Structural Studies of Actin and Actin-Binding Proteins

by

Hui Wang

B.Sc., Nanjing University, 2003
M.Sc., The University of British Columbia, 2006

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in
The Faculty of Graduate Studies (Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
July, 2009
© Hui Wang 2009
Abstract

Actin is involved in cell movement, maintaining cell shape and anchoring cytoskeletal proteins. These functions are regulated by many actin-binding proteins, including those of the gelsolin superfamily. Gelsolin superfamily members regulate actin organization by severing, capping F-actin, nucleating the formation of F-actin and/or bundling F-actin. Although abundant structures are available for gelsolin and gelsolin fragments in complexes with actin, the detailed mechanisms for gelsolin activation, and for gelsolin severing and capping of F-actin are still unknown. Structures for gelsolin family members and their complexes with actin help elucidate these mechanisms.

In this thesis, I describe the purification, crystallization and solution of the structures of the following four proteins and protein complexes. The structure of human G1-G3/actin indicates cooperative binding of Ca\(^{2+}\) in G2 and G6 is responsible for opening the G2/G6 latch to expose the F-actin binding site on G2. A new equine G1-G3/actin structure suggests G2-G3 can adopt a CapG-like conformation and reveals novel interactions between gelsolin and actin. The villin V6 structure implies a common spring-loaded activation mechanism in the gelsolin superfamily. Finally, a new actin monomer structure is the first reported for G-actin in an ATP state, without ABPs or modification. All these structures contribute to our understanding of actin’s physiological roles and their regulation by the gelsolin superfamily.
# Table of Contents

Abstract .............................................................................. ii  
Table of Contents .............................................................. iii  
List of Tables ...................................................................... vii  
List of Figures ...................................................................... viii  
List of Abbreviations ............................................................ xi  
Acknowledgements .............................................................. xiii  
Dedication ........................................................................... xv  
Co-Authorship Statement ...................................................... xvi  

1 Introduction ...................................................................... 1  
1.1 Actin ........................................................................... 1  
   1.1.1 Properties and Function ........................................... 1  
   1.1.2 Structure of Actin ................................................... 2  
   1.1.3 F-actin Dynamics ................................................... 9  
   1.1.4 Actin-binding Proteins .......................................... 10  
1.2 Gelsolin ....................................................................... 11  
   1.2.1 Properties and Function ........................................ 11
# Table of Contents

1.2.2 Structure of Gelsolin .............................................. 12  
1.2.3 Gelsolin and Actin Interaction ................................. 16  
1.3 Villin ............................................................... 19  
1.3.1 Properties and Function ......................................... 19  
1.3.2 Structure of Villin ................................................ 21  
1.3.3 Villin and Actin Interaction ..................................... 22  
1.4 CapG ............................................................... 24  
1.4.1 Properties and Function ......................................... 24  
1.4.2 Structure of CapG ................................................ 25  
1.4.3 CapG and Actin Interaction ..................................... 26  
1.5 Heat Shock Protein 27 ................................................ 27  
1.5.1 Properties and Function ......................................... 27  
1.5.2 Structures of sHsps ............................................... 28  
1.5.3 Hsp27 and Actin Interactions .................................... 29  
1.6 X-ray Diffraction .................................................... 30  
1.7 Objectives ........................................................... 32  

2 Materials and Methods .................................................. 34  
2.1 Actin Purification .................................................... 34  
2.2 Gelsolin Purification .................................................. 36  
2.2.1 Equine Plasma Gelsolin ......................................... 36  
2.2.2 Human Cytoplasmic Full Length Gelsolin ..................... 37  
2.3 Villin Purification .................................................... 38  
2.4 Purification of Protein Complexes .................................... 39  
2.4.1 Equine Gelsolin and Actin Complex .............................. 39  
2.4.2 Human Gelsolin and Actin Complex .............................. 40  
2.4.3 Villin and Actin Complex ........................................ 40
### Table of Contents

2.4.4 Hsp27 and Actin Complex .......................... 41

2.5 Protein Crystallography ............................... 42
  2.5.1 Crystallization of Proteins ......................... 42
  2.5.2 Data Collection and Data Processing ............... 44
  2.5.3 Structure Determination and Refinement .......... 45

3 Results and Discussion ................................. 48
  3.1 Structure of a Novel Equine Gelsolin and Actin Complex .. 48
    3.1.1 Comparison with Inactive G1-G3 .................... 49
    3.1.2 Comparison with Previous G1-G3/Actin Structure .. 51
    3.1.3 Comparison with Mutant CapG ....................... 52
    3.1.4 Implication for Gelsolin-Capped F-actin Model .... 53
    3.1.5 Implication for Actin Severing Mechanism .......... 55
  3.2 Structure of Human Gelsolin and Actin Complex .......... 57
    3.2.1 Comparison with Previous G1-G3/Actin Structure .. 58
    3.2.2 Gelsolin Activation by Ca\(^{2+}\) .................. 60
    3.2.3 Implications for FAF ............................... 64
  3.3 Structure of Villin Domain 6 .......................... 65
    3.3.1 Structural Comparison of V6 with G6 ............... 66
    3.3.2 V6 Interaction Interfaces ......................... 68
    3.3.3 Spring-Loaded Mechanism ............................ 73
    3.3.4 Implication for Gelsolin Activation ............... 74
  3.4 Structure of G-Actin ............................... 75
    3.4.1 Comparison with F-actin Protomer .................. 77
    3.4.2 Comparison with ADP-ribosylated Actin ............ 81
    3.4.3 Implication for Actin ATPase Activity ............. 84
    3.4.4 Implication for G-actin to F-actin transition ..... 85
List of Tables

3.1 Data collection and refinement statistics for the novel equine G1-G3/actin complex................................. 50
3.2 Data collection and refinement statistics for the human G1-G3/actin complex........................................... 59
3.3 Data collection and refinement statistics for V6.................. 67
3.4 Data collection and refinement statistics for native G-actin... 78
3.5 RMS deviations of actin Ca positions from alignments of our native actin with other actin structures............. 79

B.1 Data collection and refinement statistics for equine serum albumin....................................................... 111
# List of Figures

1.1 Structures of G-actin in an ADP state and F-actin protomer ........................................ 3
1.2 DNase I/actin structure and the derived G-actin structure ........................................... 4
1.3 G1/actin structure and the derived G-actin structure .................................................. 5
1.4 F-actin Holmes model (original) and Lorenz model .................................................... 6
1.5 F-actin Holmes model (2003) and new high resolution X-ray fiber diffraction model ........ 8
1.6 The regulation of F-actin by actin binding proteins .................................................. 11
1.7 Gelsolin structure in the absence of calcium ................................................................. 14
1.8 Three latches in inactive gelsolin .................................................................................. 15
1.9 Active gelsolin fragments G1-G3 and G4-G6 ................................................................. 16
1.10 Gelsolin G2 binds one Cd\(^{2+}\) ion in a type II Ca\(^{2+}\)-binding site ......................... 17
1.11 Structures of G1-G3/actin and G4-G6/actin ................................................................. 18
1.12 Gelsolin capped actin filament models ......................................................................... 20
1.13 Structures of villin V1 and headpiece .......................................................................... 23
1.14 Structure of mutant CapG ............................................................................................. 26
1.15 Structures of Hsp16.5 and Hsp16.9 ............................................................................. 29

2.1 Illustration of protein crystallization by vapor diffusion ............................................. 43

3.1 Schematic representation of the structure of a novel equine G1-G3/actin complex ............... 49
List of Figures

3.2 Comparison of the novel G1-G3/actin with earlier G1-G3/actin 51
3.3 Alignment of G2-G3 from our novel G1-G3/actin with mutant CapG 53
3.4 Comparison of the novel G1-G3 with mutant CapG 54
3.5 Gelsolin-capped F-actin model based on the novel G1-G3/actin and G4-G6/actin 55
3.6 An alternative mechanism for actin binding, severing and capping by gelsolin 56
3.7 Schematic representation of the structure of human G1-G3/actin complex 58
3.8 Comparison human G1-G3/actin complex with earlier published equine G1-G3/actin 62
3.9 Structural interdependence of the G2 and G6 Ca\(^{2+}\)-binding sites 63
3.10 Schematic representation of the structure of villin V6 in the absence of Ca\(^{2+}\) 66
3.11 Alignment of our Ca-free V6 structure with the Ca-bound G6 structure 68
3.12 Structural comparison of villin V6 (Ca-free) with gelsolin G6 (Ca-bound and Ca-free) 69
3.13 Comparison of the domain 4-interaction surfaces of G6 and V6 71
3.14 Comparison of the domain 2-interaction surfaces of G6 and V6 72
3.15 Domain 6 interactions with actin 73
3.16 Schematic representation of the structure of G-actin in an ATP state 76
3.17 Alignment of our actin structure with several other structures of actin 77
List of Figures

3.18 Structural comparison our G-actin with the F-actin protomer  81
3.19 Residues involved in the G-actin to F-actin transition         82
3.20 Structural comparison our G-actin with ADP-ribosylated actin  83
3.21 Comparisons of nucleotide interactions in our native G-actin
   and F-actin protomer                                         85
B.1 Schematic representation of the structure of equine serum al-
   bumin                                                      110
List of Abbreviations

ABPs — Actin Binding Proteins
ADP — Adenosine 5’-diphosphate
ADF — Actin Depolymerization Factor
ATP — Adenosine 5’-triphosphate
Buffer A — 2 mM Tris-HCl, 0.2 mM CaCl$_2$, 0.2 mM ATP, 1 mM DTT, pH 7.6-7.8
Buffer E — 50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, pH 8.0
Buffer P — 50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA and 1 mM DTT, pH 7.0
Buffer W — 50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, pH 8.0
C1-C3 — CapG domains 1 through 3
CCP4 — Collaborative Computational Project Number 4
DNase I — Deoxyribonuclease I
DTT — Dithiothreitol
EDTA — Ethylene diamine tetracetic acid
EGTA — Ethylene glycol tetraacetic acid
F-actin — Filamentous actin
FAF — Familial Amyloidosis Finnish type
GA — Complex of one gelsolin with one actin molecule
GA$_2$ — Complex of one gelsolin with two actin molecules
List of Abbreviations

GA₃ — Complex of one gelsolin with three actin molecules
G-actin — Actin monomer
G1-G6 — Gelsolin domains 1 through 6
G1-G3 — Gelsolin domains 1 through 3
G2-G4 — Gelsolin domains 2 through 4
G4-G6 — Gelsolin domains 4 through 6
Hepes — N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid
HP — Villin headpiece domain
Hsp27 — Heat shock protein 27
IPTG — Isopropyl β-D-1-thiogalactopyranoside
kDa — kiloDalton
LPA — Lysophosphatidic acid
MS/MS — Tandem mass spectrometry
NaAc — Na acetate
PCR — Polymerase chain reaction
PDB — Protein Data Bank
PEG — Polyethyleneglycol
PIP₂ — Phosphatidylinositol 4,5-bisphosphate
PPIs — Polyphosphoinositides
RMSD — Root-Mean-Square Deviation
SDS-PAGE — Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
Tris-HCl — Tris (hydroxymethyl) aminomethane hydrochloride
TMR — Tetramethyl-rhodamine-5-maleimide
V1-V6 — Villin domains 1 through 6
V4-V6 — Villin domains 4 through 6 with medium length V3-V4 linker
V6 — Villin domain 6
Acknowledgements

First of all, I would like thank my supervisor, Dr. Les Burtnick. Les brought me into the world of research and with whom I began to learn about protein sciences using biophysical methods. His expertise and research insight often serves to give me a sense of direction throughout my Ph.D research projects. Les is also the role model for me of a successful researcher in the field.

Secondly, I wish to thank Dr. Robert Robinson helping us by collecting crystallographic data at a synchrotron and analyzing these data. He also provided me with many helpful suggestions and constructive advice during the course of this work.

Next, I am very grateful to Dr. Dunja Urosev, who helped me set up experiments when I joined in this laboratory. I would also like to extend my gratitude to Anantasak (Tacky) Loonchanta and Chenguang Jiang in our group for their help in my experiments. Tacky shared his knowledge in molecular biology and made many useful suggestions to my experiments. I am also indebted to Drs. Gunnar Olovsson and Wei Chen for their help in using X-Ray Crystallography and Mass Spectrometry facilities.

I also wish to express my appreciation to Drs. Hongbin Li and Grant Mauk. As the member of my supervisory committee, Dr. Li made many valuable suggestions for this thesis. Dr. Mauk generously provided us with Hsp27 protein for Hsp27 and actin interaction research.
Acknowledgements

Sincere thanks are due to Drs. Zhan Zhang, Yi Cao, Ning Wu, Ying Li, Yakun Chen, and Zheng Yang for precious friendship during my stay at UBC. These friends helped me keep tempo of work and made life enjoyable.

I am thankful to UBC Centre for Blood Research for providing excellent research facilities and friendly environment. Funding for the UBC Centre for Blood Research X-ray Crystallography Hub is provided in part by grants from the Canada Foundation for Innovation, the Michael Smith Foundation for Health Research, the Howard Hughes Medical Institute and the Canadian Institutes of Health Research. The CBR Mass Spectrometry Suite is supported by funds from the Michael Smith Foundation for Health Research and Canada Foundation for Innovation.

Finally, I am forever indebted to my parents and my uncles for their understanding, patience and encouragement when it was most required.
Dedication

This thesis is dedicated to my mother, who supported and encouraged me to do my best in all matters of life. It is also dedicated to my wife, whose understanding, encouraging and love have provided a good basis for the thesis.
Co-Authorship Statement

The work presented in this thesis contributes to four collaborative manuscripts that are at different stages of completion. Below I outline my part and those of the other authors.

The structure of villin domain V6 (thesis Chapter 3.3) is in press in the Journal of Biological Chemistry (available online June 1, 2009, in JBC Papers in Press), "Helix straightening as an activation mechanism in the gelsolin superfamily of actin regulatory proteins" by Hui Wang, Sakesit Chumnarnsilpa, Anantasak Loonchanta, Qiang Li, Lynette Kwan, Sylvie Robine, Marten Larssen, Ivana Mihalek, Leslie D. Burtnick, and Robert C. Robinson. Drs. Burtnick and Robinson designed the experiments. I expressed and purified the C-terminal half of villin. I also crystallized villin domain V6 and collected data at the UBC Centre for Blood Research (CBR) X-ray Crystallography Suite. Dr. Robinson and I solved the structure. Dr. Robinson then collected higher resolution data on my crystals at the synchrotron facility in Taiwan. Sakesit Chumnarnsilpa and Anantasak Loonchanta made the plasmid for the C-terminal half of villin. Dr. Ivana Mihalek and her group performed the molecular dynamics simulations reported in the paper. I participated in writing the manuscript with Drs. Burtnick and Robinson.

The structure of human gelsolin domains G1-G3 in a complex with actin (thesis Chapter 3.2) is in press (accepted July 7, 2009) in Proceedings of the
National Academy of Sciences (USA), "Ca^{2+}-binding by domain 2 plays a
critical role in the activation and stabilization of gelsolin" by Shalini Nag*,
Qing Ma*, Hui Wang*, Sakesit Chumnarnsilpa, Wei Lin Lee, Maria Hernandez,
Leslie D. Burtnick, and Robert C. Robinson (*co-first authors). Drs.
Burtnick and Robinson designed the experiments. I expressed and purified
full length human gelsolin. I purified actin from rabbit muscle, then pre-
pared and crystallized the G1-G3/actin complex. I collected diffraction data
at the CBR facility and solved the structure of G1-G3/actin with the help
of Dr. Robinson. Dr. Robinson then collected higher resolution data with
my crystals at the synchrotron facility in Taiwan. Shalini Nag performed
the biochemical assays reported in the manuscript, while Qing Ma crystal-
lized human gelsolin and solved its structure. Sakesit Chumnarnsilpa made
the plasmid of gelsolin. Qing Ma, Shalini Nag and I wrote portions of the
manuscript with Drs. Burtnick and Robinson.

Our manuscript, "The structure of native G-actin" (based on thesis
Chapter 3.4), by Hui Wang, Robert C. Robinson, and Leslie D. Burtnick is
under consideration for publication in the inaugural issue of the re-named
journal, Cytoskeleton (currently Cell Motility and the Cytoskeleton), in Jan-
uary, 2010. Dr. Burtnick and I designed the experiments. I purified actin
from rabbit muscle, and prepared the Hsp27/actin complex that yielded
crystals of actin. I collected diffraction data at the CBR facility and solved
the structure of actin, with help from Dr. Robinson. Dr. Robinson then col-
lected higher resolution data with my crystals at the synchrotron facility in
Taiwan. I wrote the manuscript together with Drs. Burtnick and Robinson.

The novel structure of equine gelsolin domains G1-G3 in a complex with
actin (thesis Chapter 3.1) constitutes the major part of a manuscript in
preparation (80% complete), "Alternative structures of activated gelsolin
Co-Authorship Statement

family proteins”, by Hui Wang, Chenguang Jiang, Robert C. Robinson, and Leslie D. Burtnick. Dr. Burtnick and I designed the experiments. I purified plasma gelsolin from horse serum and actin from rabbit muscle. Next, I prepared and crystallized the G1-G3/actin complex and collected data at the CBR facility. Dr. Robinson and I solved the structure of G1-G3/actin. Then Dr. Robinson collected higher resolution data at the synchrotron facility in Taiwan. Chenguang Jiang purified and crystallized native CapG, a three-domain member of the gelsolin family of proteins. Chenguang Jiang and I wrote the manuscript, together with Drs. Burtnick and Robinson.
Chapter 1

Introduction

1.1 Actin

1.1.1 Properties and Function

As the most abundant protein in a typical eukaryotic cell, actin comes in two forms: actin monomer (G-actin) and actin filament (F-actin). F-actin is formed by polymerization of G-actin, which enables actin to generate movement that is independent of motor proteins. According to their origins, actins are classified as muscle actins (skeletal, cardiac, smooth muscle) or cytoplasmic actins. Actins from different sources and species share a high degree of sequence identity. In muscle cells, as a major component of thin filaments in sarcomeres, muscle actin comprises about 20% of the total protein mass. Muscle contraction generates tension through interaction between F-actin filaments and thick filaments composed of the motor protein myosin.

