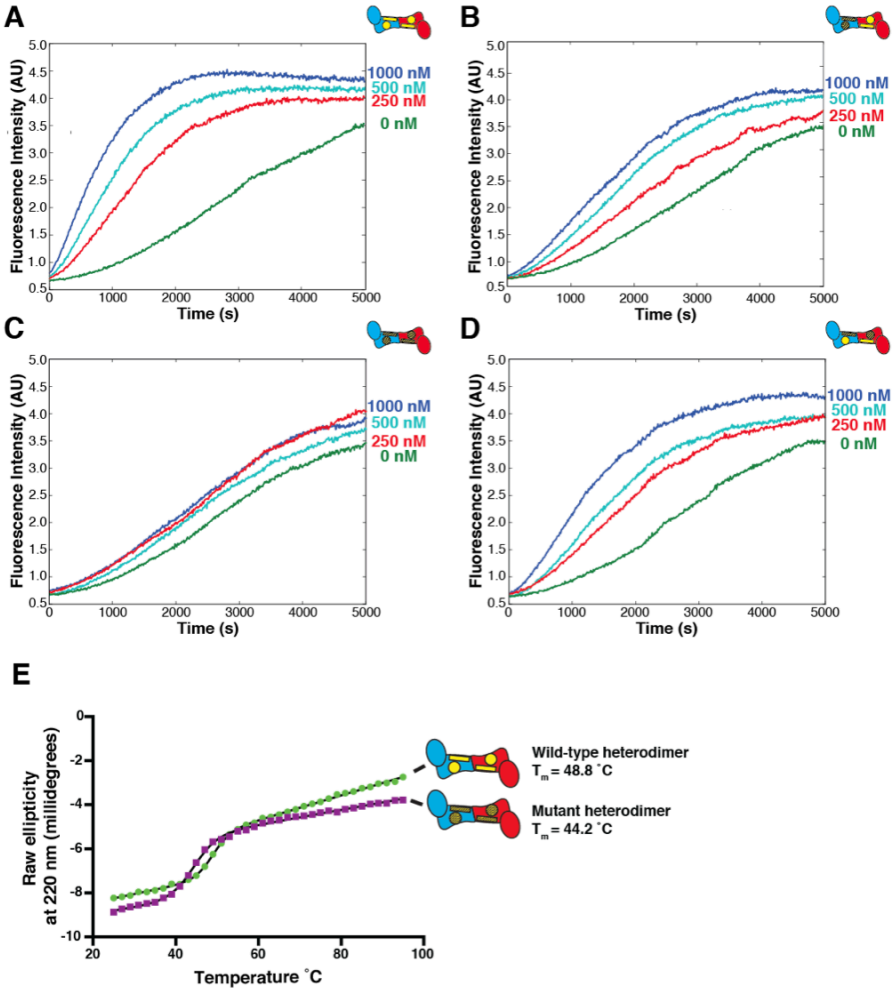


Fig. 5-11. Representative Pyrene Actin Polymerization Data from the VCD Mutant Heterodimer Experiment and Thermal Stability of VCD Heterodimer Mutants. A-D. Representative pyrene actin polymerization experiments, at 2 mM actin, 10% pyrene labeled, in KMEI, at the indicated VopL VCD concentrations. Panels differ only in the identity of the heterodimer added. **A.** Wild-type heterodimer. **B.** Heterodimer with both base patches mutated in the one subunit. **C.** Heterodimer with both patches mutated in both subunits. **D.** Heterodimer with one patch mutated in one subunit, and the other patch mutated in the other subunit. **E.** Mutant VopL heterodimers exhibit comparable stability to that of wild-type heterodimers. The plot shows raw ellipticity at 220 nm of 0.1 mg/ml samples, as a function of temperature. The data for the wild-type and mutant heterodimers are shown as green circles and purple squares, respectively. The lines represent the best fit to a six-parameter sigmoidal function. Schematic representations of the heterodimers tested, along with their associated Tm values appear to the right of the graph.

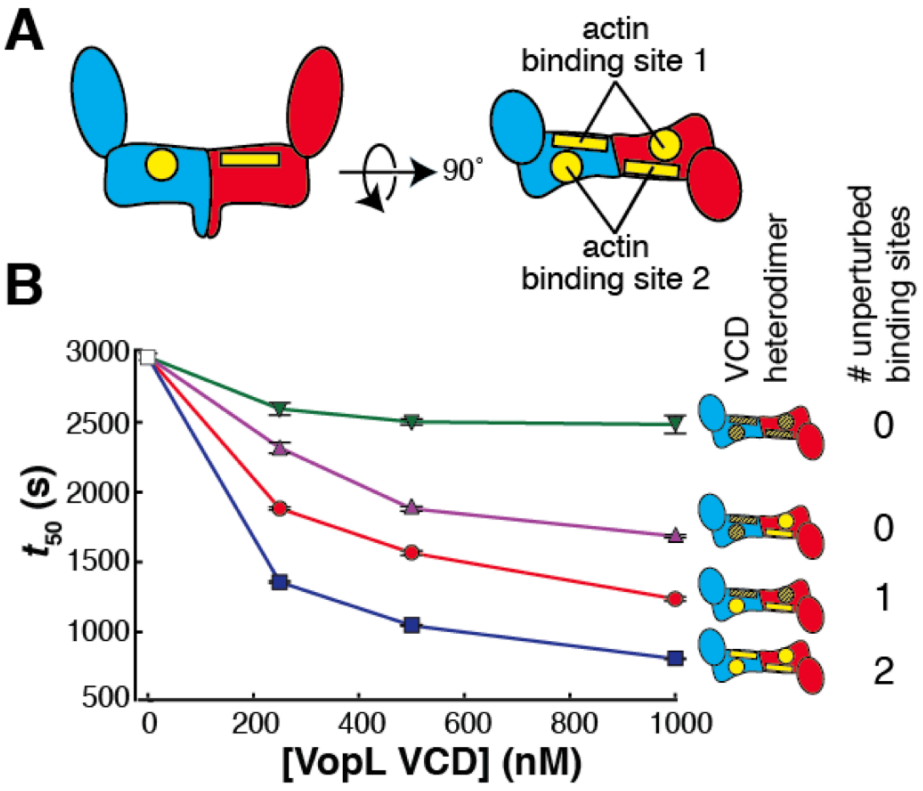
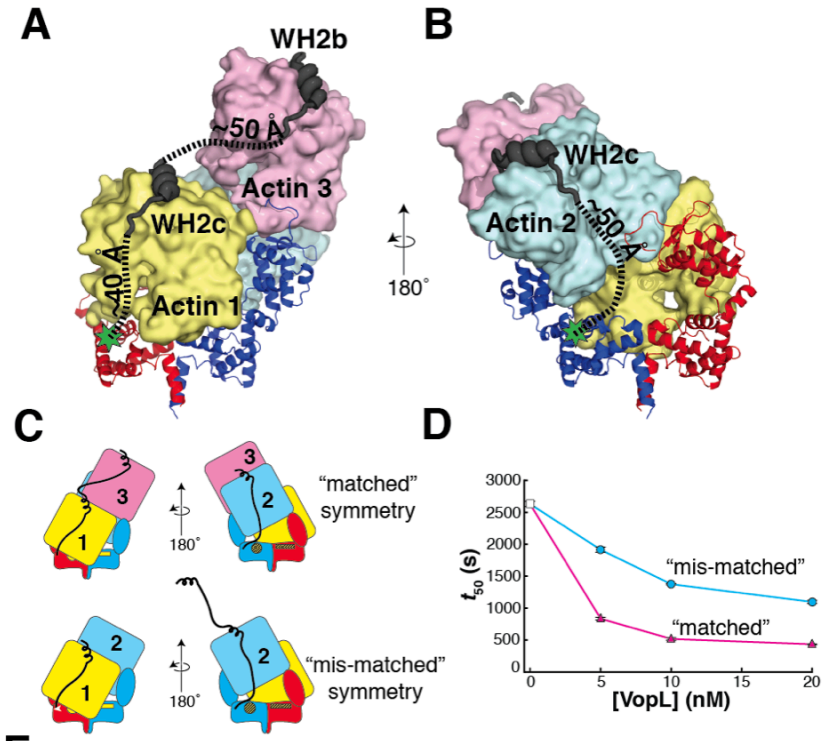


Fig. 5-12 The VopL VCD Can Bind to the Actin Trimer in Two Equivalent Orientations. A. Schematic side (left) and top (right) views of the VCD. The VCD homodimer has two symmetry-related binding sites for actin 1. We subdivide each of these binding sites into two surfaces, one residing on each VCD chain. One surface is shown as a rectangle (including residues E251, Y275, V405, and E408), and the other surface is shown as a circle (including residues N397, Y425, and R428). **B.** Actin assembly was quantified by measuring the time to half-maximal polymerization (t50). The average and SE (n = 4, error bars appear in front of the symbol) are plotted for four heterodimers possessing different numbers of unperturbed actin binding sites (wild-type sites and mutant sites are yellow and hatched, respectively). Both patches on both monomers are represented by green inverted triangles, both patches on one monomer are represented by magenta triangles, one patch on one monomer and one patch on the other monomer are represented by red circles, and the unmutated heterodimer is represented by blue squares. Experiments using 0 nM VopL (white square) were included in all data sets but are the same data.

the second patch on the other monomer, activity is only modestly decreased. These data support the structure-based prediction that the VCD nucleates actin using one of two equivalent, mutually exclusive binding sites.

