Structural Study of the WH2 Family and Filamin: Implications for Actin Cytoskeleton Regulation

ADELEKE H. AGUDA
Dissertation presented at Uppsala University to be publicly examined in C10:301, BMC, Hursagata 3, SE 751 23, Uppsala, Wednesday, November 8, 2006 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

Cellular processes like motility, chemotaxis, phagocytosis and morphogenesis are dependent on the dynamic regulation of the actin cytoskeleton. This cytoskeleton system is tightly controlled by a number of diverse actin-binding proteins (ABPs) by various mechanisms described as nucleation, polymerization, capping, severing, depolymerization and sequestration. The ABPs are grouped based on sequence identity as in the Wiskott-Aldrich Syndrome protein homology domain 2 (WH2), and the calponin homology domain (CH) containing proteins.

In this work, we elucidate the crystal structures of hybrids of gelsolin domain 1 with thymosin β4, ciboulot domain 2, and the second WH2 domain of N-WASP each bound to actin. We show that the single WH2 motif containing protein thymosin β4 in part sequesters actin by binding its pointed end via a C-terminal helix. This interaction prevents the addition of bound actin protomers to the barbed end of the filament. We propose that sequence variations in some WH2 motifs conferred F-actin binding ability to multiple repeat-containing proteins. These F-actin binding domains interact with the barbed end of a filament and the adjacent WH2 motifs are then freed to add their bound actin to the growing filament end. We demonstrate the binding of ciboulot domains 2 and 3 to both G- and F-actin and that full length ciboulot is capable of binding two actin monomers simultaneously.

We have also cloned, expressed, purified and crystallized rod domains 14-16 from the actin crosslinking protein α-filamin. Preliminary X-ray crystallography data gives us hope that we shall be able to solve the structure of this triple domain repeat.

Keywords: Actin, Thymosin, Ciboulot, N-WASP, Filamin, Calponin homology domain, WH2, Protein Crystallography

Adeleke H. Aguda. Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden

© Adeleke H. Aguda 2006

ISBN 1651-6206
ISBN 91-554-6679-6
urn:nbn:se:uu:diva-7188 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7188)
Figure 1. Model of Thymosin β4 bound to actin
A bit beyond perception’s reach
    I sometimes believe I see
That life is two locked boxes,
Each containing the other’s key

Piet Hein (1905-1996).

It is not the possession of truth but the success which attends the seeking
after it, that enriches the seeker and brings happiness to him

Max Planck (1858-1947).
To my mother

Iya ni wura
Publications

This thesis is based on the following publications: hereafter referred to by their roman numerals.

I  Edward Irobi, Adeleke H. Aguda, Mårten Larsson, Christophe Guerin, Helen L. Yin, Leslie D. Burtnick, Laurent Blanchoin, and Robert C. Robinson


II  Adeleke H. Aguda, Bo Xue, Edward Irobi, Thomas Préat, and Robert C. Robinson

The structural basis of actin interaction with multiple WH2/βthymosin motif-containing proteins. *Structure.* 2006; 14, 469-76.

III  Adeleke H. Aguda, Amos M. Sakwe, Lars Rask, and Robert C. Robinson

Structural study of α-filamin repeats 14-16. *Submitted.*

All reprints made with the permission of the publishers
Other Publications not included in the thesis


Contents

Introduction ........................................................................................................................................... 11
Background .......................................................................................................................................... 12
Methods ................................................................................................................................................ 23
Aim ....................................................................................................................................................... 27
Present Investigations ......................................................................................................................... 28
Paper I .................................................................................................................................................. 29
Paper II ................................................................................................................................................ 32
Paper III ............................................................................................................................................... 34
Discussion and future perspective ................................................................................................. 35
Conclusions ......................................................................................................................................... 38
Acknowledgements ............................................................................................................................ 39
References ........................................................................................................................................... 41
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABD</td>
<td>Actin binding domain</td>
</tr>
<tr>
<td>ABPs</td>
<td>Actin binding proteins</td>
</tr>
<tr>
<td>ACT</td>
<td>Actobindin</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine di-phosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine tri-phosphate</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CH</td>
<td>Calponin homology domain</td>
</tr>
<tr>
<td>CIB</td>
<td>Ciboulot</td>
</tr>
<tr>
<td>Cib 1</td>
<td>Ciboulot domains 1</td>
</tr>
<tr>
<td>Cib23</td>
<td>Ciboulot domains 2 and 3</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EM</td>
<td>Electron microscopy</td>
</tr>
<tr>
<td>G1</td>
<td>Gelsolin domain 1</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-beta-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>MIM</td>
<td>Missing in metastasis</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>Nw2</td>
<td>N-WASP WH2 domain 2</td>
</tr>
<tr>
<td>N-WASP</td>
<td>Neural Wiskott-Aldrich Syndrome protein</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>TTβ</td>
<td>TetraThymosin β</td>
</tr>
<tr>
<td>Tβ4</td>
<td>Thymosin β4</td>
</tr>
<tr>
<td>VASP</td>
<td>Vasodilator-stimulated phosphoprotein</td>
</tr>
<tr>
<td>WASP</td>
<td>Wiskott-Aldrich Syndrome protein</td>
</tr>
<tr>
<td>WH2</td>
<td>Wiskott-Aldrich Syndrome protein homology domain 2</td>
</tr>
</tbody>
</table>
Introduction