In the cytoplasm of both nonmuscle cells and muscle cells, actin microfilaments, together with intermediate filaments and microtubules, constitute the three molecular building blocks of the cytoskeleton. Actin microfilaments play roles in maintaining cell shape, protruding membrane forward, anchoring cytoskeletal proteins, and so on [1].

In cytoplasm, actin is in a dynamic equilibrium between G-actin and F-actin. The factors that regulate the transition between G-actin and F-actin
include intracellular ion concentration (ionic strength), pH and a number of actin-binding proteins (ABPs). Higher ionic strength (above 50 mM KCl), neutral or slightly acidic pH, and millimolar concentration of Mg$^{2+}$ contribute to the transition from G-actin to F-actin [2].

1.1.2 Structure of Actin

Structure of G-actin

G-actin is a single polypeptide chain with 375 amino acids residues, with a molar mass of 42 kDa. Because of the propensity of actin to spontaneously self-assembly in solvent conditions generally used to crystallize proteins, actin is either cocrystallized with different ABPs or subjected to chemical modification before crystallization. Known ABP/actin complex structures include DNase I/actin [3], profilin/actin [4, 5], gelsolin/actin [6–8], vitamin D binding protein (DBP)/actin [9], thymosin-$\beta$4/actin [10], ciboulot/actin [11], and so on. By modifying actin at Cys374 (Rhodamine), its nucleotide cleft (Latrunculin-A), or Arg177 (ADP-ribose), the polymerization of G-actin is inhibited. Known structures with modifications include tetramethylrhodamine-5-maleimide (TMR)-modified G-actin in an ADP state (ADP bound in the nucleotide-binding cleft) [12] (Figure 1.1A) and in an ATP analog (AMPPNP) state [13], an antiparallel actin dimer in an ATP state (ATP bound in the nucleotide-binding cleft) with bound Latrunculin-A [14], G-actin in an ATP state after ribosylation, and an antiparallel actin dimer in an ATP state after ribosylation [15].

All these G-actin structures share many common features (Figure 1.1). They have dimensions of 55 Å x 55 Å x 35 Å. Each G-actin comprises four subdomains (numbered 1-4), consisting of a right handed $\beta\alpha\beta$-unit, a muti-
1.1. Actin

Figure 1.1: Structures of G-actin in an ADP state and F-actin protomer. A: TMR-modified actin (PDB code: 1J6Z); B: F-actin protomer (PDB code: 2ZWH). Ribbon diagrams were generated using PYMOL, here and in subsequent figures (http://pymol.sourceforge.net/).

stranded $\beta$-sheet, and a $\beta$-meander. The N-terminus and C-terminus of G-actin are both located in subdomain 1. Subdomains 1, 2 and subdomains 3, 4 are related by a propeller-like twist. These four subdomains come together to surround a bound ATP or ADP. *In vivo*, the nucleotide binding site of G-actin is almost exclusively associated with ATP, while the majority of F-actin protomers contain bound ADP. The exact conformation of G-actin depends on whether there is ATP or ADP in the nucleotide-binding cleft. Along with nucleotide, a divalent metal ion, Ca$^{2+}$ or Mg$^{2+}$, binds within the cleft with high affinity. In contrast to the twisted orientation of the two halves of the G-actin structure, the F-actin protomer has a flat conformation [16] (Figure 1.1B).

Despite the similarities between these G-actins structures, some structural differences exist. Firstly, the DNase I binding loop within subdomain
2 (His40-Gly48) of TMR-modified actin in the ADP state is folded as an α-helix [12], while it is a β-strand in DNase I/actin [3] (Figure 1.2), or disordered in other structures, as in G1/actin [6] (Figure 1.3). These structural differences may be attributed in part to different nucleotides occupying the nucleotide cleft (ADP in TMR-modified actin and ATP in others), or to the binding of DNase I in subdomain 2. Secondly, a closed state and an open state of actin occur with regard to the nucleotide cleft [17]. Most G-actins exhibit the closed state, with only one profilin/actin structure in an open state [5]. The open state may be due to special treatment (soaked in salts) of actin/profilin crystals before data collection. So far, an uncomplexed G-actin structure in the ATP state without modifications has not been reported.

Figure 1.2: DNase I/actin structure (PDB code: 1ATN) (A) and the derived G-actin structure in an ATP state (B).
1.1. Actin

Model of F-actin

As the most functionally relevant form of actin in vivo, F-actin can be described with either a single-start short-pitch left-handed helical representation, or a two-start long-pitch right-handed helical structure. In terms of the latter convention, F-actin can be described as two intertwined strands that cross each other every half pitch (357 Å), with a filament radius of 23.7 Å [16]. Each half pitch includes 13 actin protomers. Each actin protomer is surrounded by four others. In each strand, the axial protomer translation is about 55 Å per actin unit. The two helical strands are axially staggered by half the axial monomer translation [1].

Because of the inherent disorder in linear aggregations of F-actin, the crystallographic structure of F-actin is still unavailable. However, several
plausible models have been proposed based on X-ray fiber diffraction [16, 18–20], cryo-electron-microscopy reconstruction [21, 22], or observed crystallo-
graphic contacts in complexed G-actin crystals [4]. The original Holmes
model [18] was created by fitting the G-actin monomer (PDB code: 1ATN)
into a 8 Å resolution F-actin electron density map determined using fiber
diffraction data from ordered arrays of filaments. In the Holmes model
(Figure 1.4A), subdomains 1 and 2 of each actin monomer are at larger
1.1. Actin

radius with respect to the helical axis than subdomains 3 and 4, which is supported by most experimental evidence. Because the diffraction pattern from the single fibrous molecule is cylindrically averaged by spinning it around the fiber axis, some structural information about each protomer may be lost, e.g. disordered subdomain 2. The Lorenz model [19] (Figure 1.4B) is a refinement of the Holmes model, which obtains a best fit of the calculated diffraction pattern to the experimental diffraction pattern. But the stereochemistry of G-actin is damaged. The latest Holmes model [20] (Figure 1.5A) was constructed by fitting TMR-G-actin structure (PDB code: 1J6Z) to X-ray fibre diffraction patterns. From this model, each actin protomer reveals a flat conformation, with an ordered α-helix in subdomain 2. Cyro-electron-microscopy has been used to yield a model by combining large numbers of cryo-EM images of ostensibly identical individual F-actin assemblies to produce a 3D reconstruction. The latest electron-microscopy model is comprised of three conformation states, with a resolution of 12.5Å [22]. The dynamic multiple conformations of F-actin are attributed to differences in the bound nucleotide, ABPs and actin isoforms studies [22]. In spite of some conformational differences, most models share common features with the Holmes model. But in the Schutt model [4], which is derived from crystallographic properties of profilin/actin crystals, subdomains 3 and 4 are arranged closer to the filament axis than subdomains 1 and 2. In 2009, a high resolution F-actin model (Figure 1.5B) was developed from high resolution X-ray fiber diffraction data to 3.3 Å in the radial direction and 5.6 Å along the equator [16]. By making use of gelsolin and a strong external magnetic field, a well-oriented actin filament sol was produced, which is a prerequisite for obtaining high resolution data. Then the model was built by tilting G-actin in the best orientation within the F-actin helix, calculating
the low-energy vibrational modes of the G-actin, and selecting the combination of modes that best fit the fiber diffraction diagram. From this model, a flat actin protomer conformation, with an open D-loop, was presented for the first time. The transition from G- to F-actin seems to involve a 20° rotation of subdomains 1 and 2 with respect to subdomains 3 and 4 about a rotation axis roughly through the hinge between the domains at the bottom of the nucleotide cleft and at right angles to the helix axis.

Figure 1.5: F-actin Holmes model (2003) and new high resolution X-ray fiber diffraction model [16]. A: Five protomers of F-actin (drawn in green, cyan, yellow, pink and purple) in the Holmes model (2003) [20]; B: Five protomers of F-actin (drawn in green, cyan, yellow, pink and purple) in the new high resolution model [16].
1.1.3 F-actin Dynamics

*In vivo*, F-actin is in a continuous state of assembly and disassembly [1]. Since each actin subunit in F-actin is oriented with its nucleotide cleft toward the same end of filament, F-actin has structural and functional polarity. When decorated with myosin subdomain 1, F-actin shows an arrowhead-like appearance. Thus, one end is called the barbed end (plus end), and the other is the pointed end (minus end). Once polymerization is initiated by nucleation, actin monomers assemble at the barbed end, which grows ten times faster than that at the pointed end. In the steady state, F-actin adds G-actin to its barbed end, while losing G-actin from the pointed end at the same rate. This process is called "treadmilling".

The polymerization and depolymerization of actin can be described in the following steps [1]: firstly, ATP-G-actin (G-actin with bound ATP) incorporates at the barbed end of an actin filament with an accompanying conformational change to the F-actin state; secondly, ATP-F-actin (F-actin composed of subunits having bound ATP) hydrolyzes into ADP-Pi-F-actin (F-actin composed of subunits having bound ADP and inorganic phosphate); subsequently, Pi (phosphate) is released, leaving ADP-F-actin (F-actin composed of subunits having bound ADP); finally, ADP-G-actin (G-actin with bound ADP) is disassembled from ADP-F-actin at the pointed end. At the barbed end, the on-rate for G-actin-ATP is faster than the off-rate, and the on-rate for G-actin-ADP is slower than off-rate. Therefore, the actin monomers at barbed end of F-actin all contain ATP. In contrast, at the pointed end, the on-rate for G-actin-ATP is slower than the off-rate, and the on-rate for G-actin-ADP is slower than the off-rate. Thus actin protoomers at the pointed end contain ADP predominantly.
1.1.4 Actin-binding Proteins

*In vivo*, an abundance of ABPs in cytoplasm regulate actin polymerization, depolymerization, and the organization of actin filaments [1]. If there were not these ABPs in cells, actin would mainly exist as filaments in physiological conditions. ABPs are classified according to how they regulate actin polymerization *in vitro*. Profilin can maintain the population of assembly-ready G-actins by sequestering G-actin. It promotes the exchange of actin-bound ADP for ATP to prime G-actin for polymerization. Thymosin-β4 can also control the availability of G-actin by sequestering G-actin. In contrast to profilin, it inhibits the exchange of ADP for ATP and, so, blocks actin polymerization. The Arp2/3 complex can nucleate actin assembly, promote branching of F-actin, and cross-link F-actin by binding to the sides of filaments. CapZ is the most abundant barbed-end capping protein. It can nucleate actin assembly, as well as capture preexisting F-actin and regulate actin assembly at the barbed end. ADF/cofilin can promote depolymerization of F-actin by binding to ADP-F-actin and ADP-G-actin; it can also cap the pointed end of tropomyosin-coated actin filaments and stabilize short-actin oligomers. DNase I can promote actin depolymerization by binding tightly to the pointed end of G-actin from the G-actin pool available for assembly. Gelsolin can bind to the sides of ADP-F-actin, sever F-actin, and block the barbed end of F-actin. All these ABPs together regulate actin-driven movement in the absence of motor proteins (Figure 1.6).
1.2 Gelsolin

1.2.1 Properties and Function

The gelsolin superfamily of ABPs can regulate actin organization by severing and capping actin filaments, and nucleating the formation of new filaments. The activities of the family members are regulated to varying degrees by calcium ions, pH, phosphoinositides and tyrosine phosphorylation [23]. This superfamily consists of at least ten members, including gelsolin, villin, CapG, adseverin, advillin, supervillin and flightless-I [24]. These proteins generally contain three or six-fold repeats of a gelsolin-like domain, with or without additional domains at the N- and/or C-termini.
1.2. Gelsolin

As the founding member of this family, gelsolin is found in a wide range of vertebrate, lower eukaryotic and plant cells. Gelsolin exists in two different isoforms: cytoplasmic gelsolin (intracellular) and plasma gelsolin (extracellular)[24]. Cytoplasmic gelsolin regulates the mobility and architecture of cells. Its activities include severing F-actin, capping F-actin ends and nucleating actin assembly. Compared with cytoplasmic gelsolin, plasma gelsolin has a 25-amino acid peptide extension at its N-terminus and a disulfide bond between Cys188 and Cys201. Plasma gelsolin can sever and cap actin filaments which are released into circulation after cell death. Vitamin D-binding protein (DBP) then can sequester actin monomers of the pointed end of these capped actin oligomers. This pair of proteins constitute the extracellular actin scavenging system that removes potentially harmful actin filaments from the circulatory system.

1.2.2 Structure of Gelsolin

Gelsolin with a molar mass of 81 kDa (cytoplasmic gelsolin) or 83 kDa (plasma gelsolin), contains two homologous halves: N-terminal (G1-G3) and C-terminal (G4-G6). Both halves are comprised of three homologous domains, each of which consists of 120-130 amino acid residues. G1 and G4, G2 and G5, and G3 and G6 are most closely related when compared pairwise based on amino acid sequence analysis and structure. In the structure of equine plasma gelsolin in the absence of Ca$^{2+}$ [25], the six domains pack into a compact globule with overall dimensions 85 Å x 55 Å x 36 Å (Figure 1.7). Each of G1 through G6 share a similar folding topology, consisting of a five or six stranded β-sheet at its core sandwiched between a 3.5-4.5 turn α-helix and a 1-2 turn α-helix, which run parallel to and perpendicular to the
1.2. *Gelsolin*

... strands in the sheet, respectively. The N-terminal and C-terminal halves are connected by a 53-residue linker, which enforces limits on the independence of the two halves, once they have been activated and move apart. Three latches hold inactive gelsolin in an autoinhibited conformation: tail latch (G2/G6 latch), G1/G3 latch and G4/G6 latch (Figure 1.8). The tail latch, between a helical extension at the C-terminus and G2, holds the two halves of gelsolin in an autoinhibited state, in which the F-actin binding site on G2 is blocked. The G1-G3 and G4-G6 latches, are formed by shared β-sheets between G1 and G3, and between G4 and G6, respectively. The binding sites for G-actin in G1 and G4 are masked within the globular arrangement of the domains in inactive gelsolin.

The activation of gelsolin is regulated by Ca$^{2+}$. Sequence and structural data indicate that gelsolin possesses eight Ca-binding sites, classified as two types: type I and type II [26]. Type I sites involve coordination of Ca$^{2+}$ between actin and a gelsolin domain. Both G1 and G4 possess one type I site. Type II Ca$^{2+}$ sites are wholly contained within gelsolin domains (Figure 1.9). Each of the six domains possesses one type II site. Occupation of each of these sites by a metal ion has been confirmed by crystallographic data. An exceptional case is that of the type II binding site of G2 in the structure of equine G1-G3 bound to actin [7], in which the site is vacant. But this site can be filled by Cd$^{2+}$ [27] (Figure 1.10) or Tb$^{3+}$ [28] ions. The vacant G2 type II site in this G1-G3/actin structure is proposed to facilitate the activation process through a mechanism that involved the transient binding of Ca$^{2+}$ [7].

At micromolar Ca$^{2+}$ levels, inactive gelsolin becomes activated via global conformation changes [29]. However, the structure of fully activated intact gelsolin is still not available. Small-angle X-ray scattering experiments indi-
1.2. Gelsolin

Figure 1.7: Gelsolin structure (PDB code:1D0N) in the absence of calcium. The six domains of are colored as: G1 (red); G2 (light green); G3 (yellow); G4 (pink); G5 (dark green); G6 (orange). The color scheme is used in the subsequent figures. This form of gelsolin does not bind actin.

cate activated gelsolin to have undergone large-scale conformation changes and highlight the critical roles played by the flexible linkers between domains [30]. So far, structures of activated fragments of gelsolin have been investigated in isolation [27, 31, 32] and in the form of complexes with G-actin [6–8, 33]. Compared with inactive gelsolin, these activated gelsolin fragments reveal dramatic conformational rearrangements (Figure 1.9). In activated G4-G6, the continuous $\beta$-sheet between G4 and G6 is torn apart; the kinked long helix in G6 is straightened; G6 is flipped over and translated by 40 Å relative to G4 and G5; new G6 contacts with G5 are established. As for activated G1-G3, the $\beta$-sheet that runs from G1 to G3 is separated
1.2. Gelsolin

Figure 1.8: Three latches in inactive gelsolin (PDB code: 1D0N). A: G2/G6 latch (tail latch); B: G1/G3 latch; C: G4-G6 latch

into two; G1 translates away from G2 to extend the reach of the G1-G2 linker to 30 Å. The G2-G3 linker shortens through adoption of a helical conformation. But the structure of the G3-G4 linker in activated gelsolin is still unknown.