WH2 Arrays Deliver Actin, which Is Organized by the VCD

Adjacent to the VCD, each VopL monomer possesses three tandem WH2 motifs that bind actin with nanomolar affinity (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011). This array exhibits little nucleation activity on its own, but when fused to the VCD, it increases the potency of VopL by over 100-fold (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011). The WH2 motifs likely enhance activity by binding to and delivering actin monomers to the VCD, compensating for the modest affinity of the VCD for actin. To understand this process, we modeled WH2 motifs onto the actins in our structure using a previously reported crystal structure of a VopL WH2-actin complex (Rebowski, Namgoong et al. 2010). WH2 motifs possess an amphipathic helix that binds in the cleft between subdomains 1 and 3 at the barbed end of the actin monomer, followed by an “LKKT motif” that binds to the actin face (Dominguez 2009, Carlier, Husson et al. 2011). Each of the three actins in the complex readily accommodates a WH2 motif without steric clashes. The 23 residue linker between the LKKT sequence of the third WH2 motif and the N terminus of the



	VCD alone	WH2 array alone	WH2 array + VCD
Affinity for actin	weak	strong	strong
Actin organization	strong	weak	strong
Actin nucleation	moderate	weak	strong

Fig. 5-13. The VopL WH2 Domains Recruit and Deliver Actin Monomers to the VCD. A. WH2b and WH2c modeled onto actin 3 and actin 1, respectively. **B.** WH2c modeled onto actin 2. In **A** and **B,** green stars indicate the position of the VCD N- terminus. Dashed lines approximate the trajectories of the paths used to estimate the distances between structured elements in the model (VCD N terminus to WH2c and WH2c to WH2b). **C.** Illustration of VopL heterodimers that harbor mutations in the VCD that disrupt binding of actin 1 to one face and, additionally, are fused to a total of three WH2 motifs, an array of two WH2 motifs on one VCD monomer and a single WH2 motif on the other. The respective VCD monomers within a heterodimer are red or blue. Actins 1 (yellow), 2 (blue), and 3 (pink) are

indicated. The unmutated face of the VCD is denoted by bright yellow patches. The mutated face of the VCD appears as yellow patches with black hash marks. Only the unmutated face could bind actin 1 in the mode observed in the crystal structure. The WH2 arrays are represented as black lines that emanate from the respective VCD monomers. In the construct with ‘‘matched’’ symmetry, the $(WH2)_2$ array de- livers actins 1 and 3 such that actin 1 contacts the unmutated face of the VCD; the single WH2 motif delivers actin 2, creating a stable trimer (Sept and McCammon, 2001). In the construct with ‘‘mismatched’’ symmetry, if the system uses the wild- type face of the VCD, there is no WH2 motif to deliver actin 3 longitudinally to actin 1; the $(WH2)_2$ array could deliver an actin longitudinally to actin 2 (not shown in cartoon), but this would not create a stable actin trimer. **D.** Actin assembly was quantified by measuring the time to half-maximal polymerization (t_{50}). The average and SE (n = 4, error bars appear in front of the symbol) are plotted for two different VopL heterodimers. Both heterodimers harbor mutations that disrupt a single face of the VCD, each possessing a single WH2 motif fused to one VCD monomer and two tandem WH2 motifs on the other. However, the two heterodimers differ from one another in how the WH2 arrays are positioned relative to the mutated surface on VCD, referred to as matched (magenta triangles) or mismatched (cyan circles) as shown in **C**. Experiments using 0 nM VopL (white square) were included in both data sets but are the same data. **E.** A summary of the roles of the VCD and the WH2 arrays in VopL-mediated actin nucleation. The cartoons represent the ensemble behavior in the presence of actin, for the VCD alone, the WH2 array alone, and full-length VopL. The rows in the table indicate the capacity to bind to actin, to organize actin into filament-like structures, and to nucleate actin filaments. Note that, for visual clarity, in **C** and **E**, we represent the unbound WH2 motif as helical, although physical data suggest that, in the free state, WH2 motifs are disordered.

VCD can span the modeled distance between these points for both actin 1 and actin 2 (~40 and ~50 Å, respectively) (Figs. 5-13A-B). We note that the structure contradicts our previous hypothesis regarding the aberrantly low activity of a fusion of the third WH2 motif with the VCD (Yu, Cheng et al. 2011), an observation we are currently working to understand. The 20 residue linker between the second and third WH2 motifs (WH2b and WH2c, respectively) is sufficiently long to span the distance between the WH2 motifs modeled onto actins 1 and 3 (Fig. 5-13A) but is insufficient to span the distance between the WH2 motifs modeled onto actins 1 and 2. Thus, it is likely that one WH2 array delivers actins 1 and 3, whereas the opposite array delivers actin 2.

In order to test this model, we returned to our heterodimer strategy. As described above (Figures 5 and 6), mutating both actin-binding patches on one face of a VCD heterodimer restricts assembly of the three actins to predominantly one orientation, dictated by binding of actin 1 to the wild-type face. In a minimal system in which a total of three WH2 motifs recruit the actin trimer, there exists a maximally active configuration in which one VCD monomer within a dimer uses two WH2 motifs to deliver actins 1 and 3 to the wild-type face, and the other VCD monomer uses a single WH2 motif to deliver actin 2. If the WH2 arrays were swapped, the actin geometry in the initial trimer would not match that of the recruiting elements because there would be no WH2 motif to recruit actin 3 (Fig. 5-13C). Thus, the “single-face” heterodimer, by forcing use of only one face of the

VCD, provides a means of examining the relationship between actin geometry controlled by the VCD and delivery of actin by the WH2 arrays of the individual subunits.

As in a previous experiment (Fig. 5-12), in order to impair one face of a VCD heterodimer, we mutated N397, Y425, and R428 on one subunit (monomer A, red in Figure 7B) and E251, Y275, V405, and D413 on its partner (monomer B, blue in Figure 7B). As shown in Figures 5-13A and 5-13B, our model predicts that WH2 motifs of monomer A (red) should deliver actins 1 and 3 to this heterodimer, whereas those of monomer B (blue) should deliver actin 2. As shown in Fig. 5-13D, when monomer A is joined to two WH2 motifs and monomer B is joined to one motif, activity of the heterodimer is high. But when the WH2 arrays are swapped so that monomer A has only one WH2 motif and monomer B has two motifs, activity is low (Fig. 5-13D). Thus, when the geometry of actins organized by the VCD is matched to the geometry of recruitment by WH2 motifs (Fig. 5-13C, top), the two elements can act synergistically to promote nucleation. But when the geometries are mismatched (Fig. 5-13C, bottom), synergy is less and activity is lower. Together, these data provide strong support for our model of WH2-mediated delivery of actin to the VCD.

We note that, in the full-length protein, the system likely proceeds through multiple complexes containing between three and six actins delivered by the WH2 arrays and is organized by the VCD according to this same general plan (Fig. 5-

13E, right). The distribution of these complexes will depend on the actin concentration and the relative rates of actin binding and nucleation, but there is not necessarily a single defined pathway involving all six WH2 motifs.

Deviations from the Canonical Filament Model

The organization of actins in the VCD complex qualitatively resembles that in an actin filament. However, some details of the structure are quantitatively distinct from the filament. First, the rotational and translational relationships between monomers do not exactly match those in an ideal filament (although we note that monomers within individual actin filaments can exhibit substantial angular deviations from ideality (Schmid, Sherman et al. 2004). Second, the actin monomers adopt conformations that more closely resemble that of the unpolymerized actin monomer than the monomer in a filament, with backbone root-mean-square deviation (rmsd) of 0.6–0.9 Å versus 2 to 3 Å, respectively. Third, the configuration in our structure positions the subdomain 2 DNase I binding loop (D loop) of actin 3 too far from the barbed end cleft of actin 1 for these elements to engage as they do in a filament (Figure S4F) (Oda, Iwasa et al. 2009, Fujii, Iwane et al. 2010). Some of these deviations may result from our use of the actin mutant to obtain crystals. However, in order for WH2 domains to bind actin and participate in nucleation, the D loop cannot be bound in the barbed end cleft. Thus,

accessibility of the cleft to WH2 domains and the absence of D loop engagement likely reflect an authentic intermediate in the nucleation pathway.

Discussion

Here, we have described the structure of the VopL VCD in complex with three actin monomers. The VCD organizes the actin monomers into a trimer that closely resembles an actin filament. This positioning of the actins is dictated by binding of each arm to an actin pair that closely resembles a short-pitch actin dimer and by sharing of one actin between the pairs. This positioning leaves all actin subdomain 1–subdomain 3 clefts accessible to the array of WH2 motifs extending from the amino terminal end of each VCD subunit.

Together, these observations suggest a mechanism of nucleation that relies on synergistic actions of the VCD and WH2 domains. In this mechanism, the VCD organizes actin monomers to closely resemble an actin filament. But because the domain has low affinity for actin, it has low nucleation activity in isolation. Conversely, the WH2 domains capture actin monomers with high affinity but are poor organizers due to their inherent flexibility. When together in the same molecule, the two elements can effectively bind and organize actins in a filament-like configuration (Figure 4). At some stage during nucleation, either before or after additional monomers bind, this structure must reorganize into a bona fide filament in which actin monomers are in the filamentous conformation and the D-loops are

engaged in the barbed end clefts. This rearrangement could weaken VCD-actin interactions and would block interactions between the WH2 amphipathic helix and actin, contributing to the observed rapid dissociation of VopL from nascent filaments (Namgoong, Boczkowska et al. 2011).