As far back as the 1940’s, fibers extracted from muscle tissue were found to consist of two distinct proteins, namely, myosin and its binding partner actin [1]. Series of investigations were conducted to characterize their interactions using various microscopy techniques which eventually yielded the now well known sliding mechanism of actin and myosin filaments in muscle contraction. It was not until the 1970’s that actin was first identified by Lazarides and Weber as a non-muscle cell protein [2-4]. Today it is well recognized that actin is a major component of a microfilament system which along with intermediate filaments and the microtubules form what is collectively described as the cytoskeleton.

The cytoskeleton system occurs throughout the plant and animal kingdoms and similar systems have been described in microorganisms. This complex network of filament systems is responsible for maintaining eukaryotic cell structure and shape [5]. It provides a framework for many cellular processes such as motility, phagocytosis, tissue morphogenesis, cell anchorage and chemotaxis [6-8]. It also forms a major portion of total cellular protein reaching up to 80% in some cells and consists of many different proteins co-operating in the organization of the dynamics of cells [9].

Knowledge about the cytoskeleton systems has increased tremendously following more than fifty years of research and many of the participating proteins are now known. The focus of this thesis will be on the actin filament system and in particular two groups of actin binding proteins described as the Wiskott-Aldrich Syndrome protein homology domain 2 (WH2) proteins and the filamins. These groups constitute a good representation of actin sequestering/polymerizing proteins (WH2) and filament crosslinking proteins (filamin). They consequently play a crucial role in regulating actin dynamics at the membrane of healthy cells and their impairment leads to disease states [10, 11].
Background

Actin

Actin is a highly conserved protein with few N-terminal amino acid sequence differences between species that are distantly separated on the evolutionary tree. Several isoforms exist in multi-cellular organisms: three groups (α, β, and γ) in humans, two in *Drosophila melanogaster* and several more in plants [12]. In humans, several groups can be expressed in the same cell, usually with one group predominant over the others [13, 14].

The basic unit is the actin monomer (G-actin), a single polypeptide chain of 375 amino acids. This 42 kDa protein is composed of two lobes (colored yellow and blue in Fig 2A) each of which can be further divided into two sub-domains. Sandwiched between the lobes is a nucleotide binding pocket that contains the ADP or ATP necessary to maintain its conformation [12, 15] and a divalent cation (Ca$^{2+}$ or Mg$^{2+}$, painted green in Fig 2A). The interface between sub-domains 1 and 3 is rich in hydrophobic residues which creates a hydrophobic pocket (Fig 2A) identified as a major surface for protein interactions with actin [16]. Actin is known to be involved in more interactions than any other protein, being able to self assemble as well as to assemble with other proteins. The interacting partners are collectively referred to as actin binding proteins (ABPs) which number over 162 [17] with new classes being identified annually. Under specific conditions (as in elevated salt concentrations), G-actin can assemble into ordered filaments (F-actin), a process that is strictly regulated in-vivo by the various ABPs. The filament is a polar right hand double helix and the two ends are referred to as the fast growing (barbed) and the slow growing (pointed) ends (Fig 2B). In the presence of ATP, actin assembly (polymerization) at the barbed end is ten times faster than at the pointed end, but equivalent in the presence of ADP. It is the coordinated assembly and disassembly of the filaments that account for the dynamic functions of the actin cytoskeleton.

Crystal structure of actin

The ability to polymerize has been a major drawback to the determination of the atomic structure of uncomplexed actin. The first structures of ABP-bound actin were resolved in the 90’s [18-20] and these revealed the importance of the DNase I binding loop and the hydrophobic pocket. These structures also highlighted the possible inter domain movement in actin as sug-
gested by the unusual ‘open’ conformation observed for the actin profilin complex. However, several questions still remained unanswered; Such as what does uncomplexed actin really look like? And, how much had its native conformation been altered by these bound ABPs.

Many more actin structures have been resolved since, most have been bound to either ABPs [21-23] or to non protein actin sequestering molecules [24-27]. Only recently, has the first structure of an uncomplexed actin been solved howbeit, a modified form of actin [28]. This structure however shows no remarkable difference to previously published data.

Contrary to G-actin, there is no crystal structure of filamentous actin but numerous models have been proposed based on fiber diffraction, electron microscopy (EM) reconstruction and crystal contact analysis [29-32]. Among these includes the Egelman, Schutt-Lindberg and Holmes-Lorenz models (Fig 2B).