G4-G6 is a calcium-dependent actin monomer binding fragment, while G1-G3 does not require calcium for activity [34]. Fluorescence and equilibrium dialysis experiments [35] showed that two high-affinity Ca$^{2+}$-binding sites are located in G4-G6, with Kd values of 0.2 µM and 2 µM. Since isolated G1-G3 can conduct actin severing in a calcium-independent manner, the occupation of these two sites was considered to be the minimum requirement to activate gelsolin. From synchrotron footprinting experiments [36], a third calcium binding site in G4-G6 with a Kd value of 100 µM has also been identified. Based on the available information [28], the two high-affinity sites were assigned to G4 (2 µM) and G6 (0.2 µM), while the low affinity one was assigned to G5 (100 µM). But it is still not clear how many calcium ions are necessary to activate Ca-free gelsolin.
1.2. Gelsolin

Figure 1.9: Active gelsolin fragments G1-G3 (A) from G1-G3/actin structure (PDB code: 1RGI) and G4-G6 (B) from G4-G6/actin structure (PDB code: 1H1V). Black spheres represent calcium ions bound at type II sites. The type II calcium ions in G1, G3, G4, G5, G6 are coordinated by an Asp, two backbone carbonyl oxygen atoms and a Glu.

1.2.3 Gelsolin and Actin Interaction

In the presence of calcium, gelsolin binds to the side of F-actin, severs F-actin, and then caps the newly generated barbed filament end. But because of the inherently dynamic nature of F-actin and the multiplicity of steps in the process [29], the structure of a gelsolin-capped filament would be hard to determine directly. A gelsolin-capped filament model has to be constructed by combining structural information from gelsolin fragment/actin complexes and the available actin model. Proteolytic fragments and sequence analysis
identified three actin binding regions in gelsolin: a calcium-independent G-actin binding sites in G1 (residues 49-72, 121-135); a calcium-dependent G-actin binding site in G4 (residues 431-454) and a calcium-independent F-actin binding site in G2 (residues 162-166) [29]. In inactive gelsolin, these actin binding sites are blocked by the G1-G3 latch, the G4-G6 latch and the tail latch, respectively. Although the structure of intact gelsolin bound to actin is still unavailable, crystal structures of the C-terminal half of human gelsolin bound to one actin and the N-terminal half of equine gelsolin bound to one actin have been solved [7, 8] (Figure 1.11). From the G4-G6/actin structure, G4 binds to a cleft between the subdomains 1 and 3 of actin, a location involved in binding several other ABPs. In addition, G6, displays a secondary actin-binding site for subdomain 3 of actin. In G1-G3/actin, G1 binds to the same location on an actin monomer as does G4. In this complex, G2 binds to subdomain 2 of actin, unlike its homologue in the G4-
Gelsolin

Gelsolin consists of six globular domains, G1-G6, which interact with actin. G5, which does not contact actin. Additionally, G3 exhibits a secondary actin-binding site, making contact with subdomain 1 of actin.

Figure 1.11: Structures of G1-G3/actin and G4-G6/actin. A: Structure of G1-G3 bound to actin (G1, red; G2, green; G3, yellow; actin: cyan) in the presence of calcium (PDB code: 1RGI) B: G4-G6 bound to actin (G4, pink; G5, dark green; G6, orange; actin: cyan) in the presence of calcium. (PDB code: 1H1V)

The interaction between gelsolin and actin is inhibited by phosphatidylinositol 4,5-bisphosphate (PIP$_2$)[23]. The PIP$_2$-binding sites (residues: 132-149 and 161-172) overlap the actin binding sites of G2 so that PIP$_2$ dislodges and uncaps actin [25]. *In vivo*, the binding of PIP$_2$ by gelsolin leads to its sequestration in an inactive state near the cell membrane. Gelsolin is released into cytoplasm after PIP$_2$ is hydrolyzed.

By superimposition of the G4-G6/actin structure and G1-G3/actin structure onto two protomers within the Holmes F-actin model, a model for a gelsolin-capped filament has been proposed [7]. This model explains how
1.3. Villin

activated intact gelsolin binds to the side of F-actin, severs F-actin, and then caps the newly generated barbed filament end. In this model, G2-G3 is placed at the junction between two longitudinally neighboring actin protomers. The long helix of G2 lies close to the binding site on the actin occupied by the analogous helix of G1 or G4, but on a different protomer. Electron microscopy studies of G2-G6/F-actin [37] also confirmed the F-actin-binding region in G2 to be attached to two longitudinally neighboring actin protomers. Because the length of the G3-G4 linker (53 amino acid residues) permits G4-G6 to interact with two different pairs of protomers across the actin filament, two variations of the model exist [38] (Figure 1.12). Based on the model, structures of gelsolin bound to one actin (GA), two actins (GA$_2$) and three actins (GA$_3$) have been proposed [38]. These complexes have been prepared and crystallized, but preliminary crystallographic analysis shows considerable disorder among some protein domains. Further investigations of these complexes are needed to test the gelsolin-capped models of F-actin and to determine the G3-G4 linker conformation.

1.3 Villin

1.3.1 Properties and Function

Epithelial cells serve as a physiological and structural barrier and are involved in vectorial transport of ions, solutes and water. Actin filaments in these cells are mainly localized in the apical cytoskeleton, where microtubules and intermediate filaments are absent. These crosslinked filaments form microvilli at the apical surface, which can increase the cell surface area and regulate its absorptive and secretory functions [39]. As member of gelsolin superfamily, villin is located in the microvilli of absorptive epithelium. It
1.3. Villin

Figure 1.12: Gelsolin capped actin filament models. A: Five protomers of F-actin (drawn in cyan) with G1-G3/actin (G1, red; G2, green; G3, yellow) and G4-G6/actin (G4, pink; G5, dark green; G6, orange). B: Four protomers of F-actin (drawn in cyan) with G1-G3/actin and G4-G6:actin, [7]

is able to regulate epithelial cell structure and function. Villin is unique in that it retains all the actin-regulatory functions of gelsolin, but it can also bundle (crosslink) actin filaments in the presence of nanomolar concentrations of Ca$^{2+}$. Although gelsolin can perform capping and severing functions at micromolar concentrations of Ca$^{2+}$, micromolar and millimolar concentrations are required for villin to exhibit capping and severing functions, respectively. However, after tyrosine phosphorylation, villin can sever actin filaments even in the nanomolar range of Ca$^{2+}$ [40]. Villin shares similar proteolytic cleavage patterns with gelsolin [41]. Like the C-terminal half of gelsolin, the binding of G-actin by the C-terminal half of villin requires
1.3. Villin

the presence of Ca\(^{2+}\). In contrast to the N-terminal half of gelsolin, actin-severing by the N-terminal half of villin is calcium-dependent. Creation of hybrid proteins demonstrates that the domains of villin and gelsolin are not interchangeable [42].

1.3.2 Structure of Villin

Villin contains two distinct parts, a gelsolin core comprised of six gelsolin-like domains (V1-V6) and a C-terminal headpiece (HP). HP is a seventh domain that is connected to the villin core by a 40-residue linker. Because of the high sequence identity (50%) between villin and gelsolin, much of what is inferred about villin structure and function is based on data from studies of gelsolin. Like gelsolin, villin exists in an inactive conformation in the absence of calcium. After association with Ca\(^{2+}\), villin undergoes a major conformational rearrangement via a so-called hinge mechanism [43]. As a result, the villin conformation becomes more asymmetric, with an overall increase in maximal dimension from 84Å to 123Å, based on hydrodynamic and spectroscopic studies [43]. Also, F-actin binding sites in the villin core are exposed, thus activating its actin-severing function. No high resolution structure of intact villin is available.

From sequential mutagenesis studies [44], six calcium binding sites have been identified in villin: two major calcium binding sites (one is type I and the other is type II) in V1, and four type II calcium sites in V2-V6. The type I calcium site in V1 (Glu25, Asp44 and Glu74) regulates the F-actin capping and severing activities, while the type II calcium site in V1 (Asp86, Ala93 and Asp61) only regulates the severing activity. The other four type II sites are involved in stabilizing villin conformation and regulating actin
severing activity. The calcium binding sites regulating F-actin severing have much lower affinity than those regulating capping activity [45].

Although abundant X-ray crystallographic structural information exists for gelsolin, structural data for villin is limited to the V1 [46] and HP domains [47, 48] (Figure 1.13). Like typical gelsolin domains, the V1 domain consists of a five-stranded $\beta$-sheet at its core, sandwiched between a pair of $\alpha$-helices, a long helix that runs roughly parallel and a short helix that lies perpendicular to the strands. HP is a very compact structure that folds around an extensive hydrophobic core. It can be described best as two subdomains with unique folding properties and differing contributions to F-actin-binding. The N-terminal subdomain is primarily made up of bends and reverse turns around a hydrophobic core defined by Phe16, Leu18, Leu21, Val22, Leu29, Pro30, Val33, Lys38, Glu39, His41, and Leu42. The C-terminal subdomain is composed of three $\alpha$-helices that pack tightly around a C-terminal hydrophobic core defined by Asp46, Phe47, Val50, Phe51, Met53, Phe58, Leu61, Leu69, Lys70, and Leu75 [47].

### 1.3.3 Villin and Actin Interaction

Villin binds to and caps barbed-ends of F-actin with high affinity in the presence of micromolar Ca$^{2+}$. The two F-actin binding sites of villin are located in the V1-V2 linker and headpiece [39]. Tyrosine phosphorylation can decrease the affinity of villin for F-actin [49]. In the presence of tropomyosin, the binding of villin to F-actin can be inhibited [50]. At millimolar concentrations of Ca$^{2+}$, villin possesses severing activity [51]. The F-actin binding site in the V1-V2 linker, which overlaps with PIP$_2$ binding sites, is involved in severing. Some basic residues have been associated with the binding of
1.3. Villin

Figure 1.13: A, Structures of villin V1 (PDB code: 2VIL); B, Structure of villin headpiece (PDB code: 1QQV). The N-terminal and C-terminal subdomains of HP are colored red and green, respectively.

F-actin, including Arg138, Lys143 and Lys145 [52].

Like gelsolin, villin can bind to G-actin and nucleate actin pointed-end growth of F-actin. By analogy with gelsolin G1 and G4, G-actin binding sites on villin are likely to exist in V1 and V4, although there is no direct evidence [39]. The nucleation activity of villin is regulated by both G-actin bindings sites in V1 and V4, while the capping activity is regulated by the V1 G-actin binding site alone.

Villin is special in that it has a bundling function which gelsolin does not. Previous studies showed F-actin binding sites in both the core and HP to be necessary to bundle F-actin, but more recent studies indicated that
1.4. CapG

villin can use only the HP F-actin binding sites to perform bundling [53]. F-actin binding sites in HP include a hydrophobic cap and a crown of mixed charged residues, which consist of Arg787, Lys815, Lys816 and Lys789 [47]. Since HP is connected to V6 by a 40-residue disordered linker, HP has been proposed to bind actin independently of the remainder of the protein [54]. Recently, a villin-crosslinked F-actin model has been proposed, based on the electron tomography experiments [55]. In it, villin bundles F-actin by using three domains simultaneously: V1, V2 and HP, even without Ca$^{2+}$. In this model, the interactions between villin and F-actin protomers are totally different from those of other actin-binding proteins (even gelsolin) with actin.

Interaction between villin and actin is regulated by calcium, phosphorylation and phosphoinositides (e.g., PIP$_2$) [56]. Higher concentrations of Ca$^{2+}$ and lower concentrations of PIP$_2$ contribute to the actin-capping activity. Higher concentrations of calcium and tyrosine phosphorylation help, while PIP$_2$ inhibits, actin-servering activity [45]. Tyrosine phosphorylation inhibits actin-nucleating activity. PIP$_2$ favors, but tyrosine phosphorylation disfavors, actin-bundling activity.

1.4 CapG

1.4.1 Properties and Function

CapG (also named macrophage-capping protein, gCap39, or Mbh1), which is located mainly in nuclei, is another member of the gelsolin/villin superfamily. Although CapG shares 49% sequence identity with the N-terminal half of gelsolin, CapG only caps actin filaments, but does not sever them. Similar to gelsolin, CapG requires Ca$^{2+}$ for activation and actin-binding. In
contrast to gelsolin, the capping by CapG is reversible by reducing the Ca^{2+} concentration to submicromolar [57]. Studies of CapG knock-out cells suggest that CapG is involved in ruffling, which is attributed to the reversible capping property of CapG [58]. Also, CapG plays an important role as a mediator of endothelial cell response to mechanical forces [24].

1.4.2 Structure of CapG

CapG contains three gelsolin-like domains: C1, C2 and C3. The structure of intact CapG is still unavailable. However, the structure for a CapG having mutations in its severing domain (residues 84-91) and WH2-containing segment (residues 124-137) is available [59] (Figure 1.14). The WH2 (Wiscott-Aldridge syndrome protein (WASP) homology domain 2) domain has the consensus sequence FKHVXPN and appears in the G1-G2 linker. It is shared by many actin regulatory proteins, such as gelsolin, WASP, thymosin-β4, etc [33]. The specific function of the WH2 domain is still under investigation. The crystal structure of this mutant CapG in the presence of Eu^{3+} contains three homologous domains, in which C1 is translated away from C2 and C3 by an extended C1-C2 linker (36 Å). The orientations of C1, C2 and C3 are strikingly different from those of G1, G2, and G3 derived from G1-G3/actin structure [7]. Each domain consists of a five-stranded β-sheet at its core, sandwiched between a pair of α-helices, a long helix that runs roughly parallel and a short helix that lies perpendicular to the strands. This structure contains two type II Eu^{3+} in conserved metal-binding sites located in domains C1 (involving Gly45 and Asp46) and C3 (involving Asp279, Asp280 and Glu304).

Based on the crystal contacts in this mutant CapG structure, a model
1.4. CapG

Figure 1.14: Structure of mutant CapG (PDB code: 1JHW). The three domains of mutant CapG are colored as red (C1), green (C2) and yellow (C3), respectively. A cyan sphere represents Eu^{3+}.

for an inactive mutant CapG has also been derived [59]. In the inactive mutant CapG model, the interface between C1 and C2 is similar to that in G1 and G2 in the Ca-free gelsolin structure. But the interface between C2 and C3 is different from that between G2 and G3 in that same structure [25].

1.4.3 CapG and Actin Interaction

CapG binds and caps F-actin in the presence of micromolar Ca^{2+}. The binding affinity of CapG for actin is much lower than that of gelsolin. In spite of native CapG not having a severing function, such an activity can be introduced into CapG by mutation of the C1-C3 linker to match the
corresponding sequence in gelsolin. The structure of CapG bound to actin is still unavailable.

Based on the mutant-CapG structure, a model of a mutant CapG capped actin filament has been proposed [59]. This model was constructed by first superimposing C1 onto a barbed end actin protomer according to the G1/actin structure [6], then rotating domains C2 and C3 about the C1-C2 linker in order to dock C2 to the next actin protomer in the same mode as in G1/actin. But this strategy ignores the inferred F-actin binding properties of C2, which should bind to actin in advance of C1. Clearly, further evidence is needed to test the validity of this model.

1.5 Heat Shock Protein 27

1.5.1 Properties and Function

Heat shock proteins (Hsps) are synthesized by cells or organisms in response to environmental and physiological stress, and act as chaperones to restabilize partially unfolded proteins. Hsps have been identified in almost all organisms from prokaryotes to mammals. Based on their size, structure and function, Hsps are divided into five major families, Hsp110, Hsp90, Hsp70, Hsp60 and small Hsps (sHsps), having molar masses of 15-43 kDa. When exposed to stress (such as heat shock), cells respond by modifying their cytoskeletal networks. As a result, microtubules, intermediate filaments and actin filaments are disorganized. Hsps are induced to overexpress locally and interact with different cytoskeletal components. Large Hsps bind mostly to the microtubules, while sHsps contribute to modulation and stabilization of actin filaments and intermediate filaments [60].

sHsps comprise an abundant and diverse family of proteins, including
1.5. Heat Shock Protein 27

Hsp27, Hsp22, Hsp20, αA-crystallins, αB-crystallins, and so on. All except αA-crystallins are highly expressed in muscles and heart, where actin expression is also high. Thus the main functions of sHsps in muscle seem to be related to interactions with actin [61]. Similar to other Hsps, Hsp27 functions as a molecular chaperone to help refold non-native proteins. Hsp27 can prevent actin filaments disaggregation during exposure to various stress stimuli. Also, Hsp27 inhibits apoptosis through an ability to interact with key components of the apoptotic signalling pathway [62].

1.5.2 Structures of sHsps

In structure, sHsps share a conserved αA-crystallin domain near the C-terminal region, which consists of about 90 amino acids and has an IgG-like fold, followed by a flexible C-terminal tail. In contrast, the N-terminal regions of sHsps differ in sequence and length and contain a WDPF motif near the N-terminus, which is involved in their oligomerization [63]. As is true for most other sHsps, Hsp27 forms large oligomers in vivo, the size of which depends on temperature, pH, ionic strength and the degree of phosphorylation of individual monomers. In a non-phosphorylated state, Hsp27 usually aggregates into sets of six tetramers, i.e. 24-mers [64]. After phosphorylation at Ser15, Ser78 and Ser82, Hsp27 24-mers dissociate into tetramers [65]. The quaternary structure of Hsp27 determines its functional properties. There is no Hsp27 monomer or oligomer structure available. So far, the crystal structures of sHsps are limited to two non-mammalian sHsps: Hsp16.5 [66] and Hsp16.9 [67] (Figure 1.15). Despite their variation in amino acid sequence, both structures conserve a compact β-sheet sandwich similar to the immunoglobulin-like fold. The sheet consists of two layers of three
and five antiparallel strands, respectively, connected by a short interdomain loop.

![Figure 1.15: Structures of Hsp16.5 (PDB code: 1SHS) (A) and Hsp16.9 (PDB code: 1GME) (B).]

### 1.5.3 Hsp27 and Actin Interactions

The detailed interactions between sHsps and actin remain to be clarified [68]. Interaction between Hsp27 and actin is regulated by phosphorylation and dephosphorylation. Previous studies showed non-phosphorylated Hsp27 monomers inhibit actin filament polymerization, while phosphorylated Hsp27 oligomers stabilize the actin filament [69]. Phosphorylated Hsp27 monomers and non-phosphorylated Hsp27 oligomers are not active [70]. Hsp25 monomer, a homologue of human Hsp27, binds to actin at a 1:1 molar ratio. It was also proposed that Hsp27 affects actin polymerization
by interacting with G-actin [61].