We posit that the structure of the VopL VCD bound to actin provides a snapshot of the nucleation process, suggesting three general mechanistic themes that describe the actions of nucleation factors. First, this structure provides strong evidence that nucleation factors act by organizing monomers into a filament-like configuration. Second, comparison to other nucleation factors suggests that the separation of actin organizing and recruiting functionalities that we observe for VopL may be general. Third, the dynamic association of a given nucleation factor with a filament end may depend on the degree to which the nucleation factor organizes actins to resemble the canonical filament structure.

A long-standing hypothesis is that actin filaments arise from filament-like nuclei and that the fundamental purpose of nucleation factors is to facilitate the formation of these structures (Dominguez 2009). The VCD-actin structure provides the strongest evidence to date that this idea is in fact correct; VopL positions three actins in a manner very similar to those in the canonical filament. Previous crystal structures of actin-nucleation factor complexes have been less compelling in this regard; actin-actin contacts have arisen from crystallographic symmetry and do not match those in the canonical filament to the degree that they do in this structure.

The crystal structure of the formin Bni1p engaged with actin (Otomo, Otomo et al. 2005) revealed that actin lies in a filament-like configuration but with an 180° rotation dictated by a crystallographic symmetry axis. The actin monomers in the structure of FMNL3 bound to actin also made contacts along a crystallographic axis but did not adopt a filament-like configuration (Thompson, Heimsath et al. 2013). Electron microscopy reconstruction of Arp2/3 complex engaged with a nucleated filament (at 26 Å) and in a soluble activated state (at ~20 Å) provides additional views of nucleation (Xu, Moseley et al. 2004, Rouiller, Xu et al. 2008) that are consistent with the idea that nucleation factors function by organizing filament-like nuclei. That several of these nucleation factors appear to induce short-pitch-like contacts strongly supports the idea that actin nucleation occurs through filament-like nuclei and that actin nucleation factors function by generating such structures.

Based on the structure and biochemistry, we propose that VopL functions via a division of labor whereby the WH2 domains bind actin monomers with high affinity and deliver them to the VCD, the latter serving to enforce a filament-like arrangement. These two distinct functionalities, present within the same molecule, give rise to potent nucleation activity (Figure 7E). Such division of labor also appears to be important to both the Arp2/3 complex and formins. In Arp2/3-mediated nucleation, Arp2 and Arp3 serve a function analogous to that of the VCD in that they act as an organizing entity of low inherent affinity for actin monomers

(Mullins, Stafford et al. 1997, Kaiser, Vinson et al. 1999, Padrick, Doolittle et al. 2011). When Arp2/3 complex is activated by WASP family protein VCA peptides, the WH2 motifs in VCA bind actin monomers with high affinity and deliver them to the Arp2/Arp3 organizing template. In formin proteins, the FH2 domain serves a function analogous to that of the VCD in that it too induces a filament-like actin nucleus (Otomo, Otomo et al. 2005). Despite high affinity for the filament barbed end, many FH2 domains exhibit only modest affinity for actin monomers or only bind a single monomer (Thompson, Heimsath et al. 2013). In some formins, such as Inf2 (Chhabra and Higgs 2006), FMNL3 (Heimsath and Higgs 2012), and mDia1 (Gould, Maiti et al. 2011), there is a WH2-like sequence C-terminal to the FH2 domain. In these proteins, the FH2 domains have weak activity on their own but have high activity in concert with the additional sequences. These data are consistent with the FH2 domains acting as an organizing template and the WH2-like sequences delivering actin monomers. Furthermore, and consistent with the overall idea, the formin mDia1 may use the nucleation factor APC as an actin recruiting factor (Okada, Bartolini et al. 2010). Taken together, this “division of labor” appears to be a mechanistic theme that applies to diverse nucleation factors.

Different nucleation factors exhibit distinct dynamic properties at the end of the nascent filament. Structural data on the Arp2/3 complex (Rouiller, Xu et al. 2008) (Rouiller et al., 2008), the formin Bni1p (Otomo, Otomo et al. 2005), and now VopL, suggest that deviations from an ideal filament, or lack thereof, may

underlie this variability. The Arp2/3 complex remains persistently associated with the pointed end of the filament it nucleates. Arp2/3 complex in a filament branch organizes the actin homologs Arp2 and Arp3 in a configuration indistinguishable from that of actins in the filament (Xu, Moseley et al. 2004, Rouiller, Xu et al. 2008). The formin FH2 dimer moves with the end of the nascent filament through cyclic release of individual formin subunits from the terminal two actins and subsequent binding of this subunit to an incoming actin monomer (Xu, Moseley et al. 2004, Goode and Eck 2007, Rouiller, Xu et al. 2008). In complex with the Bni1p FH2 domain, actins adopt a filament-like but strained configuration; relief of this strain has been invoked to explain the binding and release dynamics of the FH2 subunits at the filament end (Otomo, Otomo et al. 2005, Thompson, Heimsath et al. 2013). We have now found that VopL, which dissociates from filaments shortly after nucleation (Namgoong, Boczkowska et al. 2011), also arranges actin monomers in a configuration that deviates from the canonical filament. Together, these behaviors suggest that strain in the nucleus and/or nascent filament may play an important role in dictating the dynamic behaviors of nucleation factors. Those that are structured to bind an ideal filament may remain persistently attached, whereas those that bind strained configurations may be dynamic. Such features could be selected for during evolution to produce the specific behaviors needed for distinct biological activities.

Experimental Procedures

Protein Purification

VopL VCD (residues 247–484) was expressed in Escherichia coli and purified as previously described (Yu, Cheng et al. 2011). Briefly, VCD was expressed as a glutathione S-transferase (GST) fusion in BL21(DE3)-T1R E. coli at 18°C overnight. VCD was purified over glutathione sepharose (GE) and cleaved off the beads with TEV protease. Further purification was accomplished using SOURCE15Q ion exchange followed by Superdex 200 gel filtration (GE) chromatographies. VCD mutants were made using site-directed mutagenesis or gene synthesis (Genscript). VCD heterodimers and VopL heterodimers bearing the minimal set of three WH2 domains were produced by replacing the VopL coiled coil with one of two compatible “peptide velcro” coiled-coil pairs and by coexpressing these proteins in BL21(DE3)-T1R bacteria from two origin-compatible plasmids, pMAL-c2 (NEB) and pCDF1 (Novagen). TEV protease cleavable maltose-binding protein fusions with an engineered acidic helix (O'Shea, Lumb et al. 1993) were cloned into a pMAL-C2-derived vector (NEB). TEV protease cleavable His6 fusions with an engineered basic helix (O’Shea et al., 1993) were cloned into a pCDF1- (Novagen) derived vector. VCD heterodimers and heterodimers bearing the minimal set of three WH2 domains were purified using amylose (NEB) affinity chromatography, followed by Ni-NTA affinity chromatography (QIAGEN), and affinity tags were removed using TEV protease.

Purification to homogeneity was accomplished using SOURCE15Q and SOURCE15S ion exchange chromatography (GE) for the VCD constructs and the WH2-VCD constructs, respectively. Subsequently, heterodimers were subjected to gel filtration chromatography using Superdex 200 pg resin (GE) in KMEI buffer (50 mM KCl, 10 mM imidazole [pH 7.0], 1 mM EGTA, and1 mM MgCl2). Barbed-end-blocked His6-actin (Drosophila melanogaster 5C actin with D287A, V288A, and D289A mutations) was expressed using a recombinant baculovirus produced in Sf9 cells using the Bac-to-Bac system and a modified pFastBacHT expression vector (Invitrogen) that included an L21 enhancer sequence (Sano et al., 2002) driving expression of the mutant 5C actin. Mutant actin was purified by nickel affinity chromatography. Then, TEV protease was used to remove the His6 tag. The purification was completed using anion exchange followed by gel filtration chromatography. For crystallization, a 1:1 molar mixture of VCD and actin was dialyzed for 16 hr into 10 mM Tris (pH 8.0), 50 mM KCl, 0.2 mM ATP, 2 mM TCEP, and 0.1 mM CaCl2 and concentrated to 20 mg/ml.

A VopL WH2c peptide (199–226) bearing an introduced C-terminal tyrosine residue to aid in quantitation by absorbance was expressed in E. coli as a GST fusion. Following cell lysis, the fusion was affinity purified using 5 ml glutathione agarose following manufacturer recommendations. Following elution from glutathione agarose, the fusion was concentrated to 1 ml using a centrifugal concentrator (Amicon Ultra 15, 30000 MWCO UFC903024, Millipore), and the

peptide was cleaved from the GST fusion with TEV protease at 4°C overnight. The peptide was subsequently purified by two cycles of gel filtration chromatography (Superdex 75 10/300 GL, GE Healthcare) in KMEI buffer supplemented with 0.2 mM ATP.