![Figure 2. Models of actin. A) Uncomplexed actin. B) Modified Holmes model of the filament.](image)

**ATP hydrolysis**

Studies on the various ABPs have revealed that affinity for actin binding varies depending on the nucleotide in the cleft. Actin polymerizing proteins (arp2/3, VASP) tend to have a higher affinity for ATP-actin, while depoly-
merizing proteins (ADF/cofilin, gelsolin) bind tighter to ADP-actin. This has been partly attributed to the hydrolysis of ATP to ADP in the actin nucleotide cleft which serves as a timer for ABPs to pull apart the filament. Though this process of hydrolysis is largely unclear, it is speculated that the ATP form is responsible for the ‘closed’ conformation of actin. Two residues: Gly-158 and Ser-14 in the cleft interact with the $\gamma$-phosphate of ATP and this position is transmitted to the rest of the actin molecule through a network of side chains. This serves to bring the two halves of actin into this conformation [33]. In the ADP form, the $\gamma$-phosphate interaction does not exist and the two residues (14 and 158) consequently move apart creating the ‘open’ relative conformation of the actin halves [20].

This switch in conformation from closed to open states in the actin subunits consequently destabilizes the filament and triggers a host of depolymerizing ABPs to bind the filament [34, 35].

This conformation change is also seen in other NTPases where conserved Gly and Ser residues form hydrogen bonds with the $\gamma$-phosphate to close the nucleotide pocket [35]. Arp2 has been demonstrated to adopt the same mechanism. The crystal structure of the nucleotide free arp2/3 complex is in the open state was speculated to close in the activated state [36]. The switch to the closed conformation has been implicated in debranching of the filament.

**Regulation of actin function**

Actin filaments can undergo a range of activities, including: nucleation, polymerization, branching, cross-linking, bundling, debranching, capping, fragmenting and depolymerization. These dynamic processes are strictly regulated by the diverse ABPs many of which bind on the same loci on actin. These proteins are grouped based on their functions as listed below:

Nucleating e.g. Arp2/3 complex, formins, spire, gelsolin [36-40]
Polymerizing e.g. Profilin, ciboulot, actobindin, MIM [41-44]
Severing e.g. Gelsolin, villin, and cofilin [45-47]
Depolymerizing e.g. ADF/cofilin, twinfilin [48-50]
Capping e.g. Gelsolin, capG, capping protein and kaptin [12, 51-53]
Sequestering e.g. Thymosin $\beta$4 (Fig 1), DNase1 and profilin [54-56]
Crosslinking e.g. Filamin, $\alpha$-actinin, fimbrin [57-60]
Bundling e.g. Espin, Vasp [61-63].

An alternate classification groups the ABPs into families based on sequence similarity. These groups include for example the gelsolin-villin family, the thymosin$\beta$/WH2 family, the calponin homology domain (CH) containing proteins and the actin related proteins (Arp).
Actin dynamics at the cell membrane

In a ‘resting’ state, actin filaments undergo slow addition at the barbed end and depolymerization at the pointed end in vitro, a phenomenon known as ‘treadmilling’ [34]. This effect is enhanced in the presence of actin binding proteins such as profilin and halted in the presence of barbed end capping proteins. In vivo, these barbed ends are directed towards the membrane. In the presence of an appropriate stimulus, phosphoinositides (like PIP2) at the membrane edge bind to, and remove the barbed-end caps allowing polymerizing ABPs like profilin to add ATP-actin to this ‘growing’ end of the filament [64, 65]. WASP activates the arp2/3 complex to nucleate a new filament from the side of the existing filament thereby multiplying the free barbed ends which undergo elongation (Fig 3) [66, 67]. The rapidly elongating filaments generate a forward thrust at the membrane edge. Crosslinking of the filaments via proteins like α-actinin and filamin give additional strength to this meshwork [59]. ATP hydrolysis in older actin units renders the filament pointed ends unstable and severing proteins like gelsolin then rapidly fragment these into smaller pieces which depolymerizing proteins then recognize. ADF/cofilin can interact with these fragmented filaments and pull it apart into individual monomers [34]. This ADP-actin is then yielded to profilin which facilitates nucleotide exchange from ADP to ATP before repeating the cycle at the barbed end. This cycle ensures a constant supply of ATP-actin to the barbed end for rapid polymerization. With the removal of the polymerizing stimulus, capping at the barbed end by CapG, capping protein and gelsolin takes place. This prevents addition of actin to the filament consequently slowing down polymerization. Profilin in the absence of free barbed ends then exchanges its bound ATP-actin to Tβ4 which sequesters it into a pool of ready-to-use ATP-actin (Fig 1).
Gelsolin
Gelsolin is an 82 kDa protein that plays a crucial role in severing and capping of actin filaments. It is a six-domain protein (G1 to G6) comprising two similar halves (an N-terminal G1-3, and a C-terminal G4-6). G2 is the filament recognition domain [17, 38] and G1 and G4 are the capping domains. G2 recognizes and binds to the side of the filament consequently bringing G1 into close proximity to the protomer directly below. G1 then binds to and caps this protomer severing the filament in the process. Though not a member of the WH2 family of proteins, the ability of G1 to cap and consequently inhibit polymerization makes it an important tool in our research. Furthermore, gelsolin residues 149-152 have sequence similarity to the central region of the WH2 proteins and we utilize this similarity in designing hybrid proteins for the purpose of crystallographic studies.
WH2 Proteins