*In vitro*, non-phosphorylated Hsp27 monomers cap the barbed-ends of actin filaments and thus inhibit F-actin polymerization. By comparison of structural motifs shared by Hsp27 and actin, the PEGTLTVEAP motif in Hsp27, containing residues similar to those involved in actin-actin contacts in F-actin, is proposed to mimic an actin monomer and bind to actin at the corresponding actin-actin interface located in subdomain 3 [68]. *In vivo*, phosphorylated Hsp27 oligomer regulates F-actin dynamics and protects F-actin. A model for the protection of F-actin by sHsps was proposed [68]. In non-stressed cells, sHsps form large aggregates of non-phosphorylated monomers. In stressed cells, sHsp become phosphorylated and dissociates into small oligomers. These phosphorylated small oligomers (rod-like tetramers) would coat F-actin and protect it from breakage by actin-servering proteins.

More recently, some researchers have proposed that sHsps interact with F-actin only under stress conditions [61]. Furthermore, they found Hsp27 has no affect on the thermal unfolding of F-actin, but prevents aggregation of denatured actin. Hsp27 with three mutations, S15D/S78D/S82D, mimics the properties of phosphorylated Hsp27 *in vivo* and can form soluble complexes with denatured actin.

## 1.6 X-ray Diffraction

X-rays can interact with matter through their fluctuating electric field, which causes electrons to oscillate in position and emit X-rays with the same wavelength as the incident radiation. Because the intensity of scattered radiation is proportional to the square of the charge/mass ratio, electrons are much more efficient in this process than atomic nuclei or protons. The
3-dimensional structures of proteins allow us to understand biological processes at the most basic level. Although several methods, including NMR, electron microscopy, and X-ray fiber diffraction, are employed to study protein structures presently, X-ray crystallography is still the most commonly used technique to analyze macromolecular structure at atomic resolution. Since the wavelengths of X-ray are comparable to the scale of covalent chemical bonds and to the radius of a single atom, scattered X-rays from individual atoms contain information about the structural arrangement of the atoms in the sample under study. But X-ray scattering from a single molecule would be too weak to detect above the noise level. Crystals contain large arrays of molecules in the same orientation and diffraction from crystals can yield measurable signals. When X-rays hit atoms in protein crystals, the scattered waves interfere with one another, depending on the direction of the incoming and outgoing waves, and the positions of the electrons relative to each other. The relationship can be summarized as Bragg’s law:

\[ n\lambda = 2d \sin \theta \]

where \( d \) is the spacing between diffracting planes, \( \theta \) is the incident angle, \( n \) is any positive integer, and \( \lambda \) is the wavelength of the incident X-rays. If Bragg’s law conditions are satisfied, the scattered X-rays from individual atoms in a periodic structure combine constructively to generate a diffraction pattern.

Traditional X-ray sources are copper targets bombarded with high energy electrons, which emit CuK\( \alpha \) X-rays at a wavelength of 1.5418 Å. Synchrotron radiation is electromagnetic radiation, generated by the acceleration of charged particles (electrons or positrons) to relativistic velocities through magnetic fields. Photons are emitted in a narrow cone in the for-
ward direction, at a tangent to the orbit, when the circulating charged particles are perturbed by magnetic fields. The radiation produced may be tuned over the entire electromagnetic spectrum.

1.7 Objectives

At the start of this project, I set out to obtain X-ray crystallographic structures for intact gelsolin/actin complexes, including GA, GA$_2$ and GA$_3$. Solving these structures would support or refute existing models for gelsolin-capped F-actin. In addition, I would be able to extract the as yet unknown structure of intact activated gelsolin from the intact gelsolin/actin complex structures. In particular, I expected to learn whether the 53-residue long G3-G4 linking polypeptide plays an active role in the functions of gelsolin, or whether it simply acts as a passive tether that limits the span between the two functional halves of the intact protein. As the project developed, it became clear that my crystals would only yield high-resolution structures for fragments of the complexes. Fortunately, at about this stage of my work, constructs were made available (by our collaborators from the Robinson laboratory in Singapore) that would allow us to clone and express numerous proteins or fragments of proteins of the gelsolin/villin superfamily of actin regulators. I prepared recombinant human gelsolin to use in my crystallization trials instead of our standard natural source equine gelsolin, in the hope that the small differences in amino acid sequence would lead to crystals that would help achieve my original goals. I also cloned and expressed a fragment of gelsolin corresponding to G2-G4, which includes the G3-G4 linker, in order to learn more about the linker and its functions. In addition, I expressed and purified the C-terminal half of villin, a protein that
has approximately 50% amino acid sequence identity with gelsolin. Again, I screened conditions and grew crystals from samples containing this villin fragment in the presence and absence of actin. Because of the close homology between villin and gelsolin, I expected that structural information from villin crystals could be used to infer how gelsolin carries out its functions. Finally, in an attempt to learn more about gelsolin’s target, actin, I used a small heat shock protein, Hsp27 (an actin-binding protein of interest to Dr. Grant Mauk of the UBC Centre for Blood Research, who kindly supplied us with recombinant Hsp27), to form a complex with actin and grow crystals suitable for crystallographic analysis.
Chapter 2

Materials and Methods

2.1 Actin Purification

Preparation of muscle acetone powder

Rabbit muscle acetone powder was prepared based on a modified method by Pardee and Spudich [71]. 1 kg of rabbit back and leg muscles was minced in a blender with 3 L of cold 0.3 M KCl, 0.1 M KH$_2$PO$_4$, 0.05 M K$_2$HPO$_4$, 0.2 mM ATP, pH 6.5 to dissociate myosin. The slurry was centrifuged at 7000 rpm using Beckman polypropylene bottles and an Avanti J-20 XP (Beckman) centrifuge for one hour. The sediment, containing actin and thin filament proteins, was washed with 3 to 4 litres of cold water for 20 min, centrifuged for 30 min and washed with 3 L of cold 0.01 M NaHCO$_3$, pH 8.0, for 20 min in order to remove hemoglobin and myoglobin. After centrifugation for another 10 min, the sediment was washed with 3 L of cold water for 10 to 20 min and then centrifuged for 10 min. The sediment was twice washed with 4 litres of cold 95% ethanol and centrifuged for 10 min. The residue was dehydrated by treatment with 4 L acetone and then filtered by squeezing through a double layer of cheesecloth. This step removes lipids and denatures most non-actin proteins. Finally, the residue was spread out on filter paper to dry at room temperature (in the fumehood) for a few hours. The dried muscle powder was kept in sealed containers and stored

34
2.1. Actin Purification

at -20°C.

Preparation of actin from acetone powder

Actin was purified from rabbit skeletal muscle based on a modified method of Spudich and Watt [72]. 5 g rabbit muscle acetone powder was extracted with 50 ml of 2 mM Tris-HCl, 0.2 mM CaCl$_2$, 0.2 mM ATP, 1 mM DTT, pH 7.6-7.8 (buffer A) for 45 min on ice. Then the extract was filtered through two layers of cheesecloth and filter paper. The residue was then washed with 20 ml of buffer A and filtered again. The combined filtrates were centrifuged at 35,000 rpm (Beckman model optima L-90k ultracentrifuge and Beckman 45 Ti rotor) for one hour at 4°C. After discarding the pellet, the actin in the supernatant was polymerized overnight by the addition of KCl and MgCl$_2$ to final concentrations of 100 mM and 2 mM, respectively. The next day, tropomyosin were dissociated from actin and solubilized by addition of KCl to 0.8 M. After stirring 3 hrs at 4 °C, the actin solution was centrifuged at 35,000 rpm for 3 hours. The pellet, containing F-actin, was suspended in 10 ml buffer A using a potter homogenizer, after which the supernatant was dialyzed against three changes of 1 L of buffer A over three days. The resulting G-actin solution was centrifuged at 35,000 rpm (Beckman model optima L-90k ultracentrifuge and Beckman 70.1 Ti rotor) for 3 hours to get rid of any remaining F-actin. Finally, the G-actin solution was applied to a gel filtration column (Bio-Rad Sephacryl S300; 90 x 2.5 cm) at room temperature and eluted with buffer A at a rate of 2 ml/min. Elution was monitored by an ECONO UV absorption monitor (Bio Rad) at 280 nm. Actin was eluted as a single peak, but only the second half of the peak was collected for subsequent use, in order to avoid possible contamination.
2.2. Gelsolin Purification

2.2.1 Equine Plasma Gelsolin

Equine gelsolin was purified from horse serum (Pel-Freez Biologicals, Rogers, AR) based on a modified method reported previously [38]. Protease inhibitors leupeptin (2 mg/ml in H_2O) and pepstatin (2 mg/ml in DMSO) were added to the thawed horse serum, with the final concentration of each inhibitor being 200 ng/ml. Then the serum was dialyzed against three changes of 4 L of 25 mM Tris-HCl, 0.5 mM CaCl_2, pH 7.5 at 4 °C for three days. The dialyzed horse serum was centrifuged at 9,000 rpm (RC-5B centrifuge and Sorvall GSA rotor) for 40 mins. Solid NaCl was added to the clarified plasma to achieve a concentration of 50 mM. This solution was then mixed with 1.5 L of anion exchange resin (DEAE-Sephadex A-50, Pharmacia) in 25 mM Tris-HCl, 0.5 mM CaCl_2, 50 mM NaCl, pH 7.5. The mixture was kept at 4 °C and stirred every 30 min for 2.5 hours. In the presence of Ca^{2+}, gelsolin and positively charged proteins are left in the liquid phase because they do not bind the anion exchange resin. At the same time, serum albumin and other negatively charged proteins bind the anion exchange resin. Next, the mixture was filtered, and EDTA was added to the filtrate to a concentration of 10 mM to remove bound Ca^{2+} and return gelsolin to a negatively charged state. The pH of this mixture was adjusted to pH 7.8 prior to loading it onto an anion exchange column (36 x 4.5 cm) equilibrated with 25 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 7.8. Negatively charged gelsolin binds to...
the resin and later can be eluted with a salt gradient from 50 mM NaCl, 25 mM Tris-HCl, 1 mM EDTA, pH 7.8 to 350 mM NaCl, 25 mM Tris-HCl, 1 mM EDTA, pH 7.8 (750 ml per side). Elution was monitored with an ECONO UV monitor (Bio-Rad) at 280 nm. Then the eluant was dialyzed against two changes of 1.5 L 25 mM Tris-HCl, 1 mM EDTA, pH 8.0 at 4 °C. Finally, gelsolin was subjected to Affi-Gel Blue affinity chromatography in 25 mM Tris-HCl, 1 mM EDTA, pH 8.0. Gelsolin binds to the dye Cibacron Blue F3GA, attached to an inert agarose support matrix, and can be eluted by washing the column with 100 ml of 2.5 mM ATP, made up in the equilibration buffer. A fluorescence spectrometer (Perkin Elmer, model LS-5B) was used to detect gelsolin using an excitation wavelength of 280 nm and an emission wavelength of 340 nm. The eluant was concentrated and dialyzed against three changes of 1 L of 25 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0, over three days. The concentration of gelsolin was measured spectrophotometrically (Perkin Elmer Lambda 4B UV-Vis spectrometer) at a wavelength of 280 nm using an absorption coefficient of 1.4 ml mg⁻¹ cm⁻¹.

### 2.2.2 Human Cytoplasmic Full Length Gelsolin

Full length human gelsolin domains G1-G6 with an N-terminal 8 histidine tag was expressed in *E.coli*, BL21(DE3) from the expression vector, pSY5. First, 100 ml LB medium (10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl, pH 7.0) containing ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) was inoculated with gelsolin frozen stock culture (100 µl) and grown overnight at 37 °C with vigorous shaking. Then 1L LB medium was inoculated with 50 ml of this non-induced overnight culture. It was incubated in the shaker at 37 °C until an OD600 in the range of 0.6-0.8 was reached.
2.3 Villin Purification

At this point, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and was allowed to proceed 12 hr at 30 °C. Cells were harvested by centrifugation at 4500 rpm for 35 min using Beckman polypropylene bottles and an Avanti J-20 XP Beckman centrifuge. Then cells were resuspended in 20 ml of 50 mM Tris-HCl, 300 mM NaCl, 20 mM imidazole, pH 8.0 (buffer W). A cell-disrupter (high pressure homogenizer) (Avestin, Emulsiflex) was used to break-open the cells. Cell debris was removed by centrifugation at 20,000 rpm for 50 min (Beckman model optima L-90k ultracentrifuge and Beckman 45 Ti rotor). The supernatant was mixed with Ni Sepharose™ 6 Fast Flow medium (GE Healthcare) for 30 min and applied to a Ni-NTA column, and washed with buffer W at a rate of 1 ml/min to remove non-specifically bound proteins. Gelsolin with an 8 histidine tag, was eluted with 50 mM Tris-HCl, 300 mM NaCl, 300 mM imidazole, pH 8.0 (buffer E). Finally, gelsolin was purified by gel filtration (Bio-Rad Sephacryl S300; 90 x 2.5 cm) in 25 mM Tris-HCl, 1 mM EDTA, pH 8.0, monitored by an ECONO UV monitor (Bio Rad) at 280 nm.

2.3 Villin Purification

Villin domains V4-V6 with an N-terminal 8 histidine tag was expressed in *E.coli* strain BL21-Codonplus (DE3)-RIL cells (Stratagene) from the expression vector, pSY5. First, 30 ml LB medium containing ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml) was inoculated with V4-V6 frozen stock culture (100 µl) and grown overnight at 30 °C with vigorous shaking at 225 rpm. Then 1 L LB medium containing antibiotics was inoculated with 15 ml of this non-induced overnight culture and incubated in the shaker at 37 °C with 225 rpm shaking. Once the culture was grown to an OD600 of
2.4 Purification of Protein Complexes

0.7, 0.9 mM IPTG was added to induce the protein expression. The induced culture was incubated for 8 hours at 30 °C at 250 rpm shaking. Cells were harvested and disrupted in the same way as before. V4-V6 was first purified by Ni SepharoseTM 6 Fast Flow medium (GE Healthcare) and then by gel filtration (Bio-Rad Sephacryl S300; 90 x 2.5 cm) in 10 mM Tris-HCl, 150 mM NaCl and 5 mM CaCl₂, pH 7.5.

2.4 Purification of Protein Complexes

2.4.1 Equine Gelsolin and Actin Complex

The purification of the equine gelsolin/actin complex is based on a modified GA₂ purification method reported previously [38]. Gelsolin, in 25 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0, after dialysis, was incubated with 3 mM CaCl₂ at 4 °C for 5 min. Then gelsolin and actin were mixed at a molar ratio of 1:2 and incubated at 4 °C overnight. The GA₂ complex solution was run through a gel filtration column (Bio-Rad Sephacryl S300; 90 x 2.5 cm) at room temperature and eluted with buffer A at a rate of 2 ml/min. The GA₂ elution was monitored by a UV ECONO UV monitor (Bio-Rad) at 280 nm. The GA₂ fractions were combined and concentrated by RC-5B centrifugation at 3,000 rpm at 4 °C, first using Millipore centrifugal concentrators with a Sorvall GSA rotor, and then using Biosep centrifugal concentrators with a Sorvall SS34 rotor. The final concentration of GA₂ was 10 mg/ml, as determined by UV-Vis spectrophotometry (Perkin Elmer Lamda 4B) at 280 nm using a calculated absorption coefficient of 1.25 ml mg⁻¹ cm⁻¹. The purity of GA₂ was characterized by SDS-PAGE, which showed two bands as 83 kDa and 42 kDa.
2.4.2 Human Gelsolin and Actin Complex

The purification of the human gelsolin/actin complex is based on a modified GA purification method reported previously [38]. Recombinant human full-length gelsolin in 25 mM Tris-HCl, 1 mM EDTA, pH 8.0 purified by gel-filtration chromatography was incubated with 2 mM CaCl$_2$ at 4 °C for 5 min. Gelsolin was mixed with actin in buffer A at a molar ratio of 1:2 and then with 5 mM ethyleneglycol-bis-(β-aminoethyl)-N,N,N,N-tetraacetic acid (EGTA). On adding EGTA to a system containing GA$_2$, only one actin detaches from the complex under these conditions, leaving GA. Considering that the severing activity of the N-terminal half of gelsolin is independent of calcium control, it is assumed that the N-terminal part of gelsolin should be able to bind to one G-actin monomer in a calcium-free medium. After the resulting GA solution was incubated at 4 °C overnight, the complex was purified by gel-filtration (Bio-Rad Sephacryl S300; 90 x 2.5 cm) at room temperature and eluted with 10 mM Tris-HCl, 1 mM EGTA, 1mM DTT, pH 7.5 at a rate of 2 ml/min. The GA complex was concentrated as above. The final concentration of GA complex was 10 mg/ml as determined by UV-Vis absorption (Perkin Elmer Lambda 4B) at 280 nm using a calculated absorption coefficient of 1.3 ml mg$^{-1}$ cm$^{-1}$. SDS-PAGE showed two clean bands 81 kDa and 42 kDa, respectively.