Structure Determination

The VCD-actin mixture was crystallized from hanging drops containing 1 µl protein and 1 µl of well solution (0.1 M MMT buffer [pH 8.0] and 24% PEG1500), yielding crystals with dimensions up to 500 × 500 × 300 µm3. Crystals were flash frozen directly from the drop in liquid nitrogen. Diffraction data were collected at beamline 19-ID at the Advanced Proton Source (APS) (Argonne National Laboratory). The data were processed and scaled using HKL3000 (Otwinowski and Minor 1997). The structure was solved by molecular replacement using Phaser (McCoy, Grosse-Kunstleve et al. 2007) as implemented in Phenix (Adams, Afonine et al. 2010) with actin (Nair, Joel et al. 2008) and VCD arm fragments and VCD base/coiled-coil dimers (Yu, Cheng et al. 2011) as search models. The model was improved through iterative cycles of model building in Coot (Emsley, Lohkamp et al. 2010) and through subsequent positional and TLS atomic displacement parameter refinement in Phenix.

Biochemistry

Actin assembly assays contained 2 μM rabbit muscle actin (10% pyrene labeled) in KMEI supplemented with 100 μM ATP, following a previously described method adapted to a multiwell plate format (Leung et al., 2006). Actin spin-down assays were performed using 9 μM actin in KMEI. Polymerization was allowed to proceed for 16 hr at room temperature, and samples were centrifuged at 100,000 g for 2 hr at 20°C. Supernatant and pellet fractions were collected and analyzed using Coomassie stained SDS-PAGE gels.

Isothermal Titration Calorimetry

ITC experiments were conducted using a VP-ITC microcalorimeter (Microcal) at 20°C. Prior to each experiment, WH2 peptides and mutant actin were subjected to gel filtration chromatography to exchange them into KMEI buffer supplemented with 0.2 mM ATP. In each experiment, 100 μM of WH2 peptide was titrated into 10 μM actin in KMEI supplemented with 0.2 mM ATP. Data were analyzed and fit to a single-site model using Origin 7 for ITC (Microcal).

Individual Filament Growth Assay

Rabbit muscle actin was purified as previously described (Leung et al., 2006). Rabbit actin (10 ml of 40 μM in G-buffer without sodium azide and without DTT)

was polymerized by dialysis overnight in KMEI supplemented with 0.2 mM ATP. With gentle stirring, a 5-fold molar excess of Oregon green C5-maleimide in DMSO was added, and labeling was allowed to proceed overnight in the dark at 4°C. This reaction was then dialyzed against 2 l of G-buffer for 3 days, clarified by centrifugation and purified by gel filtration column (Superdex 200 pg 26/600, GE Healthcare). Fluorescently labeled actin seeds were prepared by adding ~0.6 μM AlexaFluor 546-phalloidin (A22236, Life Technologies) to a 4 μM solution of rabbit actin in KMEI in a final volume of 200 μl. Polymerization was allowed to proceed for 2 hr. Reactions to be imaged by TIRF microscopy were prepared with 0.5 μM unlabeled rabbit actin, 0.15 μM Oregon green-labeled actin, 10 μl of phalloidin seed stock, and either 0 nM, 25 nM, or 65 nM purified mutant actin in G-Mg buffer (2 mM TrisHCl [pH 8.0], 0.2 mM ATP, 1 mM MgCl2) in a final volume of 200 μl. Seed growth was initiated by addition of 20 μl 10× KMEI to 180 μl of the aforementioned mixture. Reactions were allowed to proceed for 10 min, at which time they were diluted 1:200 in TIRF buffer (KMEI supplemented with 15 mM glucose, 100 μg/ml glucose oxidase, 20 μg/ml catalase). Diluted filaments were imaged in flow cells, the assembly, NEM-myosin coating, and loading of which is previously described (Hung et al., 2011). Filaments were imaged using an Olympus IX-71 fitted with an Olympus TIRF arm, a PlanApo 100× oil objective (n.a. 1.45), and a Photometrics Cascade II:512 EMCCD camera. The microscope was run using Micro-Manager (Vale lab). Dual-color imaging was done by switching

between excitation lasers using shutters and was filtered using a dual-color filter cube (Semrock). Pixel size was calibrated with a micrometer. Exposure times were kept constant between conditions. All images were processed in NIH ImageJ using a Gaussian filter to reduce shot noise and were background subtracted. Filament lengths were quantified by total pixel number after simple thresholding, conversion to binary, and using the skeletonize function. 75–200 individual filament lengths were quantified for each condition.

Circular Dichroism

VopL heterodimers (wild-type and mutant) were purified as described above but purified by gel filtration into 50 mM sodium phosphate (pH 8.0), 100 mM sodium chloride, and 1 mM TCEP. 0.1 mg/ml samples were placed into 1 mm path length quartz cuvettes, and CD signal was measured at 220 nm as temperature was raised from 25° to 90°C in 2° increments. Temperature was allowed to stabilize for 10 s before measuring CD signal, and CD signal was averaged for 32 s. Data were fit to a six-parameter sigmoid, and Tm reported was the inflection point of the melting transition.

Molecular Modeling

The VopL WH2c motif (Protein Data Bank [PDB] ID 3M1F) (Rebowski et al., 2010) was modeled onto the VCD structure using Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger) by superimposing actins in the WH2-actin and VCD-actin complexes. Actin-actin rotation axes in the VCD complex were determined by aligning successive actins to each other in QtMG30 (Potterton, McNicholas et al. 2002) using the “Superpose” tool.

Ensuing Controversy

The work in this thesis describing the structure of the VopL VCD in complex with actin monomers strongly corroborates the general scheme through which VopL brings about filament formation (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011). These studies, combined with the one presented here reveal a mechanism in which VopL arranges a filament-like nucleus that, upon transition to a filament, displaces VopL and grows from its barbed end. Despite this, there is an alternative hypothesis that has led to ensuing controversy in the field.

The work leading to the alternative hypothesis utilizes VopF, a close relative of VopL, from *Vibrio Cholerae*, that bears 72% sequence similarity to VopL, with both molecules sharing the same domain structure (Vizcarra, Kreutz et al. 2011). Given the high degree of similarity and the identical domain organization, it is unlikely that these molecules act through distinct structural mechanisms. Yet in the

alternative model, Carlier and colleagues claim that the VCD does not possess any intrinsic nucleation capacity, and serves only to dimerize the WH2 arrays, which possess the core nucleation activity. It seems rather curious that they would make this claim, as their bulk polymerization data show that the isolated VCD enhances actin polymerization, albeit weakly, consistent with previous studies (Namgoong, Boczkowska et al. 2011, Yu, Cheng et al. 2011) (and moreover that we have shown numerous VCD mutants that impair activity of the full length protein). They further show, by imaging individual filaments using TIRF microscopy, that VopF severs actin filaments through the action of its WH2 array, and use bulk polymerization assays to conclude that VopF binds to filament barbed ends (rather than the pointed end as we observe in our crystal structure), and that in doing so, the protein can displace capping protein or gelsolin, bringing about subsequent barbed-end growth. The authors next suggest that, because VopF displaces gelsolin and capping protein, and because the filaments subsequently elongate, that VopF necessarily remains processively associated with the growing barbed end. (Pernier, Orban et al. 2013). It important to note that the authors never observe barbed-end binding or processive association directly, but, deduce their existence though the interpretation of bulk polymerization assays and single color TIRF microscopy of individual filaments. The authors never discuss or consider the possibility that the process of accepting additional monomers and undergoing elongation might displace VopF from the barbed end of a growing filament. The

authors also conclude that, because VopF does not bind to the pointed-end of pre-formed filaments, the pointed-end nucleation mechanism presented in this thesis is wrong. Given that the inter-subunit engagement of D-loops that occurs upon transition to a filament sterically blocks the WH2 binding site in the 1,3-cleft, it is unsurprising that the authors do not see binding to filament pointed ends, and this observation in no way precludes the pointed-end nucleation model, which posits that VopL exhibits high affinity for a penultimate nucleus rather than the bona fide filament; The engagement of D-loops that accompanies transition to a filament almost certainly decreases the affinity of the VopL WH2 arrays for actin. If it is true that VopF binds to the barbed ends of filaments and displaces of capping protein and Gelsolin, this too is rather unsurprising, as barbed ends possess a 1,3-cleft to which WH2 domains bind with nanomolar affinity. Our proposed nucleation mechanism does not preclude the possibility that VopL/F can also function as a regulator of filament barbed-end growth, although, in my opinion, the study from Pernier et. al. does not establish definitively whether or not this is the case, and doing so required a more rigorous investigation.

A study subsequent to the one from Pernier et. al. utilizes single-molecule, multi-color TIRF microscopy to observe the interactions of VopL and VopF with actin filaments (Burke, Harker et al. 2017) in the presence and absence of free actin monomers and actin binding proteins, with a goal of addressing, more rigorously, the hypotheses put forth in the earlier study. The authors found that

VopL and VopF function indistinguishably from one another, and that both nucleate filaments from the pointed end, and that, following nucleation, both remain associated with the nucleated filament briefly (~110 s) before dissociating from that filament. Following nucleation of a filament, both VopL/F and capping protein bind to the pointed and barbed ends, respectively, and there is no evidence that VopL/F displaces capping protein, in contrast to the results of Pernier et al. In addition, the study finds no evidence that VopL/F associate processively with growing filament barbed ends. The study does show that both VopL/F bind with equal preference to the barbed and pointed ends of pre-assembled filaments, but do so only in the absence of a pool of free actin monomers. Thus, the process of elongation likely displaces VopL/F associated with the barbed end of a pre-formed filament. This study strongly supports the pointed end nucleation model and speaks strongly against the notion that VopL/F function as regulators of barbed-end growth.

