Structural Investigation of Gelsolin Superfamily Members

by

Chenguang Jiang

B.Sc., East China University of Science and Technology, 2003
M.Sc., University of Regina, 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)
February 2012

© Chenguang Jiang, 2012
Abstract

We expressed three recombinant proteins in *E. coli* and purified them for analysis. Human gelsolin fragment G2-G3 retains the ability of intact gelsolin to bind to and co-sediment with filamentous actin. Tryptophan fluorescence intensity from G2-G3 is Ca\(^{2+}\)-sensitive, increasing with Ca\(^{2+}\) concentration in two steps, suggesting that the binding of Ca\(^{2+}\) at two sites induces a structural change that reduces the solvent exposure of tryptophan side chains. Iodide quenching studies of tryptophan fluorescence from the fragment indicate reduced effectiveness of the quencher when G2-G3 is in the presence of excess free Ca\(^{2+}\), consistent with the above conclusion. Screening of crystallization conditions for G2-G3 on its own and in complexes with actin yielded only small crystals of G2-G3 that did not diffract X-rays sufficiently well to yield structural information.

Human CapG possesses three domains, C1-C3, that closely resemble the three N-terminal domains of gelsolin. We show CapG to exhibit both actin filament nucleation and severing activities in light scattering and total internal reflection fluorescence assays. Crystallization screens yielded crystals of CapG suitable for X-ray diffraction analysis, which resulted in a 1.5 Å resolution model for the C2-C3 fragment of CapG. The structure reveals a very different packing of C2-C3 in comparison to that observed previously in a mutant, CapG-sev, that had been engineered to resemble gelsolin G1 in portions of C1 and in the linking peptide between C1 and C2 [Zhang *et al.* (2006) *EMBO J* 25, 4458-4467]. We conclude that these structures represent analogues of the two different activated forms of the N-terminal half of gelsolin bound to actin [Wang (2008) ].
Finally, human villin-6M consists of villin domains V4-V6 without headpiece, which closely resemble gelsolin G4-G6. Screening of crystallization conditions produced crystals from which X-ray diffraction data generated a 1.9 Å resolution model for villin domain V6. Although grown by a different procedure, the crystals yielded a structure for V6 that is identical to that reported previously [Wang et al. (2009) J. Biol. Chem. 284, 21265-21269].
Preface

The structural studies described in Chapter 3 of this thesis constitute the basis for two manuscripts in preparation for publication.

One concerns the structure of CapG domains C2-C3. I was responsible for the expression, purification and crystallization of recombinant human CapG. I collected and processed the initial X-ray diffraction data at the UBC Centre for Blood Research X-ray crystallography facility, studied the effects of CapG in actin polymerization assays, wrote the first version of the manuscript and edited subsequent versions. Collaborators from the laboratory of Dr. R. Robinson of the Institute of Molecular and Cell Biology in Singapore collected higher resolution diffraction data from my crystals at a synchrotron facility in Taiwan, and performed the total internal reflectance microscopy assays to detect a novel actin filament-severing activity of CapG.

The second manuscript concerns the proposal of two distinct activated states for gelsolin family members. The major new structure described in this manuscript is of a novel form of gelsolin domains G1-G3 bound to actin and was the work of a former graduate student in the Burtnick laboratory, Dr. Hui Wang. My contribution to the manuscript was to illuminate the similarity between my structure for CapG domains C2-C3 and the originally reported activated conformation of gelsolin G2-G3, and then the similarity between G2-G3 in the novel G1-G3/actin structure and that observed for CapG domains C2-C3 in a mutant CapG that had been engineered to be an efficient severer of actin filaments.
# Table of Contents

Abstract .................................................................................................................................. ii

Preface ................................................................................................................................... iv

Table of Contents ................................................................................................................ v

List of Tables ....................................................................................................................... viii

List of Figures .................................................................................................................... ix

List of Abbreviations ......................................................................................................... xi

1. Introduction ..................................................................................................................... 1

1.1 Structure and properties of actin .............................................................................. 1
  1.1.1 ACTIN ............................................................................................................. 1
  1.1.2 Structure and properties of G-actin .............................................................. 2
  1.1.3 Structural models of F-actin ......................................................................... 4

1.2 Structure and properties of gelsolin ....................................................................... 7
  1.2.1 Background of gelsolin ................................................................................ 7
  1.2.2 Structure of gelsolin ..................................................................................... 8

1.3 Gelsolin fragment G2-G3 and its interaction with actin ....................................... 11
  1.3.1 Proposed models of gelsolin operation on F-actin ...................................... 11
  1.3.2 The G2-G3 fragment and its interaction with F-actin ................................ 12

1.4 The gelsolin superfamily of actin regulatory proteins ........................................... 14

1.5 Villin ..................................................................................................................... 18
  1.5.1 Structure and function ................................................................................ 18
  1.5.2 Actin and ligand-binding properties of villin ............................................. 20

1.6 CapG .................................................................................................................... 22
  1.6.1 Introduction ................................................................................................ 22
  1.6.2 Investigation of CapG structure and its comparison with gelsolin .......... 24
  1.6.3 Functional properties of CapG ................................................................... 27
1.7 Research goals .......................................................................................................................... 29

2. Materials and Methods ........................................................................................................... 31
  2.1 Plasmid construction .............................................................................................................. 31
  2.2 Protein preparation ............................................................................................................... 32
  2.3 Purification of actin ............................................................................................................... 34
    2.3.1 Preparation of muscle powder for actin prep ................................................................. 34
    2.3.2 Actin preparation and actin complexes purification ......................................................... 35
  2.4 Crystallization trials ............................................................................................................. 37
    2.4.1 Hanging drop vapor diffusion crystallization method ...................................................... 37
    2.4.2 Crystallization and data collection .................................................................................. 39
  2.5 Structure determination and refinement ............................................................................. 41
  2.6 Assays for Ca$^{2+}$-induced changes ................................................................................. 43
  2.7 Actin polymerization assay ................................................................................................. 45