2.4.3 Villin and Actin Complex

Villin V4-V6 in 10 mM Tris-HCl, 150 mM NaCl and 5 mM CaCl$_2$, pH 7.5, was mixed with actin in buffer A to a molar ratio of 1:1. The resulting solution was held at 4 °C overnight prior to gel filtration (Bio-Rad Sephacryl S300; 90 x 2.5 cm), with elution by buffer A. The fractions containing the
2.4. Purification of Protein Complexes

protein complex were pooled and concentrated to 10 mg/ml, as determined by UV absorbance at 280 nm (Perkin Elmer Lambda 4B) using a calculated absorption coefficient of 1.4 ml mg$^{-1}$ cm$^{-1}$. Both SDS-PAGE and Tandem mass spectrometry (MS/MS) (QStar XL Hybrid ESI quadrupole time of flight tandem mass spectrometer, Applied Biosystems/MDS Sciex) analyses indicated the fractions contained both villin V4-V6 and actin.

2.4.4 Hsp27 and Actin Complex

Hsp27 was generously provided by Dr. Grant Mauk’s group in the Department of Biochemistry and Molecular Biology at UBC. Hsp27 (20 mg/ml) in 20 mM Tris-HCl, 100 mM NaCl pH 8.4 was mixed with actin in buffer A at 1:1 molar ratio. The resulting solution was incubated at 4 $^\circ$C overnight and then run through a gel filtration column, as above, in buffer A at room temperature. The elution was monitored by a UV ECONO UV monitor (Bio-Rad) at 280 nm. Two peaks (a medium size peak followed by a larger one) were eluted at about 1.5 hrs. MS/MS (QStar XL Hybrid ESI quadrupole time of flight tandem mass spectrometer, Applied Biosystems/MDS Sciex) analyses indicated both peaks contained both Hsp27 and actin. The first and second peak fractions were separately concentrated to 10 mg/ml and 5 mg/ml, respectively, as determined by UV-Vis spectrometry (Perkin Elmer Lambda 4B) at 280 nm using a calculated absorption coefficient of 1.344 ml mg$^{-1}$ cm$^{-1}$.
2.5 Protein Crystallography

2.5.1 Crystallization of Proteins

Proteins are complex physical-chemical systems whose properties vary as a function of many environmental influences, and our current understanding of protein crystallization phenomena is limited [73]. We have to screen crystallization conditions by changing individual parameters that affect crystal formation, finding conditions that yield crystals, and then optimizing the variable sets to obtain the best possible crystals for X-ray analysis. Proteins generally have many degrees of freedom and their crystallization must be carried out in a way that maintains a stable structure. Also, a high purity protein, both chemically and conformationally, is necessary for crystallization.

At present, vapour diffusion, microbatch, dialysis and free interface diffusion are the most common methods used to crystallize proteins. In the vapor diffusion method (Figure 2.1B), water vaporizes from the sample droplet containing purified protein, buffer, and precipitant and condenses into the larger reservoir containing similar buffers and precipitants in higher concentrations. Thus, the salt and protein concentrations in the sample droplet increase, leading to supersaturation of the protein component. On moving from a supersaturated state to a saturated state, the system may exhibit crystal formation. For optimal use in X-ray diffraction studies, well-ordered crystals should grow in the metastable zone of the phase diagram, where no further nucleation would take place (Figure 2.1A).

In this thesis, I used the hanging-drop vapor diffusion method to crystallize the following proteins.
2.5. Protein Crystallography

Figure 2.1: A, Idealized path for crystallization by vapor diffusion method on a protein solution phase diagram; B, Hanging drop crystallization set-up

1 Equine G1-G3/actin crystals were grown at 4 °C from a GA_2 sample at 10 mg/ml mixed with a reservoir solution of 4% PEG6000 (w/v), 100 mM Na Hepes (pH 6.2) and 2 mM CaCl_2.

2 Human G1-G3/actin crystals were grown at 4 °C from a GA sample at 10 mg/ml mixed with a reservoir solution of 9% PEG4000 (w/v), 100 mM NaAc (pH 4.6) and 100 mM Ca(Ac)_2 or 2% PEG8000 (w/v), 100 mM NaAc (pH 4.6) and 2 mM CaCl_2.

3 V6 crystals were grown at 4 °C from a V4-V6/actin sample at 10 mg/ml mixed with a reservoir solution of 15% PEG8000 (w/v), 100 mM NaAc (pH 5.0).

4 Actin crystals were grown at 22 °C from a Hsp27/actin sample, at 5 mg/ml mixed with a reservoir solution of 2 M (NH_4)_2SO_4, 100 mM Tris-HCl (pH 8.5).
2.5. Protein Crystallography

Before data collection, crystal are soaked in appropriate cryoprotectant in order to prevent ice crystals from forming during the process of freezing. Good cryoprotectants are small molecules such as: glycerol, PEG400, methyl-pentanediol (MPD), and sucrose. The cryoprotectant solutions for freezing equine G1-G3/actin crystals, human G1-G3/actin crystals, V6 crystals and actin crystals were as follows:

1 Equine G1-G3/actin crystals in 25% glycerol (v/v), 4% PEG6000 (w/v), 100 mM Hepes (pH 6.2) and 2 mM CaCl$_2$.

2 Human G1-G3/actin crystals in 25% glycerol (v/v), 9% PEG4000 (w/v), 100 mM NaAc (pH 4.6) and 100 mM Ca(Ac)$_2$ or in 25% glycerol (v/v), 2% PEG8000 (w/v), 100 mM NaAc (pH 4.6) and 2 mM CaCl$_2$.

3 V6 crystals in 25% glycerol (v/v), 15% PEG8000 (w/v), 100 mM NaAc (pH 5.0).

4 Actin crystals in 25% glycerol (v/v), 2M (NH$_4$)$_2$SO$_4$, 100 mM Tris-HCl (pH 8.5).

Then crystals were frozen and maintained at 100 K under a stream of nitrogen to prevent radiation damage to the protein crystals during data collection.

2.5.2 Data Collection and Data Processing

After a crystal is mounted and exposed to an intense beam of X-rays, the diffraction spots are recorded on a screen with CCD detector behind the crystal. The crystal must be rotated step-by-step through an appropriate
angle (depending on symmetry), with an image recorded at every step, in order to collect a complete data set. In order to avoid causing a "blind spot" in reciprocal space close to the rotation axis, it is customary to rotate the crystal slightly (0.5-1 degree) to catch a broader region of reciprocal space. In this thesis, all the preliminary data sets were collected by using a Rigaku RU200 rotating anode source with OSMIC mirrors and a MAR345 image plate detector at the UBC Centre for Blood Research. Higher resolution data were collected on beamline BL13B on an ADSC Quantum 315 CCD detector at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan).

The first step in processing a data set is to index, i.e. identify cell dimensions and determine which image peak corresponds to which position in reciprocal space. The next step is to integrate, i.e. convert these images containing the thousands of reflections into a single file, consisting of records of the Miller indices of reflections, and intensities for these reflections. The final step is to merge and scale, i.e. identify which peaks appear in more than one images and scale the relative images in order that they have a consistent intensity scale. In this thesis, data processing was performed by using HKL2000 software [74].

2.5.3 Structure Determination and Refinement

The data collected from a diffraction experiment is a reciprocal space representation of the crystal lattice. The position of each diffraction spot is determined by the size and shape of the unit cell, and the symmetry within the crystal. The lattice dimensions of the reciprocal lattice correspond to the reflection positions and the intensity of the reflection corresponds to the
2.5. Protein Crystallography

The intensity of reflection is proportional to the square of the structure factor amplitude: $|F(h,k,l)|^2$. The structure factor $F(h,k,l)$ is a complex number containing information relating to both the amplitude and phase of a wave.

$$F(h,k,l) = \sum_{j=1}^{\text{atoms}} f(j) \exp\left[2\pi \cdot i (hx(j) + ky(j) + lz(j))\right]$$

where $f(j)$ is the scattering factor of atom $j$ and depends on the kind of atom and the diffraction angle of the corresponding reflection $(h,k,l)$. After Fourier transformation, the electron density can be calculated as follows, if both amplitude and phase are known.

$$\rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F(h,k,l) \exp[-2\pi \cdot i (hx + ky + lz)]$$

Although the amplitude can be obtained from the intensity of reflection easily, the phase cannot be directly recorded during a diffraction experiment. Their phases must be determined by Molecular Replacement (MR), Multiple Isomorphous Replacement (MIR) or Multi-wavelength/Single-wavelength Anomalous Diffraction (MAD/SAD) methods.

Molecular replacement uses a previously solved structure, which shares reasonable sequence identity (at least 40%) with the target structure, as a search model to determine the orientation and position of a molecule in the unit cell. Molecular replacement is generally divided into two stages: rotation and translation. A Patterson map is an interatomic vector map created by squaring the structure factor amplitudes without any phase information. The vectors in the Patterson map include two categories: intramolecular vectors and intermolecular vectors. Intramolecular vectors depend only on the orientation of the molecule, and are independent of its position in the unit cell, which can be exploited in the rotation search. Intermolecular vectors...
2.5. Protein Crystallography

depend both on the orientation of the molecule and on its position, which can be exploited in the translation search, once the orientation is known. In the rotation search, the orientation of the search model is determined that produces maximal overlap with the target structure by comparing Patterson map for the unknown with Patterson maps derived for the search model in different orientations. In the translation search, the search model with optimized orientation is positioned by translating it to the best-fit coordinates within the asymmetric unit by moving the model, calculating a new Patterson map, and comparing it to the unknown-derived Patterson map.

In this thesis, all the structures were solved by molecular replacement with known structures, such as G1-G3/actin (PDB Code: 1RGI), G4-G6 (PDB Code: 1P8X) and actin (PDB Code: 2GWJ), using Molrep [75] and Phaser [76], available through the Collaborative Computational Project Number 4 (CCP4) software package [77]. Having obtained initial phases, an initial model can be built by fitting the known protein amino acid sequence into the initial electron density map by using the interactive computer graphics program: COOT [78]. This model is refined until the observed (\(F_{\text{obs}}\)) structure factors agree with the ones calculated (\(F_{\text{calc}}\)) from the model. The agreement is measured by an R-factor defined as

\[
R = \frac{\sum |F_{\text{obs}}| - |F_{\text{calc}}|}{\sum |F_{\text{obs}}|}
\]

Here, structural refinement was performed by rigid body refinement, followed by restrained refinement, with REFMAC 5 software [79], which is available through the CCP4 package [77]. A refined model, with an R-factor in the range of 15% to 25%, is regarded as the actual protein structure.
Chapter 3

Results and Discussion

3.1 Structure of a Novel Equine Gelsolin and Actin Complex

The X-ray data from earlier preparations of GA₂ did not yield complete structures for one gelsolin bound to two actins, but produced only images of the G1-G3/actin portion of GA₂ [38]. Although only one actin is visible in our crystals, our GA₂ crystals exhibit three kinds of space group: P3₁2₁, P₆₅₂₂ and P₂₁. The first one is the same as that found for G1-G3/actin [7] and may indicate that proteolysis occurred during the time required for the crystals to nucleate. The second type of crystal possesses a larger unit cell (a = b = 146.3 Å, c = 388.8 Å), of sufficient volume to accommodate GA₂, but only the G1-G3/actin portion appeared sufficiently ordered to yield diffraction [28]. The third type of crystal, which was grown and solved most recently, gives a novel G1-G3/actin structure (Figure 3.1), despite missing the C-terminal half of gelsolin and the second actin (Table 3.1). In this novel structure, G1 is bound to the cleft between subdomains 1 and 3 of actin, as in the previously reported G1-G3/actin [7] and G1/actin [6] structures. The G1-G2 linker extends up the face of actin to allow G2 to contact actin subdomain 1, with a contact area of 697 Å². The G2-G3 linker positions G3 to bind back to actin subdomain 2 and subdomain 4, with a contact area.
of 312 Å². As in the previously reported G1-G3/actin structure [7], we can locate four Ca²⁺: a type I Ca²⁺ is located between G1 and actin; two type II Ca²⁺ are located in G1 and G3; another Ca²⁺ is associated with ATP in the nucleotide cleft. The absence of Ca²⁺ in G2 may again be argued to indicate transient binding in this site [7].

Figure 3.1: Schematic representation of the structure of a novel equine gelsolin/actin complex (G1, red; G2, green; G3, yellow; actin: cyan). The four black spheres represent Ca²⁺.

3.1.1 Comparison with Inactive G1-G3

This novel structure confirms the large conformational changes expected during transformation from inactive gelsolin [25] to activated gelsolin. The
### 3.1. Structure of a Novel Equine Gelsolin and Actin Complex

<table>
<thead>
<tr>
<th>Wavelength (Å)</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a=78.3Å, b=99.2Å, c=96.8Å, α=γ=90°, β=90.1°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.0-2.50 (2.56-2.50)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>48476 (3534)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.6 (3.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.52 (99.87)</td>
</tr>
<tr>
<td>Average I/I₀</td>
<td>24.5 (3.4)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)</td>
<td>4.8 (35.4)</td>
</tr>
<tr>
<td>R&lt;sub&gt;factor&lt;/sub&gt; (%)</td>
<td>22.7 (28.0)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%) (Based on 5% of the data)</td>
<td>27.5 (37.2)</td>
</tr>
<tr>
<td>Number of molecules in asymmetric unit</td>
<td>2</td>
</tr>
<tr>
<td>Gelsolin residue range</td>
<td>27-371</td>
</tr>
<tr>
<td>Actin residue range</td>
<td>5-39, 50-365, 375</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters)</td>
<td>11186 (131)</td>
</tr>
<tr>
<td>Mean derived B-factor (waters) (Å²)</td>
<td>63.3 (60)</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.013</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.515</td>
</tr>
</tbody>
</table>

Table 3.1: Data collection and refinement statistics for the novel equine gelsolin G1-G3/actin complex. Highest resolution shell is shown in parenthesis.
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

G1/G3 latch in inactive gelsolin is disrupted so that the actin-binding site in G1 is exposed. The extended G1-G2 linker translates G1 away from G2. The kinked long helix in G3 becomes straight.

3.1.2 Comparison with Previous G1-G3/Actin Structure

![Comparison of G1-G3/actin structures](image)

Figure 3.2: Comparison of (A) the novel G1-G3/actin structure with the (B) earlier G1-G3/actin structure (PDB code: 1RGI). G1, G2, G3 are colored red, green and yellow, respectively. Black spheres represent Ca$^{2+}$.

The novel G1-G3/actin structure is different from the earlier published equine G1-G3/actin structure in at least three aspects (Figure 3.2). The relative orientation of G2 and G3 between our structure and the previous one is strikingly different. The interface between G2 and G3 consists of a $\beta$-strand (residues 158-166) in G2 interacting with a $\beta$-strand (residues 284-289) in G3, and a loop (residues 176-180) in G2 with a $\beta$-strand (residues 284-289) in G3. The contact area between G2 and G3 is 434 Å$^2$, which is only half of the contact area of G2 and G3 (815 Å$^2$) in the earlier published
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

G1-G3/actin structure [7]. Additionally, in the new structure the interfaces on actin with G2 and G3 are located in subdomain 1 and subdomains 2 and 4, respectively. G2 masks 697 Å² of the surface of actin, which is twice as larger as that covered by G2 in the previous structure (290 Å²) [7]. The disulfide bond between Cys188 and Cys201 found in inactive plasma gelsolin is preserved in the new G1-G3/actin structure, while it is absent in the previous one.

3.1.3 Comparison with Mutant CapG

G2 and G3 in the novel structure resemble C2 and C3 in the structure of an activated CapG having mutations in an engineered severing domain (residues 84-91) and WH2-containing segment (residues 124-137) [59]. There are several similarities between our G1-G3 conformation and that of the mutant CapG. Firstly, the relative orientations of G2 and G3 resemble those of C2 and C3 in mutant CapG. The interaction interface for these two domains in both proteins consists of two adjacent β-sheets, and a loop from one domain with a β-sheet from the other. After alignment of G2-G3 from our novel G1-G3/actin with mutant GapG domains C2-C3, the root-mean-square deviation (RMSD) of Cα positions of C2 and C3 from the corresponding positions in G2 and G3 are 1.56 Å and 1.09 Å, respectively (Figure 3.3).

Secondly, both structures contain two type II Ca²⁺ or Eu³⁺ in conserved metal-binding sites located in domains 1 and 3. Thirdly, the type II metal ion-binding sites of domain 2 in both proteins are empty, and may have become so after the proteins become fully activated [7].

The novel G1-G3 structure differs from the mutant CapG structure in two aspects (Figure 3.4). The positions of domains 1 are different. This
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

can be explained by the fact that our G1-G3 is complexed with actin, which
binds G1 in the hydrophobic cleft between subdomains 1 and 3. In this way,
the conformation of G1 is restricted, while mutant CapG is crystallized in
the absence of actin. The missing type I Eu$^{3+}$ in domain 1 of mutant CapG
can be explained in this way as well. In addition, the G2-G3 linker in our
structure is ordered, while the linker (residues 235-244) between domains 2
and 3 of mutant CapG is disordered.

Figure 3.3: Alignment of G2-G3 from our novel G1-G3/actin with mutant
CapG C2-C3 (PDB code: 1J72). A, Superimposed on domain 2; B, Super-
imposed on domain 3. Gelsolin and mutant CapG are colored as red and
green, respectively.