References:

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta Crystallogr D Biol Crystallogr **66**(Pt 2): 213-221.

Alberts, A. S. (2001). "Identification of a carboxyl-terminal diaphanous-related formin homology protein autoregulatory domain." J Biol Chem **276**(4): 2824-2830.

Alto, N. M. and K. Orth (2012). "Subversion of cell signaling by pathogens." Cold Spring Harb Perspect Biol **4**(9): a006114.

Anthis, N. J. and G. M. Clore (2015). "Visualizing transient dark states by NMR spectroscopy." Q Rev Biophys **48**(1): 35-116.

Balcer, H. I., K. Daugherty-Clarke and B. L. Goode (2010). "The p40/ARPC1 subunit of Arp2/3 complex performs multiple essential roles in WASp-regulated actin nucleation." J Biol Chem **285**(11): 8481-8491.

Blanchoin, L., R. Boujemaa-Paterski, C. Sykes and J. Plastino (2014). "Actin dynamics, architecture, and mechanics in cell motility." Physiol Rev **94**(1): 235-263.

Boczkowska, M., G. Rebowski, D. J. Kast and R. Dominguez (2014). "Structural analysis of the transitional state of Arp2/3 complex activation by two actin-bound WCAs." Nat Commun **5**: 3308.

Boczkowska, M., G. Rebowski, M. V. Petoukhov, D. B. Hayes, D. I. Svergun and R. Dominguez (2008). "X-ray scattering study of activated Arp2/3 complex with bound actin-WCA." Structure **16**(5): 695-704.

Burke, T. A., A. J. Harker, R. Dominguez and D. R. Kovar (2017). "The bacterial virulence factors VopL and VopF nucleate actin from the pointed end." J Cell Biol **216**(5): 1267-1276.

Campellone, K. G. and M. D. Welch (2010). "A nucleator arms race: cellular control of actin assembly." Nat Rev Mol Cell Biol **11**(4): 237-251.

Carlier, M. F., C. Husson, L. Renault and D. Didry (2011). "Control of actin assembly by the WH2 domains and their multifunctional tandem repeats in Spire and Cordon-Bleu." Int Rev Cell Mol Biol **290**: 55-85.

Cereghino, J. L. and J. M. Cregg (2000). "Heterologous protein expression in the methylotrophic yeast Pichia pastoris." FEMS Microbiol Rev **24**(1): 45-66.

Chen, C. Y., C. H. Cheng, Y. C. Chen, J. C. Lee, S. H. Chou, W. Huang and W. J. Chuang (2006). "Preparation of amino-acid-type selective isotope labeling of protein expressed in Pichia pastoris." Proteins **62**(1): 279-287.

Chereau, D., F. Kerff, P. Graceffa, Z. Grabarek, K. Langsetmo and R. Dominguez (2005). "Actin-bound structures of Wiskott-Aldrich syndrome protein (WASP)-homology domain 2 and the implications for filament assembly." Proc Natl Acad Sci U S A **102**(46): 16644-16649.

Chesarone, M. A., A. G. DuPage and B. L. Goode (2010). "Unleashing formins to remodel the actin and microtubule cytoskeletons." Nat Rev Mol Cell Biol **11**(1): 62-74.

Chhabra, E. S. and H. N. Higgs (2006). "INF2 Is a WASP homology 2 motif-containing formin that severs actin filaments and accelerates both polymerization and depolymerization." J Biol Chem **281**(36): 26754-26767.

Chik, J. K., U. Lindberg and C. E. Schutt (1996). "The structure of an open state of beta-actin at 2.65 A resolution." J Mol Biol **263**(4): 607-623.

Clark, L., J. A. Zahm, R. Ali, M. Kukula, L. Bian, S. M. Patrie, K. H. Gardner, M. K. Rosen and D. M. Rosenbaum (2015). "Methyl labeling and TROSY NMR spectroscopy of proteins expressed in the eukaryote Pichia pastoris." J Biomol NMR **62**(3): 239-245.

Dalhaimer, P., T. D. Pollard and B. J. Nolen (2008). "Nucleotide-mediated conformational changes of monomeric actin and Arp3 studied by molecular dynamics simulations." J Mol Biol **376**(1): 166-183.

Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." Nucleic Acids Res **35**(Web Server issue): W375-383.

Delaglio, F., S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax (1995). "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." J Biomol NMR **6**(3): 277-293.

Dominguez, R. (2007). "The beta-thymosin/WH2 fold: multifunctionality and structure." Ann N Y Acad Sci **1112**: 86-94.

Dominguez, R. (2009). "Actin filament nucleation and elongation factors--structure-function relationships." Crit Rev Biochem Mol Biol **44**(6): 351-366.

Dominguez, R. (2016). "The WH2 Domain and Actin Nucleation: Necessary but Insufficient." Trends Biochem Sci **41**(6): 478-490.
Dominguez, R. and K. C. Holmes (2011). "Actin structure and function." Annu Rev Biophys **40**: 169-186.
Durer, Z. A., D. S. Kudryashov, M. R. Sawaya, C. Altenbach, W. Hubbell and E. Reisler (2012). "Structural states and dynamics of the D-loop in actin." Biophys J **103**(5): 930-939.
Emsley, P., B. Lohkamp, W. G. Scott and K. Cowtan (2010). "Features and development of Coot." Acta Crystallogr D Biol Crystallogr **66**(Pt 4): 486-501.
Evangelista, M., D. Pruyne, D. C. Amberg, C. Boone and A. Bretscher (2002). "Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast." Nat Cell Biol **4**(3): 260-269.
Fan, Y., L. Shi, V. Ladizhansky and L. S. Brown (2011). "Uniform isotope labeling of a eukaryotic seven-transmembrane helical protein in yeast enables high-resolution solid-state NMR studies in the lipid environment." J Biomol NMR **49**(2): 151-161.
Fujii, T., A. H. Iwane, T. Yanagida and K. Namba (2010). "Direct visualization of secondary structures of F-actin by electron cryomicroscopy." Nature **467**(7316): 724-728.
Galkin, V. E., A. Orlova, G. F. Schroder and E. H. Egelman (2010). "Structural polymorphism in F-actin." Nat Struct Mol Biol **17**(11): 1318-1323.
Galkin, V. E., A. Orlova, M. R. Vos, G. F. Schroder and E. H. Egelman (2015). "Near-atomic resolution for one state of F-actin." Structure **23**(1): 173-182.
Gardner, K. H. and L. E. Kay (1997). "Production and incorporation of N-15, C-13, H-2 (H-1-delta 1 methyl) isoleucine into proteins for multidimensional NMR studies." Journal of the American Chemical Society **119**(32): 7599-7600.
Gardner, K. H., M. K. Rosen and L. E. Kay (1997). "Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR." Biochemistry **36**(6): 1389-1401.
Gardner, K. H., X. C. Zhang, K. Gehring and L. E. Kay (1998). "Solution NMR studies of a 42 KDa Escherichia coli maltose binding protein beta-cyclodextrin complex: Chemical shift assignments and analysis." Journal of the American Chemical Society **120**(45): 11738-11748.
Gasman, S., Y. Kalaidzidis and M. Zerial (2003). "RhoD regulates endosome dynamics through Diaphanous-related Formin and Src tyrosine kinase." Nat Cell Biol **5**(3): 195-204.
Goley, E. D. and M. D. Welch (2006). "The ARP2/3 complex: an actin nucleator comes of age." Nat Rev Mol Cell Biol **7**(10): 713-726.
Goode, B. L. and M. J. Eck (2007). "Mechanism and function of formins in the control of actin assembly." Annu Rev Biochem **76**: 593-627.

Gossert, A. D., A. Hinniger, S. Gutmann, W. Jahnke, A. Strauss and C. Fernandez (2011). "A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility." J Biomol NMR **51**(4): 449-456.

Goto, N. K., K. H. Gardner, G. A. Mueller, R. C. Willis and L. E. Kay (1999). "A robust and cost-effective method for the production of Val, Leu, Ile (delta 1) methyl-protonated 15N-, 13C-, 2H-labeled proteins." J Biomol NMR **13**(4): 369-374.

Gould, C. J., S. Maiti, A. Michelot, B. R. Graziano, L. Blanchoin and B. L. Goode (2011). "The formin DAD domain plays dual roles in autoinhibition and actin nucleation." Curr Biol **21**(5): 384-390.

Gournier, H., E. D. Goley, H. Niederstrasser, T. Trinh and M. D. Welch (2001). "Reconstitution of human Arp2/3 complex reveals critical roles of individual subunits in complex structure and activity." Mol Cell **8**(5): 1041-1052.

Graceffa, P. and R. Dominguez (2003). "Crystal structure of monomeric actin in the ATP state. Structural basis of nucleotide-dependent actin dynamics." J Biol Chem **278**(36): 34172-34180.

Griffey, R. H., C. D. Poulter, A. Bax, B. L. Hawkins, Z. Yamaizumi and S. Nishimura (1983). "Multiple quantum two-dimensional 1H--15N nuclear magnetic resonance spectroscopy: chemical shift correlation maps for exchangeable imino protons of Escherichia coli tRNAMetf in water." Proc Natl Acad Sci U S A **80**(19): 5895-5897.

Haglund, C. M. and M. D. Welch (2011). "Pathogens and polymers: microbe-host interactions illuminate the cytoskeleton." J Cell Biol **195**(1): 7-17.