3. Results ........................................................................................................................................ 46
  3.1 Protein preparation ............................................................................................................... 46
  3.2 Structure investigation by X-ray crystallography ............................................................... 48
    3.2.1 Novel structure of domains C2-C3 of CapG ................................................................. 49
    3.2.2 Structure of domain V6 of Villin-6M in the activated state ......................................... 52
  3.3 Investigation of G2-G3 by biochemical assays ................................................................. 54
    3.3.1 Fluorescence investigation of calcium-binding by isolated G2-G3 ............................ 54
    3.3.2 Actin-binding by G2-G3 ............................................................................................ 58

4. Discussion .................................................................................................................................... 65
  4.1 Comparison of CapG, mutant CapG, and gelsolin/actin complex structures .......... 65
  4.2 Common CapG/gelsolin features suggest common actin-related functions .......... 70
  4.3 Comparison of new V6 and previous V6 and G6 structures ........................................... 77
4.4 Fragment G2-G3 retains important properties of intact gelsolin ......................... 80

5. Conclusions and Future Work ....................................................................................... 83
   5.1 Conclusions .......................................................................................................... 83
   5.2 Future work .......................................................................................................... 86

References ......................................................................................................................... 89
List of Tables

Table 1. Data collection and refinement statistics for the domains C2 and C3 in crystals grown from native human CapG .......................................................... 50

Table 2. Data collection and refinement statistics for the villin V6 domain in crystals grown from recombinant native Villin-6M ............................................ 53

Table 3. Amino acid sequences of G2 and G3 of human gelsolin .................................. 55
List of Figures

Figure 1. Scheme for transition between G-actin and F-actin ......................... 2

Figure 2. Ribbon diagram of the polypeptide backbone structure of an actin monomer ................................................................. 4

Figure 3. Ribbon diagram of the backbone of intact inactive gelsolin ........... 9

Figure 4. Ribbon diagram representation of the structures of A: G1-G3/actin complex, and B: G4-G6/actin complex .................................................. 10

Figure 5. Two possible models for a gelsolin-capped actin filament ............ 12

Figure 6. Structures of G2-G3 extracted from inactive gelsolin and activated gelsolin G1-G3/actin .............................................................................. 13

Figure 7. Comparison of the structure of G1/actin with a model for G2/actin .... 15

Figure 8. Domain structures within the gelsolin superfamily. ..................... 16

Figure 9. Ribbon diagram representing the polypeptide backbone structure of chicken villin HP ................................................................. 20

Figure 10. Ribbon representation of the structure of the mutant human protein, CapG-sev, which possesses actin filament-severing activity .................. 26

Figure 11. Illustration of hanging drop crystallization method .................... 38

Figure 12. Schematic presentation of a two-dimensional phase diagram ....... 39

Figure 13. SDS-PAGE of target proteins ..................................................... 47

Figure 14. Ribbon representations of the backbone structures of domains C2-C3 ... 51

Figure 15. Schematic representation of the structure of villin domain V6 .......... 54

Figure 16. Fluorescence studies of tryptophan exposure to solvent in G2-G3 as a function of Ca$^{2+}$ concentration ......................................................... 57
Figure 17. SDS-PAGE for actin-Sepharose affinity chromatography of G2-G3 ..... 59

Figure 18. SDS-PAGE for cosedimentation assay of F-actin binding by G2-G3 ..... 60

Figure 19. Actin polymerization assays with CapG.................................................. 61

Figure 20. Effects of CapG on polymerized actin..................................................... 63

Figure 21. TIRF microscopy evidence for severing of F-actin by CapG.............. 64

Figure 22. Schematic representations of alternative activated forms for the N-terminal half of gelsolin bound to a single G-actin ........................................ 65

Figure 23. Superimposition of the α–carbon positions of the active states of G2-G3 and C2-C3 .......................................................................................................... 66

Figure 24. Conformational states of G1-G3.............................................................. 68

Figure 25. Conservation of the alternate domain 2/3 interfaces between gelsolin and CapG .................................................................................................................. 69

Figure 26. Active 1 and Active 2 binding sites on actin............................................. 72

Figure 27. Charge complementarity between G2-G3 and the actin filament........ 73

Figure 28. Exposure of PIP2-binding sites ............................................................... 75

Figure 29. Interchange between the two active conformations of G1-G3 .......... 76

Figure 30. Structural comparisons of domains V6 and G6 ..................................... 78
List of Abbreviations

2D – Two-dimensional

3D – Three-dimensional

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 CaCl₂, 0.2 mM ATP, 1 mM DTT, pH 7.6-7.8

C1 through C3 – Domains 1 through 3 of CapG

C2-C3 – A fragment of CapG consisting of domains C2 and C3

CapG – Actin-Capping proteins of the Gelsolin family

CapG-sev – A mutant CapG with severing function

CCP4 – Collaborative computational project number 4

DNase I – Deoxyribonuclease I

DTT – Dithiothreitol

EDTA – Ethylene diamine tetraacetic acid

EGTA – Ethyleneglycol-bis-(β-aminoethyl)-N, N, N’, N’-tetraacetic acid

F-actin – Filamentous actin

G-actin – Globular actin

G1 through G6 – Domains 1 through 6 of gelsolin
G1-G3 – The N-terminal half of gelsolin

G2-G3 – A fragment of gelsolin consisting of domains G2 and G3

G2-G4 – A fragment of gelsolin consisting of domains G2 through G4

G4-G6 – The C-terminal half of gelsolin

HEPES – N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid

HP – Villin headpiece domain

Hsp27 – Heat shock protein 27

IPTG – Isopropyl β-D-1-thiogalactopyranoside

kDa – kilo-dalton

LPA – Lysophosphatidic acid

NaAc – Sodium acetate

OD$_{600}$ – Apparent absorbance at 600 nm of a suspension of cells in a 1 cm path cuvet

PCR – Polymerase chain reaction

PDB – Protein data bank

PEG – Polyethyleneglycol

PIP2 – Phosphatidylinositol 4,5-bisphosphate

RMSD – Root-mean-square deviation

SDS-PAGE – Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

TIRF – Total internal reflectance fluorescence

Tris-HCl – Tris-(hydroxymethyl)-aminomethane hydrochloride

TMR – Tetramethyl-rhodamine-5-maleimide
Type I Ca$^{2+}$-binding site – Ca$^{2+}$-binding site in a gelsolin-type domain in which coordination of a bound Ca$^{2+}$ is shared between that domain and actin