The WH2 domain is a 35-44 amino acid containing motif that characteristically binds monomeric actin in a 1:1 ratio [68]. In all, 38 proteins are known to contain the WH2 domain in animals and yeast, but none has been identified in plants [69]. The domain has also been identified in baculoviral capsid proteins where it is implicated in the regulation of actin dynamics in the infection process. The most widely studied members of this family are the β-thymosins which have several isoforms, some of which are overexpressed in cancer cells. The group varies from a single WH2 domain-containing protein as in Tβ4 [70, 71] to members having multiple repeats of the domain as in actobindin (ACT) (two repeats) [72], ciboulot (CIB) (three repeats) [73] and tetrathymosin β (TTβ) (four repeats) [74]. A third group includes proteins having one or more WH2 domains in combination with other domain types, such as in missing in metastasis (MIM), VASP and WASP, (all with a single WH2 domain each), N-WASP (two WH2 domains), and SPIRE with four WH2 domains (Fig 4).

Single WH2 domains are known to sequester actin by binding G-actin and preventing its addition to either end of a filament [75]. The most well characterized member of this class, Tβ4 is the main actin sequestering molecule in the cell (existing at concentrations up to 550 μM in platelets) [76] thereby acting as a major reserve pool of actin. The bound actin can be recruited for polymerization via exchange with other ABPs like profilin which then adds it to the fast growing ends of filaments [77]. Proteins with multiple WH2 repeats function quite differently, they do not sequester the bound actin but rather add it to the barbed end of a growing filament in a similar manner to profilin [78]. Not surprisingly, members of the family that have additional protein motifs have other specialized functions like in the activation of the arp2/3 complex in the case of WASP, filament bundling by VASP and nucleation by SPIRE. Recently, a formin (INF2) was shown to exhibit a novel actin filament severing activity with its WH2 motif speculated to play a crucial role [79].

Thymosin β4

This 5 kDa, single WH2 domain containing protein (Fig 4) is the most abundant actin sequestering protein in mammalian cells. It was first isolated from bovine thymus and is present and conserved across mammalian species. It has been associated with carcinogenesis, apoptosis, blood coagulation, angiogenesis and wound healing with high levels (up to 13 μg/ml) of Tβ4 present in human wound fluid [80]. It is present in concentrations in excess of actin monomers with the highest levels attained in platelets (550μM in unactivated platelets) [34] and polymorphonuclear neutrophils (PMNs). Tβ4 directly competes with profilin and DNase I for monomeric actin probably due to overlapping binding locus on actin. However, a ternary complex of Tβ4:actin:DNase I and Tβ4:actin:profilin has been demonstrated in the pres-
ence of cross-linkers [81, 82]. Though primarily sequestering actin, Tβ4 has been shown in vitro to have weak cooperative binding to filaments at concentrations higher than 20 μM [70].

Tβ4 binds ATP-actin in a 1:1 complex and cross-linking experiments have suggested binding via an N-terminal helix and a C-terminal loop to the barbed and pointed ends of actin respectively. Nuclear magnetic resonance (NMR) structures have revealed helices at both ends of the molecule with a connecting unstructured central region.

**Ciboulot**

CIB is the 3 WH2-containing protein in *Drosophila melanogaster* that is involved in the control of brain development during metamorphosis (Fig 4). In contrast to Tβ4, CIB participates in filament barbed end elongation similar to profilin. Indeed profilin has been shown to restore axonal growth in CIB mutant *Drosophila* central brain. In addition to barbed end growth, CIB inhibits pointed end growth and caps the pointed end of filaments in the presence of barbed end cappers like gelsolin.

*Figure 4. Schematic representations of the domain structures of WH2 containing proteins*
A look at the sequence alignment of members of this family of proteins reveals three conserved regions (Fig 5). The N-terminus of Tβ4 has been shown by NMR and circular dichroism (CD) spectroscopy to form an α-helix under non-physiological conditions [83, 84]. This region is relatively well conserved in the family, thereby, suggesting the presence of an N-terminal α-helix in other members of the group. The central LKKTET motif and the C-terminal ends are mostly unstructured and are more variable in their amino acid composition [69, 83]. All three regions contain actin binding sites [85, 86] and the latter two regions have been implicated in the differences in actin binding exhibited by the various members of this group.

Crystallization of complexes of WH2 domains has been hampered by two factors: 1) The WH2 has no structural core and is therefore unusually flexible, and 2) actin polymerizes in the crystallization experiments. We focus our attention on the variable C-terminus in a bid to find the key that will unlock the mysteries surrounding this family of proteins. Hopefully, determination of the mechanisms of action of WH2 domains will give us insight not only for academic, but for therapeutic purposes in viral pathophysiology and in the race against cancer.

**Figure 5.** A sequence alignment of some members of the WH2 family.

---

**Filamins**

Filamins are members of a group of ABPs classified under the calponin homology (CH) domain-containing proteins which also includes spectrin, fimbrin, and α-actinin. Members of this family possess between one and four CH domains (Fig 8) [87]. The role of the CH domain in actin binding is still somewhat controversial with most studies indicating direct actin binding whilst others suggest a more localization function [88-92].

Mammalian and Dictyostelium discoideum filamins have been the most well studied, and crystal structural data have defined specific structural features essential to its crosslinking function.