3.1.4 Implication for Gelsolin-Capped F-actin Model

Based on the novel G1-G3/actin and G4-G6/actin structures (PDB code:
1H1V), a new gelsolin-capped actin filament model can be constructed by
superimposing these two structures onto the terminal protomers in the latest
F-actin model from high resolution X-ray fiber diffraction data [16]. In this
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

Figure 3.4: Comparison of novel G1-G3 with mutant CapG. A, our G1-G3 derived from the G1-G3/actin complex; B, The mutant CapG structure (PDB code: 1JHW). Black and cyan spheres represent Ca$^{2+}$ and Eu$^{3+}$, respectively.

new model (Figure 3.5), the interactions between the α-helix in G2 and the upper actin protomer predicted by the previous model [7] are absent. Thus the N-terminal half gelsolin only interacts with one actin protomer, via G1. It is apparent that the PIP$_2$ binding sites (residues 135-142, 160-172) are far more exposed in the new structure than those in the previous G1-G3/actin structure. Hence gelsolin in this new gelsolin-capped model seems primed to be removed. So this model might stand for one possible conformational stage before gelsolin is released from a capped filament.
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

Figure 3.5: Gelsolin-capped F-actin model based on the novel G1-G3/actin and G4-G6/actin (PDB code: 1H1V). Two protomers of F-actin (drawn in cyan) interact with G1-G3 (G1, red; G2, green; G3, yellow) and G4-G6 (G4, pink; G5, dark green; G6, orange). The third protomer (drawn in blue) does not interact with gelsolin. The F-actin coordinates are from the recent high resolution X-ray fiber diffraction model [16].

3.1.5 Implication for Actin Severing Mechanism

Based on the new gelsolin-capped F-actin model, we suggest a modified mechanism for actin binding, severing and capping by gelsolin (Figure 3.6). Initially, G2, which contains an F-actin binding site, binds to the side of an actin filament as in the original model [7], by interacting simultaneously with two longitudinally adjacent actin protomers. G2 and G3 orientations are as in the earlier published G1-G3/actin structure (PDB code: 1RGI). Then, G1 is directed by the G1-G2 linker to bind to the cleft between subdomains 1
3.1. Structure of a Novel Equine Gelsolin and Actin Complex

Figure 3.6: The new mechanism for actin binding, severing and capping by gelsolin (G3-G4 linker is not depicted). A, The initial recognition stage by G2. B, The bindings of G1, G3-G4 linker and G4 to actin. C, The resulting gelsolin-capped filament after severing.

and 3 of the lower actin protomer. The G3-G4 linker binds around the actin filament and directs G4 to bind another actin protomer across the filament from the first. Binding of G3-G4 linker and G4 exert strains to change the G2-G3 conformation to the novel G2-G3 one, as in the novel G1-G3/actin structure. Once the actin filament is severed by the pincer movements of G1 and G4, the gelsolin-capped actin filament is produced, which is the same as the new gelsolin capped F-actin model. Near the PIP$_2$-rich membranes, the interactions between gelsolin and actin are disrupted by the binding of PIP$_2$ to gelsolin. In the new structure, this would be easier than in the previously
reported one because of increased exposure of the PIP\textsubscript{2} binding sites in the G1-G2 linker and G2. The new model for a gelsolin-capped actin filament is primed for uncapping by PIP\textsubscript{2} in order to allow actin polymerization to proceed.

### 3.2 Structure of Human Gelsolin and Actin Complex

Although SDS-PAGE showed full length human gelsolin G1-G6 and actin to be present in the GA solution before crystallization, the X-ray structure from these GA crystals only reveals the N-terminal half of gelsolin binding to one actin (Figure 3.7). This indicates that the C-terminal half of gelsolin in this crystalline state is not well ordered, as in my previous studies [38]. So we regard this structure as human G1-G3 bound to actin in the presence of Ca\textsuperscript{2+} (Table 3.2) (PDB code: 3FFK). The overall structure of this recombinant human G1-G3/actin resembles the earlier equine G1-G3/actin structure [7]. G1 is bound to the cleft between subdomains 1 and 3 of actin. The G1-G2 linker and G2-G3 linker position G2 and G3 to contact actin subdomain 2 and subdomain 1, respectively. Compared with inactive gelsolin structure (PDB code: 1D0N), large-scale rearrangement of domains G1 and G3 relative to G2 happens in order to bind actin. The disruption of the central $\beta$-sheet of the G1/G3 latch and the removal of the first strand from the $\beta$-sheet in G2 exposes actin-binding surfaces and creates a new G2-G3 interface.
3.2. Structure of Human Gelsolin and Actin Complex

Figure 3.7: Schematic representation of the structure of human G1-G3/actin complex (G1, red; G2, green; G3, yellow; actin: cyan) (PDB code: 3FFK). The black spheres represent the five Ca$^{2+}$ associated G1-G3/actin (G1 type I, G1 type II, G2 type II, G3 type II, and Ca$^{2+}$ in the ATP cleft).

3.2.1 Comparison with Previous G1-G3/Actin Structure

The human G1-G3/actin structure is different from the earlier equine G1-G3/actin structure [7] in the following aspects (Figure 3.8). First of all, in the human G1-G3/actin, 5 Ca$^{2+}$ (G1 type I, G1 type II, G2 type II, G3 type II, and Ca$^{2+}$ in the ATP cleft) are present, while in equine G1-G3/actin, the G2 type II Ca$^{2+}$ is missing. Secondly, the disulfide bond between Cys188 and Cys201 is preserved as in inactive gelsolin structure [25], while the S-S bond is absent in the earlier equine G1-G3 [7]. Finally,
3.2. Structure of Human Gelsolin and Actin Complex

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.00</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>(a=101.3,\text{Å}, b=146.0,\text{Å}, c=148.1,\text{Å}, \alpha=\beta=\gamma=90^\circ)</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>30.0-3.0 (3.11-3.0)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>42382 (3009)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.1 (4.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.4 (97.5)</td>
</tr>
<tr>
<td>Average I/Io</td>
<td>13.6 (3.2)</td>
</tr>
<tr>
<td>(R_{\text{merge}}) (%)</td>
<td>9.7 (39.3)</td>
</tr>
<tr>
<td>(R_{\text{factor}}) (%)</td>
<td>22.0 (30.3)</td>
</tr>
<tr>
<td>(R_{\text{free}}) (%) (Based on 5% of the data)</td>
<td>27.5 (37.8)</td>
</tr>
<tr>
<td>Number of molecules in asymmetric unit</td>
<td>2</td>
</tr>
<tr>
<td>Gelsolin residue range</td>
<td>28-371 (Mol 1), 27-371 (Mol 2)</td>
</tr>
<tr>
<td>Actin residue range</td>
<td>3-39, 50-369 (Mol 1); 5-38, 50-370 (Mol 2)</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters)</td>
<td>11358 (313)</td>
</tr>
<tr>
<td>Mean derived B-factor (waters) (Å²)</td>
<td>42.4 (32.1)</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Table 3.2: Data collection and refinement statistics for human gelsolin G1-G3/actin complex. Highest resolution shell is shown in parenthesis. This table is taken from our paper [80].
3.2. Structure of Human Gelsolin and Actin Complex

Asp259, from the G2-G3 linker, participates in the coordination the Ca$^{2+}$ together with the conserved type II site residues (Asp187 and Glu209) and holds the G2-G3 linker in a stable conformation. Thus Thr260 in the linker forms a helix-initiating H-bond, which contracts the previously disordered G2-G3 linker into a helix and packs G3 with G2 more closely. In contrast, in the earlier equine G1-G3/actin, Asp259, without Ca$^{2+}$ bound, becomes the helix-initiating residue, allowing the G2-G3 linker to condense and bring G2 and G3 together.

3.2.2 Gelsolin Activation by Ca$^{2+}$

As shown in Figure 3.9, Ca$^{2+}$-binding to G2 and G6 directly competes for residues involved in the G2/G6 interface (Arg168 and Asp669, Arg169 and Asp670, Arg207 and Asp744 or Asp747) and induces conformational changes in both domains, which render them incompatible with the Ca-free structure. Ca$^{2+}$-binding to G2 potentially has a threefold effect on the initial activation of gelsolin via disordering the G2/G6 interface, altering contacts with the C-terminal tail, and promoting a conformational state in the G2-G3 linker that is incompatible for association with G6. In order to coordinate Ca$^{2+}$, Glu209 releases Arg168 and moves closer to Asp187, taking Asn206 with it. This leads to mobility in Arg168, and through proximity, in Arg169 resulting in the weakening of the G2-G6 contact. The slight straightening of the G2 long helix, to which Asn206 and Glu209 are attached, alters the position of Arg207. Since Arg207 interacts directly with the C-terminal tail, Ca$^{2+}$-binding to G2 may be a factor in releasing the tail during activation. Furthermore, participation of Asp259 in Ca$^{2+}$-coordination by G2 promotes order in the previously disordered G2-G3 linker. This introduces
3.2. Structure of Human Gelsolin and Actin Complex

a steric element into the activation process, since an ordered G2-G3 linker is incompatible with the position of G6 in the inactive structure.

Similarly, Ca\(^{2+}\)-binding to G6 has a direct effect on the G2/G6 interface. The carboxylate groups of Asp669 and Asp670 lock the G2/G6 interface through binding to the guanidinium group of Arg168 and Arg169, respectively. Direct competition by Ca\(^{2+}\) for the Asp670 sidechain and conformational restriction of Asp669 through coordination by its carboxylate destabilizes the G2/G6 interface. Furthermore, the straightening of the long helix of G6, required for the constitution of a functional Ca\(^{2+}\)-binding site, is achieved by the movement of the AB loop of G6 through formation of a hydrogen bond between Ser646 (from the AB loop) and Glu692 (from the Ca\(^{2+}\)-coordination sphere). The activated conformation of the G6 AB loop causes a steric clash with G2 and is, therefore, incompatible with Ca-free gelsolin.

Our structural analysis suggests that cooperative Ca\(^{2+}\)-binding to G2 and G6 is required to open the G2/G6 latch and initiate the large-scale conformational changes required for activation. This conclusion is justified by our thermal denaturation and actin depolymerization assays experiments [80]. But, eight Ca\(^{2+}\) have been identified to be associated either with gelsolin or at gelsolin/actin interfaces from crystallographic studies. Hence, many of these Ca\(^{2+}\)-binding sites are non-essential under experimental conditions. The presence of these apparently dispensable Ca\(^{2+}\)-binding sites can be justified by their possible effect on the rate of gelsolin action, achieved through stabilization of the optimal conformation and direct mediation of contact with actin at higher Ca\(^{2+}\) concentrations. Thus, gelsolin could be tuned to respond to Ca\(^{2+}\) levels over the four orders of magnitude of variation in Ca\(^{2+}\) levels found between the intra- and extracellular environments.
3.2. Structure of Human Gelsolin and Actin Complex

Figure 3.8: Comparison human G1-G3/actin complex with earlier published equine G1-G3/actin. A, Human cytoplasmic G1-G3/actin; B, Equine plasma G1-G3/actin (PDB code: 1RGI); C, Close up of the Ca-coordinating residues in G2 from human G1-G3/actin. The arrowhead points towards the peptide bond between Arg172 and Ala173, which gets cleaved in FAF; D, Close up of the Ca-coordinating residues in G2 from equine G1-G3/actin. The figure is taken and modified from our paper [80].
3.2. Structure of Human Gelsolin and Actin Complex

Figure 3.9: Structural interdependence of the G2 and G6 Ca\(^{2+}\)-binding sites. A, Schematic and electrostatic surface representations of Ca-free human gelsolin, highlighting the charged residues at the G2/G6 interface. By convention, surface painted blue is positively charged, while surface painted red is negatively charged; B, Schematic representations of Ca-bound G6 and G2, respectively taken from the structures of Ca-bound equine G4-G6/actin (PDB code: 1H1V) and human G1-G3/actin. The figure is taken from our paper [80].
The structure of Ca-free gelsolin strongly suggests that the vacant type II Ca-binding sites in G2 and G6 are set in electrostatic environments that attract cations and support local conformational rearrangements appropriate for springing the G2/G6 latch. Hence, the type II binding sites on G2 and G6 are likely to represent the 2 crucial Ca$^{2+}$ coordination sites for gelsolin functions.

### 3.2.3 Implications for FAF

Familial Amyloidosis Finnish-type (FAF), an inherited disease, is caused by a single mutation of Asp187 to Asn or Tyr in G2, which results in defects in Ca$^{2+}$-binding. FAF is characterized by the extracellular deposition of a 53-residue or 71-residue fragment of gelsolin. It has been proposed that this mutation destabilizes the gelsolin structure due to loss of Ca$^{2+}$-coordination at the type II site in G2, making it susceptible to proteolysis by furin located in the Golgi apparatus [27]. Two kinds of gelsolin fragment (Ala173-Met243 and Ala173-Arg225) are generated by subsequent extracellular proteolysis by β-gelsolinases. The accumulation of the fragments into amyloid fibrils leads to the symptoms of this disease [81].

In the human G1-G3/actin structure, Asp187 is part of the conserved calcium binding site in G2. After mutation, the G1-G6 mutants are unable to bind Ca$^{2+}$ at the G2 site, which is predicted to have three consequences with regard to Ca$^{2+}$-dependent activation. Firstly, the initial opening of the G2/G6 latch is affected because of the lack of the Ca$^{2+}$-induced conformational changes in G2 and the loss of cooperativity with the G6 Ca$^{2+}$-binding site, resulting in a decreased activation rate. Secondly, Ca$^{2+}$-induced stabilization of the G2-G3 module in the activated conformation is absent, pro-
longing the lifetime of intermediate conformations, where G1 has dissociated from G2 but G3 has not yet bound to G2. Lastly, because of impaired G2 Ca$^{2+}$-binding site in the mutant G2, the final conformation of G2-G3 would resemble that observed in the earlier reported equine G1-G3/actin structure (PDB code: 1RGI) (with unoccupied G2 Ca$^{2+}$-binding site) rather than the one in the human G1-G3/actin structure (with occupied G2 Ca$^{2+}$-binding site), if the mutant is able to become fully activated. The first step in the progression of the disease is the furin cleavage of G2 between Arg172 and Ala173 (Figure 3.8C). In inactive gelsolin this peptide bond is protected by flanking strands within the core $\beta$-sheet of G2. In the activated conformation the peptide bond is protected by G3, even in the absence of G2-bound Ca$^{2+}$. But in intermediate conformations, the cleavage site is accessible to the protease. Ca$^{2+}$-binding drives G2-G3 into its stable active conformation, minimizing the time spent in the intermediate states and reducing the possibility of cleavage by furin, which would set gelsolin on the course to fibril formation. Hence, we propose that the protease cleaves gelsolin during the activation stage when Arg172 and Ala173 become exposed, and that Ca$^{2+}$ binding to G2 in the wild type gelsolin protects it from cleavage by furin.

3.3 Structure of Villin Domain 6

The X-ray structure from V4-V6/actin complex crystals only reveals the villin V6 domain (Table 3.3) (PDB code: 3FG7). Since both SDS-PAGE and MS/MS results indicate that both V4-V6 and actin are present in solution before crystallization, the data suggest that proteolysis occurred during the time required for the crystals to nucleate. V6 (Figure 3.10), like a
3.3. Structure of Villin Domain 6

typical gelsolin-like domain, consists of a five-stranded $\beta$-sheet at its core sandwiched between a pair of $\alpha$-helices, a long helix that runs roughly parallel and a short helix that lies perpendicular to the strands. There is no calcium ion bound to V6 in this structure.

![Schematic representation of the structure of villin V6](PDB code: 3FG7). No bound metal ion is observed

3.3.1 Structural Comparison of V6 with G6

Since the amino acid sequence of V6 is approximately 50% identical with that of G6, the folding and functions of these two domains invite direct comparison. Of the six domains of gelsolin, the structure of V6 most closely resembles that of G6 in its Ca-bound form (PDB code: 1P8X). The RMSD of Ca positions is 3.0 Å (Figure 3.11). The long helix in each of V6 and the calcium ion-bound (Ca-bound) form of G6 is in an extended state, in
3.3. Structure of Villin Domain 6

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.00</td>
</tr>
<tr>
<td>Space group</td>
<td>P3₁</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a=b=47.82 Å, c=98.18 Å, α=β=90°, γ=120°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.00-2.00 (2.07-2.00)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>16865 (1676)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.8 (3.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.6 (98.8)</td>
</tr>
<tr>
<td>Average I/Io</td>
<td>11.8 (11.7)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)</td>
<td>4.8 (28.2)</td>
</tr>
<tr>
<td>R&lt;sub&gt;factor&lt;/sub&gt; (%)</td>
<td>19.7 (22.2)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%) (Based on 5% of the data)</td>
<td>24.6 (37.3)</td>
</tr>
<tr>
<td>Number of molecules in asymmetric unit</td>
<td>2</td>
</tr>
<tr>
<td>V6 residue range</td>
<td>635-740</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters)</td>
<td>1750 (160)</td>
</tr>
<tr>
<td>Mean derived B-factor (waters) (Å²)</td>
<td>38.0 (47.4)</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.013</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.499</td>
</tr>
</tbody>
</table>

Table 3.3: Data collection and refinement statistics for villin domain 6. Highest resolution shell is shown in parenthesis. This table is taken from our paper [82].
3.3. Structure of Villin Domain 6

contrast to the kinked state observed in G6 in the context of its intact calcium ion-free (Ca-free) parent protein. Comparison of the arrangement of the residues important in calcium binding in G6 (Ser646, Asn647, Asp670 and Glu692) with that of the homologous residues from V6 (Ser625, Asn626, Asp648 and Glu670) emphasizes the structural similarity to Ca-bound G6 (Figure 3.12). By contrast, in Ca-free G6, the positions and interactions of these residues are altered significantly. These comparisons suggest that the present structure of Ca-free V6 is that of an activated state despite the absence of bound calcium.

Figure 3.11: Alignment of our Ca-free V6 structure (orange) with the Ca-bound G6 structure (blue) from activated G4-G6 structure (PDB code: 1P8X).