Type II Ca$^{2+}$-binding site – Ca$^{2+}$-binding site in a gelsolin-type domain in which coordination of a bound Ca$^{2+}$ is wholly owned within that domain

Villin-6M – Domains 4 to 6 of villin without headpiece

V1 through V6 – Domains 1 through 6 of Villin

V6 – Villin domain 6
1. Introduction

1.1 Structure and properties of actin

1.1.1 Actin

Actin was discovered as a major protein component of rabbit skeletal muscle in 1942 [1] and since has been shown to be present in all eukaryotic cells, primarily in the cytoplasm, but also in the nucleus as reviewed in [2] and [3]. Based on the several reports of coexistence of multiple actin genes discovered from different species, it is commonly accepted that actin has different isoforms in higher eukaryotes, e.g., α-, β-, and γ-isoforms, which are expressed to different degrees in different tissues and cells as specifically mentioned in [4] and reviewed in [5]. Both in muscle and non-muscle cells, actin plays essential biological and physiological roles. It exists in two basic forms: a monomeric, globular form (G-actin) and a polymeric, filamentous form (F-actin) [6] that can be several micrometers in length as reviewed in [7]. F-actin is formed by polymerization of G-actin in a polarized fashion [6, 8]. The two forms coexist in a dynamic state under the control of multiple factors (Fig. 1), including intracellular ionic strength, pH and several actin-binding proteins (ABPs) as reviewed in [9]. ABPs can be subdivided into two antagonistic groups, those that favor the polymeric over the monomeric form and vice versa as reviewed in [9].
1.1.2 Structure and properties of G-actin

The first crystal structure of G-actin was obtained from co-crystals with DNase I [10]. Since then, several different non-polymerizable forms of G-actin have been crystallized as reviewed in [9]. These include numerous examples of actin co-crystallized with another protein (Fig. 2), as well as actin chemically modified to block polymerization, such as with tetramethyl-rhodamine-5-maleimide (TMR) [11]. Thus far, of the approximately 30 structures of G-actin deposited in the Protein Data Bank (PDB), only one exists of an uncomplexed, unmodified G-actin [12], determined from crystals of
G-actin grown in the presence of a chaperone protein, Hsp27. In spite of some minor conformational differences, all structures elucidated share several common features. G-actin is a 375-amino acid protein [13] consisting of four subdomains that surround a nucleotide cleft usually occupied by ATP rather than ADP [14]. ATP binds in the form of a complex with either Ca$^{2+}$ or Mg$^{2+}$ (the more dominant cytoplasmic cation in vivo) associating with the $\alpha$- and $\beta$-phosphates as mentioned in [15]. ATP-G-actin can be converted into the ADP form by hydrolysis as reviewed in [15] and [16]. There are a number of minor conformational differences in published G-actin structures that depend on the nucleotide type bound in the nucleotide-binding cleft [10, 11, 12, 17]. Although an early report [18] claimed that actin can be polymerized in a nucleotide- and divalent cation-free state in the presence of sucrose, under physiological conditions, G-actin loses its polymerizability and begins to denature in the absence of nucleotide and divalent cation [19]. ATP-G-actin adds faster and dissociates slower than ADP-G-actin at both the faster growing (“plus” or “barbed”) end and at the slower growing (“minus” or “pointed”) end of an F-actin filament [20].
Figure 2. Ribbon diagram of the polypeptide backbone structure of an actin monomer (G-Actin) (extracted from PDB ID 1RGI). This and subsequent representations of protein structures were produced using the PyMOL Molecular Graphics System (Schrödinger, LLC).

1.1.3 Structural models of F-actin

How monomers of actin join together to form the stable polymers crucial to muscle contraction, cellular motility and the cytoskeleton has been a long-standing question. Due to inherent disorder in aggregates of filaments, it has not been possible to obtain a crystal structure for F-actin. However, technical advances have permitted other methods, such as electron microscopy [21, 22] and fiber diffraction analysis [23, 24], to be used to
construct medium resolution (3.0 – 6.0 Å) models of F-actin. In the first such model, referred to as the Holmes or Heidelberg model [24], the structure of G-actin from the DNase I/G-actin complex was fit to low resolution F-actin fiber diffraction data. Actin subdomains 3 and 4 of each protomer are located close to the filament axis, while subdomains 1 and 2 are more distal and exposed to the surrounding solvent. The filaments are approximately 7 nm in diameter and can be described as two-start long pitch structures with a right-handed helical twist (Fig. 1). The two intertwined strands cross each other every half-pitch (36 nm), which involves 13 actin monomers. In each strand, the axial translation per monomer unit is ~5.5 nm, and the two helical strands are axially staggered by half the axial monomer translation [9, 24].

When decorated with myosin subfragment 1 (S1), F-actin displays an arrowhead-like appearance in electron micrographs [25]. F-actin is a polar filament with discernibly different ends. Each unit in F-actin is oriented with its cleft toward the same end of the filament as described in [26]. One end is called the barbed (or “plus”) end, while the other is named the pointed (or “minus”) end [20]. The barbed end of an F-actin filament is the preferred site for addition of ATP-G-actin [16]. The pointed-end is the slower growing end, and under certain conditions, may exhibit net loss of ADP-actin subunits [27, 28].