Filamins are structurally divided into an actin binding (ABD) headpiece and a rod domain composed of repeating Ig-like repeats (Fig 6) [93, 94].
**ABD domain**

The ABD consists of tandem calponin homology (CH) domains which bear strong sequence similarity to other CH domain containing proteins like dystrophin, utrophin and α-actinin [95-99]. This domain is responsible for F-actin binding and the ability of the α-actinin ABD to bind G-actin suggests similar ability in filamin.

There is presently no crystal structure of the ABD of filamin, but high resolution structures of the ABD of utrophin, dystrophin and α-actinin [100-102] have shown that each CH domain forms 6-7 α helices [103-105]. The long terminal helix of the first domain (CH1) connects the tandem domains and its adopted conformation has been proposed to determine the relative spatial orientation of the two domains. The unbound conformation has been described as the 'compact' conformation which is speculated on binding to actin to also adopt the 'extended' filament-bound conformation predicted by EM reconstructions [106, 107]. This switch is proposed to be facilitated by the co-operative actin binding between the two adjacent domains.

![Figure 6. Schematic diagram of filamin and some of its interactions](image)

**Actin binding sites**

By point mutations, three actin binding sites have been identified on the first and last helices of CH1 (S1 and S2 in Fig 7A) and the first helix in CH2 (S3).
S2 has been shown to be crucial, but the CH1 on its own has weak F-actin binding affinity as shown by CH2 deleted constructs. The CH2 though complementing CH1 domain on its own binds F-actin very weakly [103, 111].

Rod Domain

Though the ABD domain is similar to that of other members of this family of proteins, the rod domains vary with filamin adopting immunoglobulin (Ig) like repeats while others adopt a spectrin-like (spectrin, utrophin), calponin-like (calponin), coiled-coil (interaptin) structure, or a mix of these with/without other motifs (Fig 8) [87, 112]. 24 Ig repeats (Fig 7B) are present in human filamins while 20 and 6 are present in Drosophila melanogaster and Dictyostelium discoideum respectively [113-115]. Like many crosslinking proteins, filamins are known to oligomerize by forming dimers which is speculated to be crucial for function [116]. The nature of its dimerization however varies. Dictyostelium gelation factor forms an antiparallel dimer while human filamins adopt a parallel dimer conformation. In both cases, the dimer interface is located in the C-terminal most Ig repeat (6 in dictyostelium and 24 in human) [93, 117]. Each Ig repeat is composed of about 100 residues and the repeats are connected by short loops. A larger
amino acid insertion exists between repeats 15-16 and 23-24 in human-filamins. These forms the proteolysis susceptible hinges 1 and 2, respectively (Fig 6) [118-120].

**Figure 8.** Domain structures of CH containing proteins

Mammalian filamins interact with many proteins (Fig 6) including transmembrane proteins making them an important link between the actin cytoskeleton and the cell membrane [121-123]. It has been demonstrated that filamin provides protection against membrane receptor degradation as well as regulating receptor and co-activator functions in the nucleus [124-126]. Due to the diverse nature of its interactions, mutations in the filamin gene are consequently associated with a number of syndromes including the otopalatodigital (OPD) spectrum disorders [127-132].
Methods

Our goal as protein crystallographers is to grow good quality diffracting crystals for subsequent crystallographic analysis. To achieve this, we isolate pure protein in large enough amounts and modify the protein environment to stimulate an ordered 3-dimensional assembly of identical protein molecules. The methods we adopt in this study include: gene isolation by polymerase chain reaction (PCR), protein expression in *Escherichia coli* strains (*E. coli*), protein purification, crystallization, X-ray diffraction and model building.

**Gene isolation**

To grow good quality crystals for analysis, it is not unusual to modify the properties of the protein of interest. This can be achieved through deletions, mutations or as we adopt in some of our studies, designing hybrids of two related proteins. PCR of a gene from a cDNA and insertion into an appropriate expression vector often bypasses the problem of low yield, low purity, and isolation of multiple isoforms often encountered when purifying proteins from native sources.

The hybrid constructs in this work were designed based on the existence of a region of sequence similarity between gelsolin domain 1 (G1) and the WH2 family. The final protein product was a simple hybrid molecule without introduced extraneous amino acids.

In these studies, the expression vectors used were pHis8-3, which encodes an N-terminal eight histidine tag and pGEX-4T-3 (GE Healthcare, Sweden) which encodes an N-terminal GST tag. Both vectors also encode a thrombin protease cleavage site situated at the N-terminus of the expressed protein.
Protein expression and purification

Expression of the target protein was achieved by insertion of the cloned DNA into the pHis8-3 or pGEX-4T-3 vectors and transformation into a BL21 (DE3) pLysS expression cell strain of *E.coli*. The vector contains a T7-promoter and protein expression was achieved under isopropyl-beta-D-thiogalactopyranoside (IPTG) induction.