3.3.2 V6 Interaction Interfaces

G6 is the central domain in inactive Ca-free gelsolin, forming direct contacts with all other domains. G6 interacts most extensively with G2 and G4. Of
3.3. Structure of Villin Domain 6

Figure 3.12: Structural comparison of villin V6 (Ca-free) with gelsolin G6 (Ca-bound and Ca-free). A, Ca-free V6; B, Key residues that are involved in the putative V6 calcium-binding site; C, Ca-bound G6; D, Key residues that are involved in the G6 calcium-binding site; E, Ca-free G6, taken from the structure of whole plasma gelsolin (PDB code: 1D0N); F, Residues involved in G6 calcium-binding are dislocated in the absence of calcium. This figure is taken from our paper [82].
particular note is a continuous $\beta$-sheet between G4 and G6 that seals the G4-G6 latch, obscuring the actin-binding site on G4. Binding of calcium triggers disruption of this latch. However, only subtle changes on binding calcium are observed in the G6 interface that binds to G4. The analogous surface on V6 is globally similar to G6 displaying a characteristic alternating charge distribution suggesting a $\beta$-sheet interaction with V4, similar to that observed between G4 and G6 (Figure 3.13). The loop that exits the domain 4 strand forms a small steric clash in the Ca-bound G6 and Ca-free V6 models, suggesting that binding to domain 4 will not occur in these two conformations. Surface residues at this interface, including residues involved in the steric clash, show conservation within, yet divergence between, gelsolins and villins. Taken together, these data suggest that the manner in which V6 interacts with V4 in villin is specific yet analogous to the situation in gelsolin between G6 and G4, and that calcium-free V6 in isolation adopts an activated conformation.

In inactive gelsolin, electrostatic interactions between a negatively charged surface on G6, which includes Asp670, with Arg168 and Arg169 (in G2) fasten the tail latch over the F-actin binding site of G2 [26]. Coordination of a calcium ion by G6 competes for the side chain of Asp670 leading to the disruption of these interactions. These residues are conserved in villin (Asp648, Lys145, Arg146, respectively). Furthermore, surface residues at the interface with the N-terminal half display conservation within, yet divergence between, gelsolins and villins. This suggests that V6 will form similar but specific interactions with V1-V3 when compared to those seen between G6 and G1-G3. Superposition of the Ca-bound structure of G6 and the Ca-free form of V6 onto the structure of Ca-free G6 (Figure 3.14), with respect to G2, allows assessment of the effects of the conformational differ-
3.3. Structure of Villin Domain 6

Figure 3.13: Comparison of the domain 4-interaction surfaces of G6 and V6. A, Representation of Ca-free gelsolin showing the central position of G6 within the inactive structure (PDB code: 1D0N). The black sphere pinpoints the unoccupied calcium-binding site within G6; B, Surface charge representation of G6 from A, rotated approximately 90 degrees around the vertical axis, showing the interaction surface with G4 via a common β-sheet; C, Ca-bound G6 superimposed on G6 in B; D, Ca-free V6 superimposed on G6 in B. This figure is taken from our paper [82].

ences on interactions with domain 2. The AB loops, which connect strands A and B, adopt different conformations in the presence of the kinked versus straight helices in domains 6. This movement in the Ca-bound G6 AB loop is sterically incompatible with binding to G2 and likely a major factor in the separation of G2 and G6 during calcium activation. Similarly, the AB loop in Ca-free isolated V6 shows steric clashes in the model presented, suggesting a similar activation mechanism for villin, and in reverse, movement of the AB loop, may be needed, possibly combined with helix kinking, to adopt an inactive conformation within whole villin.

In the G4-G6/actin complex (PDB code 1H1V), there is interaction between the AB loop of G6 and an α-helical section of peptide chain (Asn307
3.3. Structure of Villin Domain 6

Figure 3.14: Comparison of the domain 2-interaction surfaces of G6 and V6. A, Arg168 and Arg169, from G2, bind to Asp670 and a negatively charged surface on G6 in the Ca-free form of gelsolin. The AB-loop is distant from G2 in this conformation. B, Calcium-binding by G6 involves direct coordination to Asp670. Superposition of Ca-bound G6 on A reveals that the AB-loop moves to clash with G2 in this model. C, Superposition of Ca-free V6 on A creates a model that demonstrates that Asp648 is similarly placed to activate villin via calcium-binding or to secure the Ca-free structure through binding to V2. In this model the AB-loop clashes with domain 2 and, as such, would have to move to bind V2. This figure is taken from our paper [82].

through Asp321) in actin subdomain 3. This contact forms during the actin filament-severing process and contributes to disruption of actin-actin contacts during that process [26]. Superimposition of the coordinates for Ca-free G6 from inactive intact gelsolin onto those for Ca-bound G6 in the activated structure of the G4-G6/actin complex (Figure 3.15) reveals how a transformation of G6 in this complex to a Ca-free state would disturb the interaction with actin. However, superimposition of the coordinates of Ca-free V6 onto
3.3. Structure of Villin Domain 6

Figure 3.15: Domain 6 interactions with actin. A, The structure of G4-G6/actin (PDB code: 1H1V) shows that the actin helix comprising residues 307-321 contacts the AB-loop of G6. B, Surface charge representation of Ca-bound G6 showing the interaction with actin residues 307-321. Ca-bound G6 is oriented as in Fig. 1C. C, Superposition of Ca-free V6 onto B produces a model that suggests that V6 might retain the ability to interact with these actin residues. D, Similar modeling of the transformation of G6 to a Ca-free state disturbs that interaction. This figure is taken from our paper [82].

those for Ca-bound G6 in G4-G6/actin suggests that contact between V6 and the actin helix could be maintained. From this, we infer that domain 6 from this family of proteins, once freed of interactions that prevent straightening of its long helix, may bind actin and induce severing at suboptimal calcium concentrations.

3.3.3 Spring-Loaded Mechanism

The kink in the helix of the Ca-free G6 can be understood in terms of the context in which it was discovered, folded within the compact globular structure of intact Ca-free gelsolin. The bend in the long helix of G6 (and in that of its analogue, G3) exists for steric reasons, to avoid clashing with the
3.3. Structure of Villin Domain 6

long helix of the adjacent domain, G4 (G1, in the case of G3) in the folded protein. Our molecular dynamics calculations also suggest that this helix in the isolated G6 domain prefer to be straight in the absence of Ca\(^{2+}\) and other gelsolin domains [82]. In the present structure of V6, in the absence of V4, there is no need for the helix to assume the less energetically favored kinked form. The evidence above enforces predictions from sequence alignments that the structures of gelsolin and the V1-V6 body of villin are homologous. As such, the long helix of V6 within Ca-free villin may well be bent. We suggest that the bent helix behaves energetically like a cocked spring.

The role of binding calcium, in the context of the whole protein, is to release interdomain contacts. One consequence of this is the straightening of otherwise bent \(\alpha\)-helices. The potential energy held in the bent helices of the inactive proteins is tapped to do work during activation, opening latches that would otherwise block actin-binding surfaces.

3.3.4 Implication for Gelsolin Activation

Millimolar calcium ion concentrations, such as those present in blood plasma, activate gelsolin to be extremely efficient at severing undesirable extracellular actin filaments. These calcium levels cannot be attained intracellularly, yet gelsolin is able to function as an actin filament regulator in this environment. Furthermore, at low pH, gelsolin is able to sever actin in the complete absence of calcium, while villin has quite different calcium requirements to those of gelsolin [45]. The importance of calcium in activating this protein superfamily appears to vary with molecular and environmental setting, despite the conservation in calcium-binding residues.

The structure of Ca-free V6 and its similarity to that of Ca-bound G6,
suggests an explanation to this diversity of activating conditions. V6 and G6 adopt active conformations, in the absence of their other domains, through inbuilt spring mechanisms involving their long helices. During activation, the extended $\beta$-sheet shared between G4 and G6 and the interaction between G2 and G6 (and V4, V2 with V6) are torn apart. A portion of the energy cost to do so could be regained by straightening of the kinked helix of G6, and by analogy, V6. The remainder of the energy may be recouped from the binding of calcium ions and the formation of new contacts, such as between G5 and G6. Any activation, calcium or otherwise, that releases domain 6 from the other domains, within these proteins, will thus allow domain 6 to adopt this active conformation that can interact with actin. Hence, regardless of which particular circumstance initiates activation, domain 6 will be able to reach its activated state. In this regard, we speculate that the long helix in domain 6, and by analogy in domain 3, within this superfamily of proteins acts as a cocked spring, which can be triggered to straighten under a variety of circumstances.

### 3.4 Structure of G-Actin

X-ray crystallographic result from Hsp27/actin complex revealed each crystallographic asymmetric unit contains only one actin molecule, which is consistent with a solvent fraction of 60%. There is no electron density for Hsp27 present in crystal structure. However, MS/MS analysis (QStar XL Hybrid ESI quadrupole time of flight tandem mass spectrometer, Applied Biosystems/MDS Sciex) indicates that both Hsp27 and actin are present in crystallization drops which produced Hsp27/actin crystals. Since the unit cell provides enough space ($a=b=95.7\,\AA$, $c=96.9\,\AA$) to accommodate an
3.4. Structure of G-Actin

Figure 3.16: Schematic representation of the structure of G-actin in an ATP state (PDB code: 3HBT). Four subdomains of actin (colored cyan) come together to surround an ATP (ball-and-stick representation) and a Ca$^{2+}$ (black sphere) in the nucleotide cleft. The D-loop region in subdomain 2 is disordered. The structure was obtained with Hsp27 in the protein solution drop.

Hsp27 molecule in addition to one actin (still retaining a reasonable solvent percentage), we can not exclude the possibility that Hsp27 is present, although either disordered or adopting multiple low occupancy sites, resulting in a lack of attributable electron density. Based on the available structural information, we claim our structure to be G-actin in an ATP state, without ABPs or modification (Table 3.4) (PDB code: 3HBT). The overall crystal structure of our G-actin is very similar to those previously presented for
3.4. Structure of G-Actin

chemically modification G-actin, for G-actin bound in protein complexes, and for crosslinked G-actin structures [15]. In our structure (Figure 3.16), one ATP molecule and one Ca$^{2+}$ are located in the nucleotide cleft, which exists in a closed state. The D-loop is disordered, as in most actin structures. When superimposing our structure with previous ones (Figure 3.17), RMSD values for Cα positions are in the 1 Å range (Table 3.5).

Figure 3.17: Alignment of our actin structure with several other structures of actin. These actins are colored: the present structure (cyan), ADPR-actin (red), TMR-actin (blue), F-actin protomer (yellow), G1-G3/actin (green). The PDB codes for these structures are listed in Table 3.5.

3.4.1 Comparison with F-actin Protomer

The F-actin protomer structure is derived from the high resolution X-ray fiber diffraction F-actin model published recently [16]. By comparison of our
3.4. Structure of G-Actin

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>1.00</td>
</tr>
<tr>
<td>Space group</td>
<td>P3221</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a=b=95.7Å, c=96.9Å, α=β=90°, γ=120°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.0-2.70 (2.80-2.70)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>13778 (1348)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.6 (1.8)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.1 (93.1)</td>
</tr>
<tr>
<td>Average I/Io</td>
<td>10.8 (1.6)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>8.9 (40.0)</td>
</tr>
<tr>
<td>Rfactor (%)</td>
<td>20.3 (28.4)</td>
</tr>
<tr>
<td>Rfree (%) (Based on 5% of the data)</td>
<td>26.6 (35.5)</td>
</tr>
<tr>
<td>Number of molecules in asymmetric unit</td>
<td>1</td>
</tr>
<tr>
<td>Actin residue range</td>
<td>5-39, 51-375</td>
</tr>
<tr>
<td>Non-hydrogen atoms (calcium ions)</td>
<td>2931 (1)</td>
</tr>
<tr>
<td>Mean derived B-factor (waters) (Å²)</td>
<td>47.8 (42.1)</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.011</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.37</td>
</tr>
</tbody>
</table>

Table 3.4: Data collection and refinement statistics for the native G-actin.

Highest resolution shell is shown in parenthesis.
### 3.4. Structure of G-Actin

<table>
<thead>
<tr>
<th>Actin Structure (PDB code)</th>
<th>RMSD (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADPR-actin (hexagonal) (2GWJ)</td>
<td>0.224</td>
</tr>
<tr>
<td>ADPR-actin dimer (Orthorhombic) (2GWK, Chain A)</td>
<td>0.286</td>
</tr>
<tr>
<td>ADPR-actin dimer (Orthorhombic) (2GWK, Chain B)</td>
<td>0.292</td>
</tr>
<tr>
<td>TMR-actin (ADP) (1J6Z)</td>
<td>0.665</td>
</tr>
<tr>
<td>F-actin protomer (2ZWH)</td>
<td>1.440</td>
</tr>
<tr>
<td>G1-G3/actin (1RGI)</td>
<td>0.449</td>
</tr>
<tr>
<td>DNase I/actin (1ATN)</td>
<td>0.497</td>
</tr>
<tr>
<td>TMR-actin (AMPPNP) (1NWK)</td>
<td>0.428</td>
</tr>
<tr>
<td>Actin dimer with LATRUNCULIN A (1LCU, chain A)</td>
<td>0.620</td>
</tr>
<tr>
<td>Actin dimer with LATRUNCULIN A (1LCU, chain B)</td>
<td>0.592</td>
</tr>
</tbody>
</table>

|Table 3.5: RMS deviations of actin Cα positions from alignments of our native actin with other actin structures|
3.4. Structure of G-Actin

G-actin with the F-actin protomer, two dramatic differences are obvious (Figure 3.18). First, the D-loop region (His40-Gly48) in subdomain 2 of the F-actin protomer is present as an open loop, while the same region in our structure is disordered. The difference may be attributed to the difference in the nucleotide bound to the nucleotide cleft (ADP in F-actin protomer versus ATP in our structure) [13], and to the difference in the respective interaction environments (the F-actin protomer is in the context of an actin filament). From the high resolution F-actin model, the D-loop region is located in the hydrophobic groove between subdomains 1 and 3 of an adjacent actin protomer. By interacting Val43 and Met44 of one protomer with Leu346 and Phe375 of next, the D-loop becomes stabilized. The loss of such interaction in our G-actin molecule explains its disordered state, as in the majority of actin structures. The second difference lies in the twist angle between the two halves of the actin structure in our structure and the flat conformation in the F-actin protomer. After superimposing both structures on subdomain 4, a rotation of subdomains 1 and 2 relative to subdomains 3 and 4 is required to transition between G- and F-actin forms of the protomer [16].

The G to F-actin conformational transition has been modelled by comparing the G-actin-TMR structure in an ADP state with the F-actin protomer [16]. Based on the fact that G-actin exists predominantly in an ATP state in cytoplasm, the conformational transition between G-actin in an ATP state and the F-actin protomer may be more relevant physiologically. The investigation of the boundary between the two major domains during the G to F-actin transition shows many side chain rearrangements are involved (Figure 3.19). Firstly, the interaction between Pro109 and Val163 in G-actin is released after transition to a flat conformation; secondly, new salt
3.4. Structure of G-Actin

Figure 3.18: Structural comparison our G-actin with the F-actin protomer. A, The structure of native G-actin with the subdomains labelled. B, The structure of the F-actin protomer (PDB code: 2ZWH) for comparison, oriented to maximize overlap at subdomain 4. C, 90 degree rotation around a vertical axis relative to A. D, 90 degree rotation around a vertical axis relative to B. The arrow indicates the twist of subdomains 1 and 2, relative to subdomains 3 and 4, needed to transition between G- and F-actin.

Bridge interactions form, involving Glu72 and Arg206, Asp187 and Arg183, Asp179 and Arg177 in the F-actin protomer, while such interactions do not exist in G-actin.

3.4.2 Comparison with ADP-ribosylated Actin

The high degree of similarity between our G-actin and ADP-ribosylated G-actin invites comment. ADP-ribosylation of Arg177 inhibits the polymerization of G-actin, which provides a way to crystallize G-actin. We compared our G-actin with the ADP-ribosylated actin structure from both monomer and dimer forms [15] (Figure 3.20). From the crystal packing pattern of our G-actin (P3_2121 crystal form), a crystallographic two fold rotation operation causes the close contact of the C-termini of one actin molecule with a symmetry-related molecule. The distance between sulfur atoms of the
3.4. Structure of G-Actin

Figure 3.19: Sidechain rearrangements at the boundary between the two major domains involved in the G to F-actin conformational transition. A, the present G-actin in an ATP state; B, the F-actin protomer (PDB code: 2ZWH)

two Cys374 residues is 3.53 Å. In the ADP-ribosylated actin as a monomer (PDB code: 2GWJ, P3_21 crystal form), the distance between sulfur atoms of the corresponding two Cys374 residues is 3.60 Å. In the disulfide-linked dimer form of ADP-ribosylated G-actin (PDB code: 2GWK, P2_12_12 crystal form), a covalent bond is evident between the two Cys374 residues, with a bond length of 2.21 Å. Despite the similarities, differences exist. In the ADP-ribosylated actin structures, although the density for ADP-ribose was not visible, the authors used SpvB (an ADP-ribosyltransferases) to cova-
3.4. Structure of G-Actin

Figure 3.20: Structural comparison our G-actin with ADP-ribosylated actin. Two molecules related by crystallographic symmetry are detailed in red and purple. The enlargement shows the close contact of the C-termini, where the Cys374 to Cys374 distance, between sulfur atoms, is 3.53 Å. Comparisons of this region with ADP-ribosylated actin as a monomer (PDB code: 2GWJ) or disulfide-linked dimer (PDB code: 2GWK) are indicated.

 lently transfer ADP-ribose to actin before crystallization [15]. MS experiment and higher temperature factors in the Arg177 sidechain distal atoms confirmed the presence of ADP-ribose. In our G-actin, which was not ADP-ribosylated, the temperature factors for the Arg177 sidechain distal atoms averaged 55 Å², which is comparable to those proximal to the backbone and to those for the same atoms in other reported structures of G-actin.