G-actin polymerization is essentially a condensation process that includes three steps as described in [16]. The first is formation of an unstable dimer, which is more likely to dissociate than to associate with a third actin monomer. The second is formation of a
trimeric species that can act as a nucleus for the third step, a rapid elongation phase due to addition of G-actin units. A plethora of interdomain conformational changes within G-actin occur during transformation to the so-called F-monomer conformation, \textit{i.e.}, the structure of an individual actin unit within an F-actin filament, which are key to the transition. Our understanding of the importance of each is not yet clear. Oda \textit{et al.} [29] recently proposed a model for F-actin based on X-ray fiber diffraction studies of highly oriented F-actin filaments in an intense magnetic field. The model attains a resolution of 3.3 Å in the radial direction and 5.6 Å along the filament axis and retains the basic features of the Holmes model. However, the resolution is sufficient to trace the path of the polypeptide backbone of each actin protomer. According to this model, the transformation from G- to F-actin requires an approximately 20° twist of the outer domain (subdomains 1 and 2) with respect to the inner domain (subdomains 3 and 4) about a rotation axis roughly at right angles to the F-actin helix axis and through the protein at the bottom of the nucleotide cleft within each actin unit. This rotation activates the ATP hydrolysis activity observed in F-actin by repositioning key residues of G-actin with respect to the \( \gamma \)-phosphate group of the bound ATP. Also, the conformational change from G-actin to F-actin alters the site on actin to which the muscle protein myosin binds during muscle contraction [30, 31]. Strikingly, the flattening of the F-actin monomer resulting from this rotation makes it more closely resemble the bacterial actin analog MreB [32], providing support for the proposal that actin is an important bridging molecule between prokaryotic and eukaryotic organisms in evolutionary history.
1.2 Structure and properties of gelsolin

1.2.1 Background of gelsolin

Gelsolin, an 82-84 kDa protein with six homologous subdomains, is the founding member of a superfamily of actin-binding proteins (ABPs) that regulate actin filament assembly and disassembly as reviewed in [33] and [34]. In a calcium-ion dependent manner, gelsolin can bind to the side of an F-actin filament, sever it in a nonhydrolytic manner, and cap the newly generated barbed end to block re-annealing or monomer addition [35]. The binding by gelsolin of specific phospholipids, such as phosphatidylinositol 4,5-bisphosphate (PIP2), induces “uncapping” to re-expose the fast-growing barbed end of the filament [36, 37].

There are two dominant isoforms of gelsolin in vivo [38]. Plasma gelsolin functions in extracellular fluid, where free calcium ion concentrations are relatively high (1-2 mM) [39], to scavenge actin that escapes from dead or dying cells [40]. Cytoplasmic gelsolin regulates actin dynamics within cells, where localized free calcium ion concentrations can vary from $10^{-7}$ to $10^{-5}$ M [41]. Plasma gelsolin consists of 755 amino acids and differs from cytoplasmic gelsolin only in possessing a 25-amino acid extension at its N-terminus [38, 42, 43]. In addition, plasma gelsolin possesses a disulfide bond connecting Cys188 and Cys201, which stabilizes the second domain of gelsolin in its activated conformation [44].
1.2.2 Structure of gelsolin

Gelsolin is composed of six similarly folded domains (G1-G6, named sequentially from the N-terminus), each of which contains 120-130 amino acid residues [38, 42, 43]. Inspection of the crystal structure [45] and analysis of the amino acid sequence [38] identify G1 and G4, G2 and G5, and G3 and G6 as most similar to each other, suggesting that gelsolin arose by triplication of a gene encoding an ancestral gelsolin-like domain, followed by a duplication of that gene to generate the current six-domain protein [38, 42, 43]. In the structure of Ca$^{2+}$-free, inactive gelsolin, the six domains fold into a globular conformation in which its filament side-binding site on domain G2 is blocked by a latch helix at the C-terminus (tail latch) [45]. Furthermore, actin monomer-binding sites on G1 and G4 are hidden within the compact globular arrangement of domains (Fig. 3). Upon binding Ca$^{2+}$, inactive gelsolin undergoes large-scale conformational changes to fully expose these sites [46, 47]. Although the structure of fully activated intact gelsolin is not yet available, the structures of the two isolated halves of activated gelsolin have been solved in the form of complexes with G-actin (Fig. 4) [46–48]. Structural analysis of inactive gelsolin and of its activated fragments suggests activation to be a multi-step process induced and regulated by calcium ions. It involves release of non-covalent attachments between and within the two halves of gelsolin, including opening of the tail latch and subsequent release of the G1-G3 and G4-G6 latches to expose the actin binding sites on G2, G1, and G4, respectively [48]. Exposure of the actin-binding site on G2
allows attachment of gelsolin to the side of a filament through domains G2 and G3 [47].
Once anchored, flexible polypeptide linkers from G2 to G1, and from G3 to G4, enable G1 and G4 to locate and bind to their respective recognition sites on F-actin to complete the severing function. The actin binding sites on the N- and C-terminal halves of gelsolin are proposed to bind to actin protomers across the filament from each other in the two-start helical representation of F-actin [47].

Figure 3. Ribbon diagram of the backbone of intact inactive gelsolin (PDB ID 1D0N). Color scheme used in this and subsequent gelsolin figures: G1, red; G2, light green; G3, yellow; G4, pink; G5, dark green; and G6, orange.)
Figure 4. Ribbon diagram representation of the structures of A: G1-G3/actin complex (PDB ID 1RGI), and B: G4-G6/actin complex (PDB ID 1H1V). In both, actin is colored cyan, ATP is shown in a ball and stick representation, with the associated calcium ion represented by a pink sphere to distinguish it from gelsolin-bound calcium ions, shown as gray spheres. Gelsolin domains G1 through G6 and actin subdomains 1 through 4 are indicated.

Gelsolin possesses two classes of calcium-binding site [46, 49, 50]. Type I sites are shared between actin and gelsolin domains, and are possessed by domains G1 and G4. They mediate the strength of interactions between gelsolin and actin through direct participation at binding interfaces; they also may play a role in release of the G1-G3 and G4-G6 latches during the activation of gelsolin [49–51]. Type II calcium-binding sites are completely contained within gelsolin. Occupation of these sites disrupts intra-protein interactions characteristic of inactive gelsolin in favor of new interactions in the coordination sphere of the calcium ion, as found in activated gelsolin [48, 49].
1.3 Gelsolin fragment G2-G3 and its interaction with actin

1.3.1 Proposed models of gelsolin operation on F-actin

G2-G3 is a 209 amino acid fragment of gelsolin with a molecular weight of ~23 kDa [38, 42]. As mentioned, following activation of gelsolin, this fragment binds and anchors gelsolin to the side of F-actin to initiate the severing and capping processes.