Purification of the protein from a homogenized cell lysate was typically initiated by incubating the homogenate in pre-washed coated (nickel or glutathione) beads which bind the affinity tag on the expressed recombinant protein. The tagged protein was eluted after careful washing of the bound beads and in the GST-tagged protein, the tag removed by cleavage with thrombin. The cleaved protein was then subjected to further purification steps such as ion exchange chromatography and/or gel filtration. The columns used in this work were the Resource Q (filamin), Superdex 75 (uncomplexed Tβ4, Cib23 and Nw2), Superdex 200 (all WH2:actin complexes and filamin) and Sephacryl 300 (for actin) all from GE Healthcare, Sweden.

Actin purification was performed from rabbit skeletal muscle acetone powder using a modified version of the Spudich and Watt purification method [133].

Crystallization

Crystallization is the process whereby millions of identical copies of the same molecule are arranged on top and beside each other in an ordered 3-dimensional array. This is essential for any meaningful crystallographic data collection to take place and indeed is often the bottleneck in X-ray crystallography. Several crystallization methods are known, including, the microbatch and dialysis methods, but the most common is based on the principle of vapor diffusion as illustrated by the hanging drop method. The hanging drop experiment consists of a well containing a precipitant solution at an appropriate pH (with or without an additive) and a cover slip on which a drop of a mixture of the protein solution and the precipitant solution (usually in a 1:1 ratio) is placed. The well is then sealed with the cover slip with the aid of grease and incubated at a chosen temperature. Vapor diffusion proceeds from the hanging drop into the reservoir solution thereby increasing the protein and precipitant concentration in the drop until a state of supersaturation is reached. At this state, the protein molecules assemble in an ordered way (nucleate) to form a crystal and continue to grow until the protein in solution is exhausted or the crystal order is lost (Fig 9). However, to achieve crystal growth, many trials and optimization of the conditions are required as proteins generally tend to form disorganized aggregates. We adopted the microbatch and the vapor diffusion methods in our crystallization trials.
Data collection and model building

Prior to data collection, the crystals were flash frozen in liquid nitrogen after incubation in a protective cryo solution consisting of the mother liquor (in the well) supplemented with a variable concentration of glycerol (10-20%). Data were collected at the Max lab synchrotron source in Lund, Sweden, the European Synchrotron Radiation Facility in Grenoble, France and the in-house Rigaku/MSC FR-E superbright X-ray source at Biopolis, Singapore. Processing of images and data merging were carried out with MOSFLM and SCALA (CCP4 1994). The phasing problem was solved by molecular replacement using G1:actin complex [21] (paper I) and G1-Tβ4:actin (paper II) as search models in the AMORE programme. The model obtained was then passed through repeated cycles of refinement in REFMAC5 followed
by cycles of manual refinement in O [134]. A validation of the quality of the final models was done in PROCHECK (CCP4, 1994). Cartoon figures of the solved structures were made using MOLSCRIPT.
The aim of this work is to determine the crystal structure of members of the WH2 family of proteins and a fragment of the crosslinking protein filamin. From their structures, we attempt to decipher how they influence the dynamics of the actin cytoskeleton at the cell membrane.
Present Investigations

**Construct design**

We constructed hybrids of gelsolin domain one with the C-terminal half of Tβ4 (paper I), the C-terminal half ciboulot domain 2-3 and C-terminus of N-WASP domain 2 (paper II) hereafter referred to as G1- Tβ4, G1-Cib23 and G1-Nw2, respectively (Fig 10). The purpose of the gelsolin fragment is to prevent actin from polymerizing and to stabilize the interaction between the WH2 domain with actin.

*Figure 10. Schematic diagram of hybrid constructs*
The structure of the G1- T\(\beta\)4:actin complex (Fig 11) reveals G1 bound to actin in the characteristic manner with the T\(\beta\)4 portion extending up the face of actin ending in an \(\alpha\)-helix that binds above the nucleotide cleft between sub domains 2 and 4 of actin. A model of the full-length T\(\beta\)4 was built by superposing this structure on a ciboulot domain 1 (Cib1): actin structure [78] after removing the G1 portion. This model shows the amphipathic N-terminal helix bound to the barbed end of the actin monomer with the central region extending up its face to bind to the C-terminal helix at the pointed end of the monomer thus preventing the monomer from binding to a filament.

To determine the possibility of a conformational change arising in T\(\beta\)4 as a result of the hybrid design, we superimposed the actin structures and found that the residues of structural overlap within the LKKTET motif between T\(\beta\)4 and Cib1 also superimpose. A stereo view of this superposition (Fig 12) reveals an excellent fit suggesting that the T\(\beta\)4 in the hybrid maintains its functional wild type conformation.
Figure 11. Crystal structure of G1-Tβ4:actin complex and model of Tβ4 bound to actin.
Figure 12. Stereo view of the structural overlap within the LKKET motif in G1-Tβ4 and ciboulot.
In this paper, we examined the structures of two proteins that have multiple WH2 repeats with special focus on the C-terminal half and central regions of Cib23 and N-WASP domain 2.

Cib23 in solution had been previously described as not binding to G-actin [78]. The design of the constructs used in those experiments however excluded half of the N-terminal helices of these domains. These results raised questions about the necessity of triple domains if only one was functional. We therefore made a hybrid of this region with G1 in a similar fashion as in paper I. A hybrid of G1 with Nw2 was also constructed for analysis.