Halabi, N., O. Rivoire, S. Leibler and R. Ranganathan (2009). "Protein sectors: evolutionary units of three-dimensional structure." Cell **138**(4): 774-786.

Hansen, A. P., A. M. Petros, A. P. Mazar, T. M. Pederson, A. Rueter and S. W. Fesik (1992). "A practical method for uniform isotopic labeling of recombinant proteins in mammalian cells." Biochemistry **31**(51): 12713-12718.

Heimsath, E. G., Jr. and H. N. Higgs (2012). "The C terminus of formin FMNL3 accelerates actin polymerization and contains a WH2 domain-like sequence that binds both monomers and filament barbed ends." J Biol Chem **287**(5): 3087-3098.

Hertzog, M., C. van Heijenoort, D. Didry, M. Gaudier, J. Coutant, B. Gigant, G. Didelot, T. Preat, M. Knossow, E. Guittet and M. F. Carlier (2004). "The beta-thymosin/WH2 domain; structural basis for the switch from inhibition to promotion of actin assembly." Cell **117**(5): 611-623.

Hino, T., T. Arakawa, H. Iwanari, T. Yurugi-Kobayashi, C. Ikeda-Suno, Y. Nakada-Nakura, O. Kusano-Arai, S. Weyand, T. Shimamura, N. Nomura, A. D. Cameron, T. Kobayashi, T. Hamakubo, S. Iwata and T. Murata

(2012). "G-protein-coupled receptor inactivation by an allosteric inverse-agonist antibody." Nature **482**(7384): 237-240.

Holmes, K. C., D. Popp, W. Gebhard and W. Kabsch (1990). "Atomic model of the actin filament." Nature **347**(6288): 44-49.

Hung, R. J., C. W. Pak and J. R. Terman (2011). "Direct redox regulation of F-actin assembly and disassembly by Mical." Science **334**(6063): 1710-1713.

Irobi, E., A. H. Aguda, M. Larsson, C. Guerin, H. L. Yin, L. D. Burtnick, L. Blanchoin and R. C. Robinson (2004). "Structural basis of actin sequestration by thymosin-beta4: implications for WH2 proteins." EMBO J **23**(18): 3599-3608.

Joel, P. B., P. M. Fagnant and K. M. Trybus (2004). "Expression of a nonpolymerizable actin mutant in Sf9 cells." Biochemistry **43**(36): 11554-11559.

Kabsch, W., H. G. Mannherz, D. Suck, E. F. Pai and K. C. Holmes (1990). "Atomic structure of the actin:DNase I complex." Nature **347**(6288): 37-44.

Kaiser, D. A., V. K. Vinson, D. B. Murphy and T. D. Pollard (1999). "Profilin is predominantly associated with monomeric actin in Acanthamoeba." J Cell Sci **112 (Pt 21)**: 3779-3790.

Kelleher, J. F., S. J. Atkinson and T. D. Pollard (1995). "Sequences, structural models, and cellular localization of the actin-related proteins Arp2 and Arp3 from Acanthamoeba." J Cell Biol **131**(2): 385-397.

Kim, A. S., L. T. Kakalis, N. Abdul-Manan, G. A. Liu and M. K. Rosen (2000). "Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein." Nature **404**(6774): 151-158.

Kofuku, Y., T. Ueda, J. Okude, Y. Shiraishi, K. Kondo, T. Mizumura, S. Suzuki and I. Shimada (2014). "Functional dynamics of deuterated beta2 -adrenergic receptor in lipid bilayers revealed by NMR spectroscopy." Angew Chem Int Ed Engl **53**(49): 13376-13379.

Kudryashov, D. S., E. E. Grintsevich, P. A. Rubenstein and E. Reisler (2010). "A nucleotide state-sensing region on actin." J Biol Chem **285**(33): 25591-25601.

Kudryashov, D. S. and E. Reisler (2013). "ATP and ADP actin states." Biopolymers **99**(4): 245-256.

Kumar S, S. G., Suleski M, Hedges SB. (2017). "Timetree: A resource for Timelines, Timetrees, and Divergence Times." Mol Biol Evol.

Kursula, P., I. Kursula, M. Massimi, Y. H. Song, J. Downer, W. A. Stanley, W. Witke and M. Wilmanns (2008). "High-resolution structural analysis of mammalian profilin 2a complex formation with two physiological ligands: the formin homology 1 domain of mDia1 and the proline-rich domain of VASP." J Mol Biol **375**(1): 270-290.

Lee, J. Y., I. L. Urbatsch, A. E. Senior and S. Wilkens (2002). "Projection structure of P-glycoprotein by electron microscopy. Evidence for a closed conformation of the nucleotide binding domains." J Biol Chem **277**(42): 40125-40131.

Li, F. and H. N. Higgs (2005). "Dissecting requirements for auto-inhibition of actin nucleation by the formin, mDia1." J Biol Chem **280**(8): 6986-6992.

Li, Z., W. Leung, A. Yon, J. Nguyen, V. C. Perez, J. Vu, W. Giang, L. T. Luong, T. Phan, K. A. Salazar, S. R. Gomez, C. Au, F. Xiang, D. W. Thomas, A. H. Franz, J. Lin-Cereghino and G. P. Lin-Cereghino (2010). "Secretion and proteolysis of heterologous proteins fused to the Escherichia coli maltose binding protein in Pichia pastoris." Protein Expr Purif **72**(1): 113-124.

Liverman, A. D., H. C. Cheng, J. E. Trosky, D. W. Leung, M. L. Yarbrough, D. L. Burdette, M. K. Rosen and K. Orth (2007). "Arp2/3-independent assembly of actin by Vibrio type III effector VopL." Proc Natl Acad Sci U S A **104**(43): 17117-17122.

Lu, S. Y., H. Jang, S. Muratcioglu, A. Gursoy, O. Keskin, R. Nussinov and J. Zhang (2016). "Ras Conformational Ensembles, Allostery, and Signaling." Chemical Reviews **116**(11): 6607-6665.

Machesky, L. M., N. B. Cole, B. Moss and T. D. Pollard (1994). "Vaccinia virus expresses a novel profilin with a higher affinity for polyphosphoinositides than actin." Biochemistry **33**(35): 10815-10824.

Maciver, S. K. and A. G. Weeds (1994). "Actophorin preferentially binds monomeric ADP-actin over ATP-bound actin: consequences for cell locomotion." FEBS Lett **347**(2-3): 251-256.

Manseau, L. J. and T. Schupbach (1989). "cappuccino and spire: two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the Drosophila embryo." Genes Dev **3**(9): 1437-1452.

Mass, R. L., R. Zeller, R. P. Woychik, T. F. Vogt and P. Leder (1990). "Disruption of formin-encoding transcripts in two mutant limb deformity alleles." Nature **346**(6287): 853-855.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." J Appl Crystallogr **40**(Pt 4): 658-674.

Miyazawa-Onami, M., K. Takeuchi, T. Takano, T. Sugiki, I. Shimada and H. Takahashi (2013). "Perdeuteration and methyl-selective (1)H, (13)C-labeling by using a Kluyveromyces lactis expression system." J Biomol NMR **57**(3): 297-304.

Morgan, W. D., A. Kragt and J. Feeney (2000). "Expression of deuterium-isotope-labelled protein in the yeast pichia pastoris for NMR studies." J Biomol NMR **17**(4): 337-347.

Mossakowska, M., J. Moraczewska, S. Khaitlina and H. Strzelecka-Golaszewska (1993). "Proteolytic removal of three C-terminal residues of actin alters the monomer-monomer interactions." Biochem J **289 (Pt 3)**: 897-902.

Mullins, R. D. and T. D. Pollard (1999). "Structure and function of the Arp2/3 complex." Curr Opin Struct Biol **9**(2): 244-249.

Mullins, R. D., W. F. Stafford and T. D. Pollard (1997). "Structure, subunit topology, and actin-binding activity of the Arp2/3 complex from Acanthamoeba." J Cell Biol **136**(2): 331-343.

Nair, U. B., P. B. Joel, Q. Wan, S. Lowey, M. A. Rould and K. M. Trybus (2008). "Crystal structures of monomeric actin bound to cytochalasin D." J Mol Biol **384**(4): 848-864.

Namgoong, S., M. Boczkowska, M. J. Glista, J. D. Winkelman, G. Rebowski, D. R. Kovar and R. Dominguez (2011). "Mechanism of actin filament nucleation by Vibrio VopL and implications for tandem W domain nucleation." Nat Struct Mol Biol **18**(9): 1060-1067.

Ngo, K. X., N. Kodera, E. Katayama, T. Ando and T. Q. Uyeda (2015). "Cofilin-induced unidirectional cooperative conformational changes in actin filaments revealed by high-speed atomic force microscopy." Elife **4**.

Noguchi, T. Q., N. Kanzaki, H. Ueno, K. Hirose and T. Q. Uyeda (2007). "A novel system for expressing toxic actin mutants in Dictyostelium and purification and characterization of a dominant lethal yeast actin mutant." J Biol Chem **282**(38): 27721-27727.

Nolen, B. J., R. S. Littlefield and T. D. Pollard (2004). "Crystal structures of actin-related protein 2/3 complex with bound ATP or ADP." Proc Natl Acad Sci U S A **101**(44): 15627-15632.