Superposition of the structures, respectively, of G1-G3/actin [47] and G4-G6/actin [48] onto two terminal actin units in the Holmes model for F-actin to cap it can be achieved in two ways [47]. Both place G2-G3 at the junction between two longitudinally neighboring actin protomers on one strand in the two-start helical model for F-actin, as previously predicted [51]. The long helix of G2 would be located close to a site on actin that is analogous to the previously characterized binding site for the corresponding helix of G1 [47, 50, 52]. The nearly 50 residue long G3-G4 linker polypeptide would permit G4 to bind to its G1-type binding site on one of two possible actin protomers across the F-actin helix from the anchoring position of G2-G3 (Fig. 5). While this model is attractive and remains viable in light of higher resolution F-actin models now available, further investigation of G2-G3 complexes with actin is required to verify the predictions for how G2 interacts with two actin protomers simultaneously.
Figure 5. Two possible models for a gelsolin-capped actin filament. A: Four ADP-actin protomers are drawn in shades of gray, with G1–G3/actin (PDB ID 1RGI) and G4–G6/actin (PDB ID 1H1V) overlaid onto the barbed end. Purple spheres mark Asp371 of G3 and Met412 of G4, a gap of 63.1 Å to be bridged by the G3–G4 linker, which is modeled in purple. B: Five ATP-actin protomers are drawn in cyan, blue, and gray, with G1–G3:actin (PDB ID 1RGI) and G4–G6:actin (PDB ID 1H1V) overlaid onto the barbed end. Purple spheres mark Asp371 of G3 and Met412 of G4, a gap of 101.3 Å. Gray spheres mark bound calcium ions.

1.3.2 The G2-G3 fragment and its interaction with F-actin

Biochemical studies originally suggested that G2 is the sole gelsolin domain that can bind to the side of an actin filament [43, 51, 53]. Subsequently, the structure of
G1-G3/actin showed that G2 and G3 share a common interface and that both contact actin, with G2 primarily at subdomain 2 and G3 extending down to subdomain 1 on the same actin protomer [47]. Upon activation of gelsolin, the opening of the G1-G3 latch releases G1 to translate away from G2 by extension of the G1-G2 linker to approximately 30 Å [47]. In contrast, the G2-G3 linker becomes shorter through adoption of a helical conformation, facilitating a tight interaction between these domains [47]. G2 and G3 cover 290 and 309 Å$^2$, respectively, of actin’s surface, and the contact area between G2 and G3 is 815 Å$^2$ (Fig. 6) [47]. Once G2-G3 is anchored to the side of a filament, the G1-G2 linker and the G3-G4 linker wrap themselves over the surface of the filament to direct G1 and G4 towards their appropriate binding positions.

Figure 6. Structures of G2-G3 extracted from inactive gelsolin (PDB ID 1D0N) and activated gelsolin G1-G3/actin (PDB ID 1RGI). The G1-G3 latch is closed in inactive gelsolin.
The position of G2 on actin in the G1-G3/actin complex differs from that suggested previously, which was predicted to resemble the one observed for actin with G1 (Fig. 7) [54, 55]. However, a G1/actin-like contact between G2 and actin remains a possible second binding mode between the proteins, supported by the models for gelsolin-capped F-actin (Figs. 5 and 7) [47]. Sequence alignment indicates that G2 may use Arg221 to form a salt bridge to Glu167 on actin, the same residue integral to the interactions of actin with G1 [47]. Overlaying of our G1–G3/actin structure onto the Holmes F-actin model reveals that the long helix of G2 lies close to the binding site on actin occupied by the analogous helix of G1, but on a different actin protomer. Furthermore, the Ile103 residue of G1, located in the center of the G1-actin interface, binds to the hydrophobic patch between actin subdomains 1 and 3. A similar residue, Leu211, on G2 can be placed within a triplet of residues (210-212) suggested to mediate the binding of gelsolin to F-actin [55, 56].

1.4 The gelsolin superfamily of actin regulatory proteins

As previously described, actin binds a substantial number of proteins collectively called Actin Binding Proteins (ABPs). More than 260 distinct and separate ABPs have been discovered as reviewed in [9], without counting their isoforms and synonyms. Many of the known ABPs bind to overlapping locations on the surface of actin,
introducing competition among their respective functions, should the ABPs co-exist in common subcellular compartments. Some ABPs bind with positive cooperativity, tending to form ternary complexes, whereas others bind with negative cooperativity. Moreover, some ABPs are capable of crosslinking actin filaments, while others enable filaments to interact with other cytoskeletal elements, such as microtubules or intermediate filaments as reviewed in [9] and [57].

![Image of G1 and G2 proteins with annotation](image)

**Figure 7.** Comparison of the structure of G1/actin (extracted from PDB ID 1RGI) with a model for G2/actin (extracted from the model in Fig. 5). Golden sphere represents a bound Ca$^{2+}$ ion.

Rapid assembly and disassembly of actin filaments is a characteristic of many actin-based processes and involves a wide spectrum of ABPs as reviewed in [57]. Given the range of ABP structures and functions, attempts at classification are bound to be somewhat arbitrary. There are always some ABPs that do not fit into any particular group.
and the boundaries between groups tend to be smeared. However, considering the roles of ABPs that facilitate either assembly or disassembly of actin filaments, some consensus classifications emerge. Among these, the gelsolin superfamily of eukaryotic proteins control actin organization by severing filaments, capping filament ends and nucleating actin assembly [9, 56]. The family includes, but is not limited to, gelsolin, villin, adseverin, severin, CapG, advillin, supervillin and flightless I (Fig. 8) [57], each containing three or more 120-amino acid structural repeats of a recognizable gelsolin domain.