Figure 13. Structure of A) G1-Cib23: actin and G1-Nw2: actin complexes. B) 90° rotation relative to A.
The structure of the G1-Cib23 (yellow green in Fig 13) bound to actin shows the C-terminal of domain 2 binding in an analogous manner as Tβ4. There was however no density for domain 3. The structure of G1-Nw2 on the other hand shows the C-terminus of Nw2 (dark blue) twisting away from actin and is not bound to its surface (Fig 13). The G1-Nw2: actin complex also revealed that actin adopts a more open form than for G1-Cib23.
In this paper we have studied the cross-linking protein filamin for crystallographic analysis. A construct consisting of repeats 14-16 were expressed, purified, crystallized and X-ray data collected. The high resolution data (Table 1) gives us hope that we will be able to solve the structure of this triple repeat.

Table 1. Statistics of preliminary data analysis

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>24.62 – 1.95 (2.06 – 1.95)</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.5418</td>
</tr>
<tr>
<td>Unit cell parameters</td>
<td>a=50.63,</td>
</tr>
<tr>
<td></td>
<td>b=52.10,</td>
</tr>
<tr>
<td></td>
<td>c=98.46,</td>
</tr>
<tr>
<td></td>
<td>α=β=γ=90°</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Observed reflections</td>
<td>204864 (26685)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>18692 (2598)</td>
</tr>
<tr>
<td>R_merge</td>
<td>0.071 (0.330)</td>
</tr>
<tr>
<td>Completeness</td>
<td>95.4 (92.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>11.0 (10.3)</td>
</tr>
</tbody>
</table>
Discussion and future perspective

The crystal structures of the WH2 proteins presented here indicate that this family of proteins generally binds to actin via three regions. The differences in their ability to interact with actin appears to depend on the amino acid composition at their central and C-terminal regions, as the N-termini are well conserved. Members of the group that have the central and C-terminus substituted or deleted show variations in their actin binding ability. This is evidenced in the Nw2 domain where the C-terminal helix is substituted with a connecting and an acidic region (CA), a known activator of the arp2/3 complex that induces branching of the filament. This CA region, due to this modification, has been a source of controversy, with some researchers speculating it does not bind directly to actin. The crystal structure observed in paper II appears to give support to this theory. However, more recent data suggest that the actin binding site of the CA region is closer to the C-terminus and therefore not included in the hybrid design. Though single WH2-motif containing proteins typically sequester actin monomers by capping both ends of the monomer, the presence of multiple motifs results in a switch to barbed end assembly and pointed end capping. Hertzog et al. [78] observed by mutagenesis experiments that the strength of the C-terminal helix is a major determinant of sequestering versus barbed end elongation, and the present structures in this study are in agreement. This is further evident from the fact that some domains in multiple WH2 domain containing proteins are exclusively F-actin binders. By point mutations, Van Troys et al. [74] identified TetraThymosin β (TTβ) domain 3 to be exclusively F-actin binding. We propose that this F-actin binding domain is crucial for the function of TTβ by anchoring it to a growing filament end. Monomers bound to TTβ are consequently aligned along the filament length bringing them into close proximity to the barbed end (Fig 14). The binding of the F-actin domain is expected to cause conformational changes in both its C and N-terminal helices, which may be relayed to the next repeat, dragging its C-terminal helix off. Once these helices come off, the actin-actin contact becomes possible thereby adding the monomer to the filament. This chain of conformational changes continues through to the first domain thereby adding another monomer to the filament and completely releasing the TTβ which can then bind to three other monomers and the cycle is repeated. The TTβ F-actin binding domain can bind at the pointed end; however there is no
mechanism to remove the C-terminal-most capping helix, and filament capping results [74]. Based on the similarity in function between the multidomain containing proteins, we propose the existence of such an F-actin binding domain in ciboulot (either domains 2 or 3) and actobindin and show by cib:actin crosslinking experiments (paper II) that these multidomain proteins bind monomers and interact with a filament via their F-actin binding domains. Furthermore, a model of the ciboulot domains superposed on the actin filament model reveals a longitudinal interaction along each strand of the filament as the only allowable mode of interaction (Fig 5 in paper II). Actobindin has two complete and one half WH2 domain and we propose that the interaction of this half domain with F-actin is responsible for the change in conformation that removes the cap in domain two of this molecule allowing it to elongate a filament.

Figure 14. Schematic representation of multiple WH2 domains with actin filaments

The WH2 proteins studied in this thesis mostly belong to the first two groups of WH2 containing proteins (discussed on page 17). The third group is the proteins with extra non-WH2 protein motifs and these constitute members of the family that have acquired other specialized functions. These molecules represent an exciting challenge for structural biology.