Nolen, B. J. and T. D. Pollard (2007). "Insights into the influence of nucleotides on actin family proteins from seven structures of Arp2/3 complex." Mol Cell **26**(3): 449-457.

Nygaard, R., Y. Zou, R. O. Dror, T. J. Mildorf, D. H. Arlow, A. Manglik, A. C. Pan, C. W. Liu, J. J. Fung, M. P. Bokoch, F. S. Thian, T. S. Kobilka, D. E. Shaw, L. Mueller, R. S. Prosser and B. K. Kobilka (2013). "The dynamic process of beta(2)-adrenergic receptor activation." Cell **152**(3): 532-542.

O'Shea, E. K., K. J. Lumb and P. S. Kim (1993). "Peptide 'Velcro': design of a heterodimeric coiled coil." Curr Biol **3**(10): 658-667.

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

Okada, K., F. Bartolini, A. M. Deaconescu, J. B. Moseley, Z. Dogic, N. Grigorieff, G. G. Gundersen and B. L. Goode (2010). "Adenomatous polyposis coli protein nucleates actin assembly and synergizes with the formin mDia1." J Cell Biol **189**(7): 1087-1096.

Ollerenshaw, J. E., V. Tugarinov and L. E. Kay (2003). "Methyl TROSY: explanation and experimental verification." Magnetic Resonance in Chemistry **41**(10): 843-852.

Ono, S. (2010). "Dynamic regulation of sarcomeric actin filaments in striated muscle." Cytoskeleton (Hoboken) **67**(11): 677-692.

Otomo, T., C. Otomo, D. R. Tomchick, M. Machius and M. K. Rosen (2005). "Structural basis of Rho GTPase-mediated activation of the formin mDia1." Mol Cell **18**(3): 273-281.

Otomo, T., D. R. Tomchick, C. Otomo, M. Machius and M. K. Rosen (2010). "Crystal structure of the Formin mDia1 in autoinhibited conformation." PLoS One **5**(9).

Otomo, T., D. R. Tomchick, C. Otomo, S. C. Panchal, M. Machius and M. K. Rosen (2005). "Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain." Nature **433**(7025): 488-494.

Otterbein, L. R., P. Graceffa and R. Dominguez (2001). "The crystal structure of uncomplexed actin in the ADP state." Science **293**(5530): 708-711.

Otwinowski, Z. and W. Minor (1997). "Processing of X-ray diffraction data collected in oscillation mode." Methods Enzymol **276**: 307-326.

Padrick, S. B., H. C. Cheng, A. M. Ismail, S. C. Panchal, L. K. Doolittle, S. Kim, B. M. Skehan, J. Umetani, C. A. Brautigam, J. M. Leong and M. K. Rosen (2008). "Hierarchical regulation of WASP/WAVE proteins." Mol Cell **32**(3): 426-438.

Padrick, S. B., L. K. Doolittle, C. A. Brautigam, D. S. King and M. K. Rosen (2011). "Arp2/3 complex is bound and activated by two WASP proteins." Proc Natl Acad Sci U S A **108**(33): E472-479.

Padrick, S. B. and M. K. Rosen (2010). "Physical mechanisms of signal integration by WASP family proteins." Annu Rev Biochem **79**: 707-735.

Paul, A. S. and T. D. Pollard (2009). "Energetic requirements for processive elongation of actin filaments by FH1FH2-formins." J Biol Chem **284**(18): 12533-12540.

Paul, A. S. and T. D. Pollard (2009). "Review of the mechanism of processive actin filament elongation by formins." Cell Motil Cytoskeleton **66**(8): 606-617.

Paunola, E., P. K. Mattila and P. Lappalainen (2002). "WH2 domain: a small, versatile adapter for actin monomers." FEBS Lett **513**(1): 92-97.

Pellegrin, S. and H. Mellor (2005). "The Rho family GTPase Rif induces filopodia through mDia2." Curr Biol **15**(2): 129-133.

Pernier, J., J. Orban, B. S. Avvaru, A. Jegou, G. Romet-Lemonne, B. Guichard and M. F. Carlier (2013). "Dimeric WH2 domains in Vibrio VopF promote actin filament barbed-end uncapping and assisted elongation." Nat Struct Mol Biol **20**(9): 1069-1076.

Pervushin, K., R. Riek, G. Wider and K. Wuthrich (1997). "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution." Proc Natl Acad Sci U S A **94**(23): 12366-12371.

Pickford, A. R. and J. M. O'Leary (2004). "Isotopic labeling of recombinant proteins from the methylotrophic yeast Pichia pastoris." Methods Mol Biol **278**: 17-33.

Pollard, T. D. (2003). "The cytoskeleton, cellular motility and the reductionist agenda." Nature **422**(6933): 741-745.

Pollard, T. D. (2007). "Regulation of actin filament assembly by Arp2/3 complex and formins." Annu Rev Biophys Biomol Struct **36**: 451-477.

Pollard, T. D. and J. A. Cooper (1986). "Actin and actin-binding proteins. A critical evaluation of mechanisms and functions." Annu Rev Biochem **55**: 987-1035.

Pollard, T. D. and J. A. Cooper (2009). "Actin, a central player in cell shape and movement." Science **326**(5957): 1208-1212.

Porta, J. C. and G. E. Borgstahl (2012). "Structural basis for profilin-mediated actin nucleotide exchange." J Mol Biol **418**(1-2): 103-116.

Potterton, E., S. McNicholas, E. Krissinel, K. Cowtan and M. Noble (2002). "The CCP4 molecular-graphics project." Acta Crystallogr D Biol Crystallogr **58**(Pt 11): 1955-1957.

Pruyne, D., M. Evangelista, C. Yang, E. Bi, S. Zigmond, A. Bretscher and C. Boone (2002). "Role of formins in actin assembly: nucleation and barbed-end association." Science **297**(5581): 612-615.

Qualmann, B. and M. M. Kessels (2009). "New players in actin polymerization--WH2-domain-containing actin nucleators." Trends Cell Biol **19**(6): 276-285.

Quinlan, M. E., J. E. Heuser, E. Kerkhoff and R. D. Mullins (2005). "Drosophila Spire is an actin nucleation factor." Nature **433**(7024): 382-388.

Quinlan, M. E., S. Hilgert, A. Bedrossian, R. D. Mullins and E. Kerkhoff (2007). "Regulatory interactions between two actin nucleators, Spire and Cappuccino." J Cell Biol **179**(1): 117-128.

Rebowski, G., S. Namgoong, M. Boczkowska, P. C. Leavis, J. Navaza and R. Dominguez (2010). "Structure of a longitudinal actin dimer assembled by tandem w domains: implications for actin filament nucleation." J Mol Biol **403**(1): 11-23.

Robinson, R. C., K. Turbedsky, D. A. Kaiser, J. B. Marchand, H. N. Higgs, S. Choe and T. D. Pollard (2001). "Crystal structure of Arp2/3 complex." Science **294**(5547): 1679-1684.

Rodal, A. A., O. Sokolova, D. B. Robins, K. M. Daugherty, S. Hippenmeyer, H. Riezman, N. Grigorieff and B. L. Goode (2005). "Conformational changes

in the Arp2/3 complex leading to actin nucleation." Nat Struct Mol Biol **12**(1): 26-31.

Rodnick-Smith, M., Q. Luan, S. L. Liu and B. J. Nolen (2016). "Role and structural mechanism of WASP-triggered conformational changes in branched actin filament nucleation by Arp2/3 complex." Proc Natl Acad Sci U S A **113**(27): E3834-3843.

Rogers, S. L., U. Wiedemann, N. Stuurman and R. D. Vale (2003). "Molecular requirements for actin-based lamella formation in Drosophila S2 cells." J Cell Biol **162**(6): 1079-1088.

Rosen, M. K., K. H. Gardner, R. C. Willis, W. E. Parris, T. Pawson and L. E. Kay (1996). "Selective methyl group protonation of perdeuterated proteins." J Mol Biol **263**(5): 627-636.

Rosenzweig, R. and L. E. Kay (2014). "Bringing dynamic molecular machines into focus by methyl-TROSY NMR." Annu Rev Biochem **83**: 291-315.

Rouiller, I., X. P. Xu, K. J. Amann, C. Egile, S. Nickell, D. Nicastro, R. Li, T. D. Pollard, N. Volkmann and D. Hanein (2008). "The structural basis of actin filament branching by the Arp2/3 complex." J Cell Biol **180**(5): 887-895.

Rould, M. A., Q. Wan, P. B. Joel, S. Lowey and K. M. Trybus (2006). "Crystal structures of expressed non-polymerizable monomeric actin in the ADP and ATP states." J Biol Chem **281**(42): 31909-31919.

Ruschak, A. M. and L. E. Kay (2010). "Methyl groups as probes of supra-molecular structure, dynamics and function." Journal of Biomolecular Nmr **46**(1): 75-87.

Sagot, I., S. K. Klee and D. Pellman (2002). "Yeast formins regulate cell polarity by controlling the assembly of actin cables." Nat Cell Biol **4**(1): 42-50.

Schmid, M. F., M. B. Sherman, P. Matsudaira and W. Chiu (2004). "Structure of the acrosomal bundle." Nature **431**(7004): 104-107.

Schutt, C. E., J. C. Myslik, M. D. Rozycki, N. C. Goonesekere and U. Lindberg (1993). "The structure of crystalline profilin-beta-actin." Nature **365**(6449): 810-816.