**Figure 8.** Domain structures within the gelsolin superfamily. The conserved gelsolin-like domains that characterize the family are numbered 1-6. The headpiece (HP) domain also is conserved in proteins of the villin subfamily within this grouping (based on [58]).
Most gelsolin superfamily members have either three or six such domains, while some contain additional domains not related in any way to gelsolin. For instance, CapG (originally identified as a macrophage actin capping protein that lacked severing activity) has three gelsolin-like domains, C1-C3, that are analogous to gelsolin domains G1-G3 [59]. On the other hand, villin (95 kDa and named due to its localization in microvilli) [60] contains six domains, V1-V6 [61], analogous to gelsolin domains G1-G6, and regulates actin dynamics [62] through a Ca\(^{2+}\)-dependent N-terminal half [63]. However villin contains an additional actin-binding headpiece (HP) domain [64] that confers on it an ability to bundle actin filaments. Many of the functions of the family members overlap, although their sensitivity to Ca\(^{2+}\) and other signals may vary. The fusion of non-gelsolin domains into some of these proteins extends their ranges of function beyond those normally associated with gelsolin in the remodeling of actin filaments, giving them unique roles in cell motility, control of apoptosis, regulation of phagocytosis, and even in the regulation of gene expression [65]. Taken together, these diverse activities make the gelsolin superfamily one of the most versatile sets of regulators of actin functions.
1.5 Villin

1.5.1 Structure and function

Villin is a 95 kDa member of the gelsolin superfamily [60], having six homologous copies of a gelsolin-type domain [61, 64] (~120 amino acids, and ~14 kDa each), denoted V1-V6 (Fig. 8). Due to the limited amount of structural data for villin, much has been inferred about villin structure and function based on its close relation to gelsolin. It is hypothesized that villin and gelsolin have similar tertiary structures by the fact that villin has ~45% amino acid sequence identity with gelsolin over the six domains [58]. Support for this proposal comes from NMR structural studies on individual domains of villin in solution, which show them to fold in a manner typical of gelsolin domains [65]. In contrast, the HP domain, a 76 amino acid “headpiece” at villin’s C-terminus, has yielded an X-ray crystallographic structure that is unrelated to any gelsolin domain, while being connected to V6 by a 40-residue disordered linker [66]. This enables it to bind to actin independently of the rest of villin, and explains the ability of villin to cross-link F-actin filaments, a property not shared by gelsolin [58, 66, 67].

As an unusually versatile actin-modifying protein, villin can regulate both assembly and disassembly of actin filaments through its nucleating, crosslinking (bundling), severing (in response to high Ca$^{2+}$ levels or following tyrosine phosphorylation), and barbed-end capping activities [68]. Villin heads a subfamily of proteins within the gelsolin superfamily of proteins, those that include a C-terminal HP domain (Fig. 8). As
the HP domain includes a distinct actin-binding surface, the members of this subfamily are capable of crosslinking actin filaments [66].

Atomic resolution structural data for intact villin are not available, but it is generally accepted that villin, like gelsolin, undergoes a global calcium-induced conformational change from a compact globular structure to become more asymmetric [63]. Based on amino acid sequence homology with gelsolin, two potential latch sequences (one between V4 and V5, and the other between V6 and HP) have been identified that might be able to form salt bridges with residues in the V2 domain in inactive villin to stabilize its compact geometry as reviewed in [58]. The binding of Ca\(^{2+}\) would release the latches and, specific to villin, induce an outward shift of the HP domain into solution. While direct visualization of such a compact folded state for intact inactive villin is still not possible, solution NMR structures of isolated domain V1 and of a fragment consisting of the V6-plus-HP domains [65], together with the recent crystal structures of both a wild-type and a mutant villin HP domain (Fig. 9) [66], permit models to be postulated. The structural data confirm the overall V1 and V6 folds to be very similar to those observed in the corresponding gelsolin domains G1 and G6 [65, 66], and suggest the 40 amino acid linker between V6 and HP to be unfolded and disordered [66].
Figure 9. Ribbon diagram representing the polypeptide backbone structure of chicken villin HP (PDB ID 2RJY).

A 3D model for the structure of villin bound to F-actin has been proposed on the basis of electron tomography of 2D arrays of F-actin/villin complexes in which villin crosslinks F-actin filaments at low calcium concentrations [69]. The 2D array map is well fit by an extended villin molecule that is similar in structure to that proposed for calcium-activated gelsolin in small-angle X-ray scattering studies [70].

1.5.2 Actin- and ligand-binding properties of villin

Discrimination among villin’s disparate actin-modifying functions is achieved by sensitivity to the binding of Ca$^{2+}$ and polyphosphoinosides, as for gelsolin, but also by a post-translational modification, tyrosine phosphorylation [71, 72].

Actin-binding by villin is completely dependent on Ca$^{2+}$. Investigations in vitro indicate that villin in a 1 µM free Ca$^{2+}$ environment can cap actin filaments at their
barbed ends with high affinity (binding constant: $10^{11}$ M$^{-1}$) through interactions that involve two actin monomer-binding sites located, respectively, on V1 and V4 [73]. These capped filaments can serve as nuclei for pointed-end filament elongation. Severing of actin filaments by villin requires significantly higher calcium concentrations (100-200 μM) [73]. Different calcium-binding sites control capping and cutting and the sites that regulate severing have a much lower affinity for Ca$^{2+}$ than the ones that regulate barbed-end filament-capping as mentioned in [72, 73]. Normally, 100-200 μM Ca$^{2+}$ concentrations are not found in live cells, but they might occur in apoptotic cells as described in [74].

Tyrosine phosphorylation in villin can decrease its affinity for F-actin and activate severing at sub-micromolar free Ca$^{2+}$ concentrations, suggesting that phosphorylation might play an important role in regulating villin conformational changes [75]. Other F-actin binding proteins, e.g., tropomyosin, also can decrease villin’s binding to F-actin [76].

Two F-actin binding sites have been identified in villin, one in V1-V2 and the other in HP. The site in the villin core contains a common F-actin-binding KKEK motif [77], whereas binding to HP involves a “hydrophobic cap” mechanism with different residues of HP contributing to actin-binding [78].