VASP, a member of this third class of proteins has a single WH2 domain sandwiched between a polyproline and an F-actin binding site. This protein which forms a homotetramer has been observed to shield the barbed end of the filament from capping proteins [135, 136]. By so doing it is speculated to interact with the barbed end and rapidly bind profilin-actin via its polyproline region. Profilin then exchanges its actin to the WH2 which adds it to the growing filament barbed end [137]. This results in rapid filament elongation, a process that is further speeded up by the ability of profilin to dimerize on binding to VASP thereby increasing the rate of actin delivery to the barbed end [138]. Several pathogens like *listeria* recruit VASP hence taking advantage of this mechanism during their infective processes [139, 140].
Another group of proteins containing four repeats of the WH2 domain have been identified as potent actin nucleators [141]. This group which includes spire in Drosophila, and spir 1 and 2 in humans, achieve filament nucleation independent of the arp2/3 complex. Quinlan et al. demonstrated that each WH2 repeat in spire recognizes and interacts with both G- and F-actin. They proposed a model showing all four WH2 repeats cooperatively nucleating a strand of the filament with the C-terminal most repeat capping the pointed end.

These are two examples of specialized function by WH2 containing proteins. As new members are being identified in the group, the diversity in their functions becomes even more apparent. The data presented in this thesis have significantly added to our knowledge of this family of proteins. However, to fully understand the role of the WH2 domain in actin cytoskeleton regulation further studies of its individual members are required.

X-ray structures of the spire:actin might help us to further understand the interaction of multiple repeats along the strand of a filament. Such a structure might shed more light on the significance of the C-terminal sequence variations in the switch from sequestering to filament elongation and test our proposed mechanism on filament pointed end capping by some members.

Biochemical and crystallographic studies on VASP is anticipated to yield information on the cooperative interaction of the WH2 domain with other protein motifs and shed more light on the actin bundling property of this protein.

We expect our filamin crystallographic data to result in a high resolution structure of the triple repeats 14-16. These repeats are expected to be structurally similar to the previously published repeats of the molecule. However, we hope to elucidate a novel structure of the hinge 1 and possibly identify the calpain proteolytic site in the region. We hope through this study, to further understand how filamin interacts with its binding partners and through these interactions anchor transmembrane receptors to the cytoskeleton. X-ray structures of this triple repeat in complex with the calcium sensing receptor (CaR), FAP 52, TRAF 1 and 2 (Fig 6) may reveal conformational changes in the filamin rod domain essential for its interactions with these proteins. An understanding of these interactions might act as pointer in the development of a new range of therapeutic drugs targeted at the OPD spectrum disorders.
Conclusions

The present structures allow a better understanding of the mechanism of actin sequestration by the single WH2 domain via both barbed and pointed end capping of the monomers, and have shed light on the possible mode of filament elongation by multiple domains through inherent F-actin binding properties. This model has been partially validated by cross-linking experiments of ciboulot with actin.

High resolution X-ray diffraction data of repeats 14-16 of human filamin A suggests that it will be possible to solve the structure of this triple immunoglobulin repeat.
Acknowledgements

My utmost gratitude goes foremost to my supervisor Ass. Prof. Robert C. Robinson for his mentorship and for patiently trying to make a decent crystallographer out of me. I am beginning to believe that you might yet succeed.

To my second supervisor Prof. Lars Rask, and my examiner Prof. Kristofer Rubin for all the scientific discussions.

It would have been impossible to complete this study without the fantastic collaborations I have enjoyed with all my co-authors, especially Prof. Leslie Burtnick; your sabbatical year with us was an enriching experience for me both scientifically and socially.

My appreciation goes to all past and present members of the structural biology group at IMBIM and BR lab in Singapore for their cooperation, support and friendship these last few years.

A big thank you to all IMBIMites for the great atmosphere of cooperation without which the graduate program would not have been the pleasant journey it has been.

I would like to express my sincere gratitude to Prof. Alwyn Jones for making synchrotron trips a comfortable experience and to Terese Bergfors for all the crystallization tips.

Thank you Amos Sakwe for weathering the storm and giving a listening ear when I have needed to let out some steam; and for the many late night runs to the gas station.

My unreserved appreciation goes to Dr. Adedayo Obafisoye for your continuous and unconditional support and encouragement; you are a friend as close as a brother.

I am eternally grateful to the Abifayo family for giving me a home and a family so far from home.
I acknowledge all my friends in Uppsala and Singapore for all the fun times. A lack of space will not allow me to list everyone, but you know who you are.

Thank you Albert, Maria and Mårten for enduring the horrifying experience of proofreading this thesis in its early chaotic stages.

I wish to express my profound gratitude to the Swedish Medical Science Research Council and Agency for Science Technology and Research (A*STAR), Singapore for financial support.

My love goes to my family; my mum, Dupe, Yemi and SOJ for patiently enduring my prolonged absence from home in my quest for a higher degree. Thank you Mr. Dada for always making your home accessible to me every time I have needed an emergency getaway from the long Swedish winters.

Finally, it is not unto him that willeth or runneth........ I thank the Almighty for health and the many blessings I have enjoyed over the years of this study.

It is finished …… At last
References


tion and depolymerization of actin in a Ca2+-sensitive manner. Plant Cell Physiol 46, 1690-1703.


the Arp2/3 complex on actin polymerization. Biochemistry 38, 15212-15222.


tion of embryonic muscle cells. Biochim Biophys Acta 1175, 243-249.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 182

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)