Schwyter, D., M. Phillips and E. Reisler (1989). "Subtilisin-cleaved actin: polymerization and interaction with myosin subfragment 1." Biochemistry **28**(14): 5889-5895.

Sept, D. and J. A. McCammon (2001). "Thermodynamics and kinetics of actin filament nucleation." Biophys J **81**(2): 667-674.

Sheterline, P., J. Clayton and J. C. Sparrow (1995). "Actin." Protein Profile **2**(1): 1-+.

Shimamura, T., M. Shiroishi, S. Weyand, H. Tsujimoto, G. Winter, V. Katritch, R. Abagyan, V. Cherezov, W. Liu, G. W. Han, T. Kobayashi, R. C. Stevens and S. Iwata (2011). "Structure of the human histamine H1 receptor complex with doxepin." Nature **475**(7354): 65-70.

Skillman, K. M., K. Diraviyam, A. Khan, K. Tang, D. Sept and L. D. Sibley (2011). "Evolutionarily divergent, unstable filamentous actin is essential for gliding motility in apicomplexan parasites." PLoS Pathog **7**(10): e1002280.

Skillman, K. M., C. I. Ma, D. H. Fremont, K. Diraviyam, J. A. Cooper, D. Sept and L. D. Sibley (2013). "The unusual dynamics of parasite actin result from isodesmic polymerization." Nat Commun **4**: 2285.

Smith, B. A., K. Daugherty-Clarke, B. L. Goode and J. Gelles (2013). "Pathway of actin filament branch formation by Arp2/3 complex revealed by single-molecule imaging." Proc Natl Acad Sci U S A **110**(4): 1285-1290.

Smith, B. A., S. B. Padrick, L. K. Doolittle, K. Daugherty-Clarke, I. R. Correa, Jr., M. Q. Xu, B. L. Goode, M. K. Rosen and J. Gelles (2013). "Three-color single molecule imaging shows WASP detachment from Arp2/3 complex triggers actin filament branch formation." Elife **2**: e01008.

Smock, R. G., O. Rivoire, W. P. Russ, J. F. Swain, S. Leibler, R. Ranganathan and L. M. Gierasch (2010). "An interdomain sector mediating allostery in Hsp70 molecular chaperones." Mol Syst Biol **6**: 414.

Splettstoesser, T., F. Noe, T. Oda and J. C. Smith (2009). "Nucleotide-dependence of G-actin conformation from multiple molecular dynamics simulations and observation of a putatively polymerization-competent superclosed state." Proteins **76**(2): 353-364.

Strauss, A., F. Bitsch, G. Fendrich, P. Graff, R. Knecht, B. Meyhack and W. Jahnke (2005). "Efficient uniform isotope labeling of Abl kinase expressed in Baculovirus-infected insect cells." J Biomol NMR **31**(4): 343-349.

Strugatsky, D., K. E. Gottschalk, R. Goldshleger, E. Bibi and S. J. Karlish (2003). "Expression of Na+,K+-ATPase in Pichia pastoris: analysis of wild type and D369N mutant proteins by Fe2+-catalyzed oxidative cleavage and molecular modeling." J Biol Chem **278**(46): 46064-46073.

Strzelecka-Golaszewska, H., M. Mossakowska, A. Wozniak, J. Moraczewska and H. Nakayama (1995). "Long-range conformational effects of proteolytic removal of the last three residues of actin." Biochem J **307 (Pt 2)**: 527-534.

Szent-Gyorgyi, A. G. (2004). "The early history of the biochemistry of muscle contraction." J Gen Physiol **123**(6): 631-641.

Thompson, M. E., E. G. Heimsath, T. J. Gauvin, H. N. Higgs and F. J. Kull (2013). "FMNL3 FH2-actin structure gives insight into formin-mediated actin nucleation and elongation." Nat Struct Mol Biol **20**(1): 111-118.

Ti, S. C., C. T. Jurgenson, B. J. Nolen and T. D. Pollard (2011). "Structural and biochemical characterization of two binding sites for nucleation-promoting factor WASp-VCA on Arp2/3 complex." Proc Natl Acad Sci U S A **108**(33): E463-471.

Ti, S. C. and T. D. Pollard (2011). "Purification of actin from fission yeast Schizosaccharomyces pombe and characterization of functional differences from muscle actin." J Biol Chem **286**(7): 5784-5792.

Verkhusha, V. V., M. M. Shavlovsky, O. V. Nevzglyadova, A. A. Gaivoronsky, A. V. Artemov, O. V. Stepanenko, I. M. Kuznetsova and K. K. Turoverov (2003). "Expression of recombinant GFP-actin fusion protein in the methylotrophic yeast Pichia pastoris." FEMS Yeast Res **3**(1): 105-111.

Vizcarra, C. L., B. Kreutz, A. A. Rodal, A. V. Toms, J. Lu, W. Zheng, M. E. Quinlan and M. J. Eck (2011). "Structure and function of the interacting domains of Spire and Fmn-family formins." Proc Natl Acad Sci U S A **108**(29): 11884-11889.

Volkmann, N., C. Page, R. Li and D. Hanein (2014). "Three-dimensional reconstructions of actin filaments capped by Arp2/3 complex." Eur J Cell Biol **93**(5-6): 179-183.

von der Ecken, J., M. Muller, W. Lehman, D. J. Manstein, P. A. Penczek and S. Raunser (2015). "Structure of the F-actin-tropomyosin complex." Nature **519**(7541): 114-117.

Vranken, W. F., W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue (2005). "The CCPN data model for NMR spectroscopy: development of a software pipeline." Proteins **59**(4): 687-696.

Wegner, A. and J. Engel (1975). "Kinetics of the cooperative association of actin to actin filaments." Biophys Chem **3**(3): 215-225.

Werner, K., C. Richter, J. Klein-Seetharaman and H. Schwalbe (2008). "Isotope labeling of mammalian GPCRs in HEK293 cells and characterization of the C-terminus of bovine rhodopsin by high resolution liquid NMR spectroscopy." J Biomol NMR **40**(1): 49-53.

Wider, G. and K. Wuthrich (1999). "NMR spectroscopy of large molecules and multimolecular assemblies in solution." Curr Opin Struct Biol **9**(5): 594-601.

Williamson, M. P. (2014). "Using chemical shift perturbation to characterise ligand binding (vol 73, pg 1, 2013)." Progress in Nuclear Magnetic Resonance Spectroscopy **80**: 64-64.

Winter, D. C., E. Y. Choe and R. Li (1999). "Genetic dissection of the budding yeast Arp2/3 complex: a comparison of the in vivo and structural roles of individual subunits." Proc Natl Acad Sci U S A **96**(13): 7288-7293.

Woychik, R. P., R. L. Maas, R. Zeller, T. F. Vogt and P. Leder (1990). "'Formins': proteins deduced from the alternative transcripts of the limb deformity gene." Nature **346**(6287): 850-853.

Xu, X. P., I. Rouiller, B. D. Slaughter, C. Egile, E. Kim, J. R. Unruh, X. Fan, T. D. Pollard, R. Li, D. Hanein and N. Volkmann (2012). "Three-dimensional

reconstructions of Arp2/3 complex with bound nucleation promoting factors." EMBO J **31**(1): 236-247.

Xu, Y., M. Liu, P. J. Simpson, R. Isaacson, E. Cota, J. Marchant, D. Yang, X. Zhang, P. Freemont and S. Matthews (2009). "Automated assignment in selectively methyl-labeled proteins." J Am Chem Soc **131**(27): 9480-9481.

Xu, Y., J. B. Moseley, I. Sagot, F. Poy, D. Pellman, B. L. Goode and M. J. Eck (2004). "Crystal structures of a Formin Homology-2 domain reveal a tethered dimer architecture." Cell **116**(5): 711-723.

Yang, C., L. Czech, S. Gerboth, S. Kojima, G. Scita and T. Svitkina (2007). "Novel roles of formin mDia2 in lamellipodia and filopodia formation in motile cells." PLoS Biol **5**(11): e317.

Yeung, P. S. and K. J. Boor (2004). "Epidemiology, pathogenesis, and prevention of foodborne Vibrio parahaemolyticus infections." Foodborne Pathog Dis **1**(2): 74-88.

Yu, B., H. C. Cheng, C. A. Brautigam, D. R. Tomchick and M. K. Rosen (2011). "Mechanism of actin filament nucleation by the bacterial effector VopL." Nat Struct Mol Biol **18**(9): 1068-1074.

Zahm, J. A., S. B. Padrick, Z. Chen, C. W. Pak, A. A. Yunus, L. Henry, D. R. Tomchick, Z. Chen and M. K. Rosen (2013). "The bacterial effector VopL organizes actin into filament-like structures." Cell **155**(2): 423-434.

Zheng, X., K. Diraviyam and D. Sept (2007). "Nucleotide effects on the structure and dynamics of actin." Biophys J **93**(4): 1277-1283.

Zhuravleva, A., E. M. Clerico and L. M. Gierasch (2012). "An interdomain energetic tug-of-war creates the allosterically active state in Hsp70 molecular chaperones." Cell **151**(6): 1296-1307.

Zigmond, S. H., M. Evangelista, C. Boone, C. Yang, A. C. Dar, F. Sicheri, J. Forkey and M. Pring (2003). "Formin leaky cap allows elongation in the presence of tight capping proteins." Curr Biol **13**(20): 1820-1823.