Direct association of villin with PIP2 inhibits both the actin filament-capping [72, 79] and severing functions of villin [79, 80]. In contrast, the actin filament-bundling function of villin is enhanced by interaction with PIP2, while it is inhibited by tyrosine
phosphorylation [72, 79]. The only actin-modifying function of villin that is not regulated by PIP2 is nucleation of pointed-end filament elongation. In spite of the significant effects of binding of PIP2 to villin on villin’s activities, the interaction does not result in gross changes to villin’s conformation. There are three proposed PIP2 binding sites on villin, two in the core and one in HP [79, 81]. All three sites are conserved among related actin-binding proteins, including gelsolin, adseverin, advillin, and CapG. Recent studies demonstrated that villin can interact with relatively hydrophilic phospholipids, such as lysophosphatidic acid (LPA), with a higher affinity than with PIP2. PIP2 and LPA compete for the same binding sites in villin [82]. Interaction of villin with LPA inhibits all actin-modifying functions of villin. On the other hand, in vivo, villin sequestration of LPA could result in modulation not only of actin filament dynamics, but also of LPA-induced intracellular signaling [82].

1.6 CapG

1.6.1 Introduction

With other gelsolin superfamily members, CapG, a 38 kDa protein, shares the properties of binding to the barbed end of actin filaments with high affinity and of binding to actin monomers and nucleating actin assembly; however, it is the only known member of the gelsolin family reported to lack the ability to sever actin filaments as
reviewed in [83]. CapG is the now agreed upon name for what different investigators previously named macrophage-capping protein [83], gCap39 [84], and Mbh1 [85]. “Cap” signifies the barbed-end actin filament-capping activity of the protein, while “G” designates it as a member of the gelsolin superfamily, sharing 49% amino acid sequence identity with the N-terminal half of gelsolin [83-85]. CapG is found at levels comparable to gelsolin in many cell types. The concentration of CapG differs from gelsolin in platelets where CapG is not detected, and in macrophages and dendritic cells where CapG is more abundant than gelsolin, representing 1% of the total cytoplasmic protein [86]. Nucleic acid sequencing of CapG cDNA [59] indicates that CapG is homologous to the N-terminus of gelsolin and is organized into three structural domains of equal molecular weight (C1-C3, from N- to C-terminal) (Fig. 8).

CapG is functionally similar to CapZ [87, 88], a heterodimeric protein that blocks actin filament barbed ends in mammalian cells [89]. CapZ has two subunits: α, 36 kDa, and β, 32 kDa [90]. CapZ interactions with actin are unaffected by changes in calcium ion concentration, whereas CapG requires mM levels of Ca\(^{2+}\) in order to bind actin [84]. CapG is reported to bind a single Ca\(^{2+}\) at a site localized between domains C2 and C3, with an affinity that is lower than reported for gelsolin or villin. The binding of Ca\(^{2+}\) results in a more open conformation of CapG [91]. CapG caps the barbed end of filaments at much lower calcium concentrations than required for interaction with G-actin, suggesting that under the calcium conditions found inside cells CapG functions primarily as a capping rather than a G-actin binding protein [84, 92].
1.6.2 Investigation of CapG structure and its comparison with gelsolin

Like severin and fragmin (Fig. 8) [58], CapG consists of three gelsolin-type structural domains, in contrast to the six found in gelsolin, adseverin, and villin. The fact that the shorter members of the gelsolin family retain both calcium-sensitivity and gelsolin-like activities invites comparisons between them and the separated halves of gelsolin. The severing efficiency of the isolated N-terminal half of gelsolin (G1–G3) is concentration dependent, whereas that of intact gelsolin is not, which suggests that cooperativity between the N- and C-terminal halves of gelsolin is required for full efficiency as reviewed in [93]. Members of another well-studied family of actin-binding proteins, ADF-cofilin, which share negligible amino acid sequence homology with gelsolin but comprise one or two domains each folded into a structure that closely resembles a gelsolin domain, also bind F-actin cooperatively and require multiple binding events for severing to occur as reviewed in [94]. These findings suggest that the three-domain gelsolin family members should retain a concentration-dependent severing activity and, indeed, such reports exist for severin and fragmin [59, 92]. However CapG has been reported to be unable to sever actin filaments at concentrations normally associated with such an activity [83].

CapG’s reported inability to sever filaments has invited rigorous comparison of its primary structure with those of gelsolin-like severing proteins. The region in which CapG most differs is located between the first and second domains, which in gelsolin
contain the consensus sequence FKHVXPN, known as the WH2 motif, found in a number of actin regulatory proteins [95]. In all severing members of the gelsolin family, a polar charged amino acid (His in gelsolin) is followed by a hydrophobic residue (Val in gelsolin) in the WH2 sequence as reviewed in [95]. However, in CapG one amino acid insertion and three substitutions are found in the consensus sequence [96]. Although the functional importance of this conserved sequence is not fully understood, and its role in actin filament severing has not been defined, it has been suggested as a potential focus for investigation by mutagenesis as mentioned in [96]. Simultaneously changing amino acids 84–91 (associated with severing by gelsolin) and 124–137 (WH2-containing segment) in CapG to the corresponding sequences found in gelsolin produces a mutant, CapG-sev, that is capable of efficiently severing and capping actin filaments. CapG-sev, compared to CapG, possesses increased actin monomer-binding affinity and an additional actin monomer-binding site [91, 96].

Crystals of CapG-sev reveal a domain-swapped dimer in which there exists a tight interaction between domains C2 and C3, and the C1-C2 linker is extended to position domain C1 away from the C2-C3 fragment (Fig. 10) [91]. Alterations in the sequence of the WH2 segment and insertion of a single amino acid are postulated to explain the lack of severing activity reported for native CapG.
Figure 10. Ribbon representation of the structure of the mutant human protein, CapG-sev, which possesses actin filament-severing activity (extracted from the domain-swapped dimeric structure deposited as PDB ID 1J72).

This arrangement of domains C1-C3 in CapG-sev is significantly different from what is observed in G1-G3 structures reported for either inactive Ca\textsuperscript{2+}-free gelsolin [45] or for its activated Ca\textsuperscript{2+}-bound form bound to G-actin [47]. The CapG C1 calcium-binding site is more exposed in CapG-sev than the comparable one in gelsolin G1, leading the authors of the study to speculate that in CapG-sev, the calcium-binding event that initiates activation involves domain C1, whereas in gelsolin it involves G3 [91, 96]. Hence the different packing arrangements, and a possible explanation for earlier reports that CapG requires lower Ca\textsuperscript{2+} concentrations than gelsolin to initiate activation [84]. In addition, the shorter linker between domains 1 and 2 in CapG-sev relative to gelsolin results in models where it is tightly stretched across the longitudinal axis of one
actin monomer in CapG-decorated F-actin filaments. This may have a bearing on earlier observations that CapG is able to bind and/or dissociate rapidly from F-actin in response to calcium concentration changes in vivo [83].