 Although the Hsp27 was used for cocrystallization, Hsp27 is not present in the crystals that grew. Given the chaperone activities of Hsp27 with respect to actin, Hsp27 appears to have provided a mechanism to remove
actin monomer from the bulk solution by crystallization rather than polymerization.

### 3.4.3 Implication for Actin ATPase Activity

Actin is a member of a superfamily of ATPases that consist of two domains connected by a hinge, with a nucleotide binding site located in the cleft between the two domains. The transition between closed and open states of the nucleotide binding cleft allows nucleotide change. In actin, Gln137 plays an important role in the ATPase activity. In the nucleotide cleft of our structure (Figure 3.21), Gln137 interacts with ATP indirectly by hydrogen bonding to a water molecule, one of five in the coordination sphere of the Ca adjacent to the $\gamma$-phosphate and $\beta$-phosphate of ATP. Three additional water molecules participate in a H-bond network that involves the $\gamma$-phosphate of ATP and the side chains of Asp154, Gln137 and His161. Such electrostatic interactions in the cleft stabilize the twisted conformation and slow the ATPase rate of actin [16]. But in the nucleotide cleft of the F-actin protomer, Gln137 interacts directly with the bound calcium ion, together with the ADP $\beta$-phosphate. Although the resolution of the structure does not allow inclusion of the water molecules that complete the coordination sphere, the conformational transition between twisted domains in G-actin and a flattened F-actin protomer undoubtedly contributes to the enhancement of ATP hydrolysis that is displayed by F-actin, probably through redeployment of the set of interactions that involve Gln137.
3.4. Structure of G-Actin

Figure 3.21: Comparisons of nucleotide interactions in our native G-actin and F-actin protomer. A, ATP interactions with G-actin close to the γ-phosphate. The calcium ion (black) is coordinated by the γ-phosphate and β-phosphate from ATP and 5 conserved water molecules (green). Three additional conserved water molecules that lie close to the γ-phosphate (purple). B, ADP interactions with the F-actin protomer. Gln137 and the ADP β-phosphate coordinate the calcium ion. The resolution of the structure does not allow inclusion of the water molecules that complete the coordination sphere.

3.4.4 Implication for G-actin to F-actin transition

Our G-actin structure confirms the twisted relative conformation of the two domains to be characteristic of monomeric, unmodified native G-actin, and that a relative rotation between them must occur to give the flattened conformation observed for an F-actin protomer. Comparing our G-actin structure with that of the F-actin protomer suggests a possible mechanism for the transition between G- and F-actin. Initially, G-actin with bound ATP
exists in a twisted state *in vivo*. Once G-actin adds to the barbed end of an actin filament, the conformation of that protomer changes from a twisted state to a flat state by a relative 20 degree rotation between the two domains in order to strengthen the interactions between F-actin protomers, e.g., as discussed for the interprotomer interactions of the D-loop. Also, as a result of this change, actin catalyses ATP hydrolysis to ADP and Pi by orienting and repositioning Gln137 relative to the $\beta$- and $\gamma$-phosphates of ATP. Finally, the release of Pi leaves the ADP-F-actin structure that characterizes the pointed end of actin filaments, and which is prone to interact with activators of filament disassembly, such as gelsolin and coflin.
Chapter 4

Conclusions and Future Work

4.1 Conclusions

By combining molecular biology, protein purification and X-ray crystallography, we succeeded in solving structures for two gelsolin/actin complexes, a villin fragment and G-actin. These structures provide us with new insights into the mechanisms of actin filament assembly and disassembly, regulatory activities played by gelsolin/villin superfamily proteins, and activation mechanisms for superfamily members.

Our novel equine G1-G3/actin structure illuminates new gelsolin/actin interfaces and a G2-G3 orientation that resembles mutant CapG. After superimposing the novel G1-G3/actin conformation and the G4-G6/actin structure onto the terminal protomers in the new F-actin model, we arrive at a new gelsolin-capped actin filament model, which may be used to modify our previous proposal for the mechanism for actin binding, severing and capping by gelsolin.

Our human G1-G3/actin structure reveals a calcium ion bound to the type II binding site in G2, which was vacant in a previously published equine G1-G3/actin structure. Structural studies show cooperative binding of Ca$^{2+}$
in G2 and G6 is required for gelsolin initial activation. The loss of Ca\textsuperscript{2+} -binding activity by G2 prolongs the lifetime of partially activated intermediate conformations in which the furin-susceptible cleavage site is exposed.

Our villin domain V6 structure displays a straight helix, as in the activated form of G6, even in the absence of Ca\textsuperscript{2+}. Comparision of V6 with G6 in intact calcium ion-free gelsolin and G6 in a calcium ion-bound state suggests that V6 adopts an active conformation in the absence of the other villin domains through an inbuilt spring mechanism. Within the intact protein, a bent V6 helix, when triggered by Ca\textsuperscript{2+}, straightens and helps push apart adjacent domains to expose actin-binding sites. Within inactive gelsolin superfamily proteins, the long helix in domain 6 (or in domain 3) acts as a cocked spring.

Our actin monomer presents the first actin structure that is neither chemically modified nor in a complex with ABPs. Structural comparision of our G-actin with the F-actin protomer suggests a possbile mechanism for the transition between G- to F-actin. As G-actin assembles onto a growing actin filament, the conformation of that protomer changes from a twisted to a flat state by a relative 20 degree rotation between the domains. As a result of redeployment of Gln137 in the F-actin structure, the rate of actin-catalysed ATPase activity is enhanced by three order of magnitude.

4.2 Future Work

Although we made progress in structural studies of actin and gelsolin/villin family members, many questions are not answered yet. Firstly, the structures for both intact activated gelsolin and intact gelsolin/actin complexes are still unavailable. All my gelsolin/actin complexes reveal only the N-
terminal half of gelsolin bound to one actin, with the C-terminal half of
gelolin and its bound actin absent. So, searching for new conditions to crys-
tallize gelsolin/actin complexes must continue, with the goal being to provide
a full structure of activated gelsolin with actin. Our gelsolin-capped actin
filament models still need more experimental evidence. Cocrystallization of
gelsolin with cross-linked actin dimers may help test these models. For ex-
ample, the crosslinker N-(4-azido-2-nitrophenyl) putrescine (ANP) [83] can
crosslink Gln41 and Cys374 from two adjacent actin protomers to give an
actin dimer. Since the two actins in this dimer have the same orientation as
in F-actin, the structure of such a gelsolin/actin dimer would provide new
evidence to test our gelsolin-capped actin filament models.

Secondly, the details of the actin filament capping and severing mecha-
nisms of gelsolin are still not clear. It is also unclear how the length of linker
connecting the N-terminal and C-terminal halves (i.e. G3-G4 linker) affects
the actin-binding, severing and capping activities of gelsolin/villin family
members. So the structural role of the G3-G4 linker must be determined. It
would be useful to clone, express and purify recombinant gelsolin fragments
(e.g. G2-G4, G3-G4, etc), crystallize them in the presence and absence of
calcium ions, and investigate their interactions with actin. Although I did
not manage to express the G2-G4 portion of gelsolin in soluble form using
either pSY5 or pSY7 vectors, further optimization of the purification con-
ditions, e.g., by changing buffers and/or adding specific salts or metal ions,
might succeed. The structure of G3-G4 as an isolated fragment can be stud-
ied by both X-ray crystallography and NMR, if a soluble fragment can be
obtained. Future work should also pay attention to studying gelsolin/actin
complexes in which the length of G3-G4 linker is shortened by recombinant
DNA techniques.
Fourthly, the mechanism of interaction between villin (or villin fragment), and actin is not known. Villin is unique because it has an actin-bundling function, in addition to capping, servering and nucleating functions. A recent villin-crosslinked actin filament model [55] from electron microscopy studies suggests novel interaction between V1, V2 and actin. More experiments will be needed to test this model. Cocrystallization of villin (or villin fragment) with actin could clarify the proposed model. In parallel with this goal, villin also should be studied in the same way as gelsolin, e.g. by crystallization with different V3-V4 linker lengths.

In regard to actin, it is not known how the delivery of G-actin to the barbed end of a filament drives the observed change in the orientation of the two domains from twisted to flat. Perhaps molecular dynamics studies...
4.2. Future Work

will help answer this question.
Bibliography


Chapter 4. Bibliography


[27] S. L. Kazmirski, R. L. Isaacson, C. An, A. Buckle, C. M. Johnson,
Chapter 4. Bibliography


[82] H. Wang, S. Chumnarnsilpa, A. Loonchanta, Q. Li, L. Kwan, S. Robine, M. Larssen, I. Mihalek, L. D. Burtnick, and R. C. Robinson. Helix straightening as an activation mechanism in the gelsolin superfamily of

Appendix A

The Human Cytoplasmic Gelsolin G2-G4 Fragment

A.1 Construction of the G2-G4 pSY5 plasmid

Based on the known DNA sequence of the human gelsolin gene (GenBank Accession No.CAA28000), specific primers containing SfiI and EcoRI restriction recognition sites for domains G2 to G4, were designed to encompass from Gly137 in gelsolin G2 to Leu511 in gelsolin G4. The PCR reaction was carried out in a 50 µl solution that consisted of plasmid template (100 ng), 1 µl (10 µM) of each primer, 1 µl (10 mM) of dNTPs, 5 µl of 10x PCR buffer (200 mM Tris-HCl, pH 8.8 at 25 °C, 100 mM KCl, 100 mM (NH₄)₂SO₄, 20 mM MgSO₄, 1 mg/ml nuclease-free BSA, 1% Triton X-100), and 3 U of Taq DNA polymerase (Biolabs). The forward primer is G2 5’-GGCCCGGGCGGCGCCGCTGAAGTACAAGAAAGG-3’, and the reverse one is G4 5’-GAATTCTTACAGGCTCATGAGGTGGTGCGGGGCTCC-3’. An initial cycle of 5 min at 94 °C, followed by 30 cycles of denaturation for 1 min at 94 °C, annealing for 1 min at 54 °C, and elongation for 1 min 50 s at 72 °C, and a final extension step at 72 °C for 7 min were performed in an automated thermal cycler (MJ Research PTC-200 Peltier Thermal Cycler). The amplification products were analyzed by electrophoresis on a
A.1. Construction of the G2-G4 pSY5 plasmid

1% agarose gel stained with ethidium bromide. Specific PCR products of approximately 1125-bp (1.125 kbp) were obtained and cloned into pSY5, a modified pET21d plasmid expression vector designed to encode an N-terminal eight-histidine tag followed by a PreScission protease cleavage site. First, the PCR product was purified by PCR purification kit (Qiagen). Both the pSY5 plasmid vector and PCR product were digested by SfiI and EcoRI restriction enzymes (Fermentas) according to the company’s instructions. Ligation was performed in 30 ml of reaction buffer. 100 ng of the digested vector was incubated together with 100 ng of PCR product in a solution containing 3 ml of 10x T4-DNA ligase buffer (300 mM Tris-HCl, pH 7.8 at 25 °C, 100 mM MgCl2, 100 mM DTT, and 10 mM ATP) and 3 U of T4-DNA ligase enzyme (Promega) at room temperature overnight. The recombinant G2-G4 plasmid was initially propagated in *E. coli* XL1Blue (Novagen). 2 μl of ligation solution was added to 50 ul of XL1Blue competent cells. The transformation solution was incubated on ice for 30 min, heated pulse for 45 sec at 42 °C, and immediately placed on ice for 2 min. After transformation, the solution was mixed with 250 μl of LB, and incubated at 37 °C at 225 rpm for 1 hr. The mixture was spread on LB/Agar/Antibiotics plate and incubated upside down at 37°C for 12 hrs. The colonies from the LB/Agar/Antibiotics plate were picked up and inoculated with 5 ml of LB and 100 μg/ml ampicillin. After incubation overnight at 37 °C, while shaking at 225 rpm, recombinant G2-G4 plasmids were prepared by Miniprep kit (Qiagen) according to the company’s manual. Once the identity of the cloned G2-G4 was confirmed by DNA sequencing, the recombinant G2-G4 plasmid was then propagated in *E. coli* Rosetta cells (Novagen) using the same procedures as above. The overnight culture grown from the correct colonies on the plate was mixed with 25% glycerol and stored as a frozen
A.2. Construction of the G2-G4 pSY7 plasmid

G2-G4 was initially overexpressed from the pSY5 plasmid as described for full-length gelsolin. However, all G2-G4 protein was found in the inclusion body regardless of induction temperature: 37 °C, 30 °C, room temperature. Therefore, pSY7, a modified pET21d plasmid expression vector designed to encode an N-terminal MBP (Maltose Binding Protein), and an eight-histidine tag followed by a PreScission protease cleavage site, was used in order to increase the solubility of G2-G4. The G2-G4 in pSY5 plasmid was digested by SfiI and EcoRI restriction enzymes (Fermentas) and purified by Gel extraction kit (Qiagen) according to the company’s manual. The ligation between the G2-G4 insert and digested pSY7 plasmid vector was performed as described above. The new recombinant G2-G4 plasmid was initially propagated in E.coli XL1Blue cells (Novagen) and then in E. coli Rosetta cells (Novagen) as described above.

A.3. G2-G4 Expression and Purification

After optimization, G2-G4 was overexpressed from the pSY7 plasmid in E.coli Rosetta cells (Stratagene). 30 ml LB medium containing 100 mg/ml ampicillin and 34 mg/ml chloramphenicol was inoculated with G2-G4 frozen stock culture (100 µl) and grown overnight at 30 °C with shaking at 225 rpm. The next day, 1 L LB medium was inoculated with 15 ml of this non-induced overnight culture. It was incubated in the shaker at 37 °C until an OD600 in the range of 0.6-0.8 was reached. Once protein expression was induced with
A.3. G2-G4 Expression and Purification

0.9 mM IPTG, the culture was incubated for 8 hours at 30 °C. Cells were harvested and disrupted in the same way as above. G2-G4-MBP was firstly purified by Ni Sepharose™ 6 Fast Flow medium (GE Healthcare) followed by dialysis against 50 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA and 1 mM DTT (buffer P) overnight. Self-expressed PreScission protease was added to G2-G4-MBP to cleave the G2-G4 and MBP. The pH of this solution was adjusted to pH 8.0 by dialysis against 50 mM Tris-HCl, 1 mM EDTA, pH 8.0, and then it was loaded into an anion exchange column (36 x 2 cm) equilibrated with 50 mM Tris-HCl, 1 mM EDTA, pH 8.0. The protein was eluted with a salt gradient from 50 mM Tris-HCl, 1 mM EDTA, pH 8.0 to 1 M NaCl, 50 mM Tris-HCl, 1 mM EDTA, pH 8.0, as two separate peaks. Each peak was further purified by gel filtration (Bio-Rad Sephacryl S300; 90 x 2.5 cm) in 10 mM Hepes, 100 mM NaCl and 0.1 mM EGTA, pH 7.5. Unfortunately, the purified fractions only contained MBP, as confirmed by MS/MS analysis (QStar XL Hybrid ESI quadrupole time of flight tandem mass spectrometer, Applied Biosystems/MDS Sciex). This indicates that G2-G4 fragment precipitates upon being G2-G4 cleaved from MBP.
Appendix B

Structure of Equine Serum Albumin

During purification of equine gelsolin from serum, ATP is used to elute gelsolin from an Affi-Gel Blue affinity chromatography column. However, sometimes ATP failed to elute a detectable gelsolin fraction. In such situations, 1 M NaCl was applied to elute bound protein from the column. The gelsolin prepared in this way contains a major contaminant at a molar mass of 67 kDa. Initially, I took this contaminant to be a proteolytically cleaved gelsolin fragment (68 kDa) and attempted to crystallize it. Crystals grew at 4 °C, from a 10 mg/ml protein solution mixed with a reservoir solution of 2 M (NH$_4$)$_2$SO$_4$, 100 mM NaAc (pH 4.60), 5 mM EGTA. Prior to X-ray data collection, the crystals were transferred into a cryoprotectant solution, 25% glycerol, 2 M (NH$_4$)$_2$SO$_4$, 100 mM NaAc (pH 4.60), 5 mM EGTA, and flash frozen in liquid nitrogen. The diffraction from these crystals is consistent with the space group P6$_1$, with unit cell parameters of a=b=96.6 Å, c=144 Å; α=β=90°, γ=120°. The protein structure determined from these crystals is that of equine serum albumin, and is very similar to that of human serum albumin (PDB code: 1AO6). It consists of three homologous domains, which each comprise two similar-folded subdomains.
Figure B.1: Schematic representation of the structure of equine serum albumin
### Appendix B. Structure of Equine Serum Albumin

<table>
<thead>
<tr>
<th>Wavelength (Å)</th>
<th>1.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P6₁</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a=b=96.6 Å, c=144 Å, α=β=90°, γ=120°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>23.85-2.45 (2.53-2.45)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>27864 (1400)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.5 (3.7)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Average I/I₀</td>
<td>28.7 (3.6)</td>
</tr>
<tr>
<td>Rₘₑᵣₑ₂ (%)</td>
<td>6.6 (31.4)</td>
</tr>
<tr>
<td>Rₓᵣᵢᵣ (%) (Based on 5% of the data)</td>
<td>25.9 (32.4)</td>
</tr>
<tr>
<td>albumin residue range</td>
<td>5-583</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters)</td>
<td>4534 (196)</td>
</tr>
<tr>
<td>Mean derived B-factor (waters) (Å²)</td>
<td>44.21 (41.83)</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.008</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.130</td>
</tr>
</tbody>
</table>

Table B.1: Data collection and refinement statistics for equine serum albumin. Highest resolution shell is shown in parenthesis.