1.6.3 Functional properties of CapG

Although both CapG and gelsolin are known regulators of actin filament dynamics, gene knockout experiments to produce mice that are either CapG-null or gelsolin-null result in mice that survive to maturity and display a generally normal phenotype with apparently normal reproductive capabilities as reviewed in [97]. Given the importance of severing and capping in actin dynamics, together with the existence of several different proteins that exhibit one or both of these activities to a significant extent, it is not surprising that gelsolin-null or CapG-null mice are able to survive [97]. Overlapping distributions and functions of different ABPs provide for back-up in cases where the normal regulatory protein is absent as reviewed in [9]. Indeed, the response of gelsolin-null mice to a variety of stresses is dramatically less efficient than in gelsolin-positive mice. For example, gelsolin-null mice exhibit prolonged bleeding times due to defects in gelsolin-dependent shape changes of platelets in response to activating conditions [98].

CapG is conserved throughout all vertebrate species. In most cell types (platelets are an exception), CapG is expressed at moderately high levels whereas in macrophages, it is
an abundant protein, representing 1% of total cytoplasmic protein [59, 86]. Macrophages are one of the most dynamic of living cells. Their dorsal surfaces rapidly change in shape, creating outward protrusions of the peripheral membrane (ruffles) as reviewed in [99]. Ruffling is believed to be the result of localized actin filament assembly [97] and the extent of ruffling correlates with oscillations in the concentration of intracellular calcium ions, suggesting that ruffling is dependent on the activities of calcium-sensitive proteins. CapG binds PIP2 with high affinity [84] and cells in which CapG has been up-regulated exhibit increased receptor-mediated phosphoinositide turnover and calcium signalling, whereas loss of CapG correlates with opposite effects. CapG gene knockout experiments show that it plays a central role in regulating the cell ruffling response [97]. This may be because CapG is able to rapidly and reversibly cap the barbed-end of actin filaments in response to fluctuations in Ca$^{2+}$ concentration, while capping by gelsolin is not easily reversed simply by a reduction in Ca$^{2+}$ concentration. As ruffling is associated with rapid shifts in actin filament length, this distinctive ability of CapG makes it a likely candidate to moderate the phenomenon as reviewed in [92, 97] and described in [91, 96].

The actin-binding activities of both gelsolin and CapG are regulated by µM Ca$^{2+}$. But, while the activation of CapG is reversed by lowering the Ca$^{2+}$ concentration to submicromolar levels [84], gelsolin does not release actin in submicromolar Ca$^{2+}$; instead, phosphoinositide binding is required for release [100]. In addition, the affinity of CapG for actin is considerably lower than that of gelsolin (K$_d$ in the nM to µM range, as compared to pM for gelsolin) [91]. In addition to participation in ruffling, CapG in
macrophages also is involved in phagocytosis, a major function of these cells, as well as in vesicle “rocketing” in their cytoplasm [97]. When macrophages are treated with lanthanum followed by zinc to antagonize secretary processes [98], a small percentage of the resulting vesicles begin to move through the cytoplasm, being propelled by actin filament tails [101, 102]. Even though the mechanism behind this activity is unclear, this movement requires assembly of actin filaments at their barbed ends, and CapG facilitates the process.

To understand how CapG affects these multiple actin-based motile functions, in vitro data show that CapG can be inhibited from capping the barbed ends of actin filaments either by lowering Ca$^{2+}$ concentrations to the nanomolar range [92] or by increasing PIP2 concentrations to the submicromolar range [84]. Therefore, receptor-induced cyclic changes in Ca$^{2+}$ and/or PIP2 local concentrations in vivo would be expected to modulate CapG activity.

X-ray crystallographic structures are available for a mutant CapG (CapG-sev) (Fig. 10) [91], for intact inactive gelsolin (Fig. 3) [45] and for the separated halves of activated gelsolin, each bound to an actin monomer (Fig. 4) [47, 48]. Absent from this set is the structure of native CapG and that of any complex between it and actin.

1.7 Research goals

Based on our previous successes in crystallization and structural elucidation of intact
inactive gelsolin and of the complexes of the two halves of activated gelsolin with actin, and on subsequent failures to crystallize full-length activated gelsolin with or without actin, we set out to crystallize the fragment of gelsolin, G2-G3, that initiates contact with F-actin, both on its own and in complexes with actin. The goals were to determine atomic resolution details of the gelsolin activation process and of the interactions involved between gelsolin and F-actin.

Because of the uncertainties inherent to protein crystallography, additional crystallization projects with overlapping goals were undertaken in parallel, using proteins known to have similar functions and expected to have similar structures to gelsolin. Specifically, I expressed and purified CapG and villin and subjected them to analogous crystallization trials in resting and activating conditions, in both the presence and absence of actin.

With each gelsolin fragment or family member expressed, purified and subjected to crystallization trials, I conducted biochemical assays to test their functions in order to complement the crystallographic results. In particular, as evidence mounted for existence of more than one activated conformation for gelsolin family proteins, we decided to reinvestigate claims of the inability of CapG to sever actin filaments.
2. Materials and Methods

2.1 Plasmid construction

Plasmids incorporating either Villin-6M or CapG were jointly prepared and generously supplied by Anantasak Loonchanta, a student in the Burtnick laboratory at UBC, and Sakesit Chumnarnsilpa, a graduate student in the laboratory of our collaborator, Dr. Robert Robinson (Institute of Molecular and Cell Biology, Singapore).

Based on the known human villin gene sequence (Genbank: P09327), specific primers containing SfiI and EcoRI restriction recognition sites for domains V4 to V6, with a medium sized linker polypeptide preceding domain V4, were designed to encompass from Lys360 in the middle of the villin V3-V4 linker to Ser720 on the C-terminal side of villin V6 [121]. Amplification by polymerase chain reaction (PCR) and cloning into pSY5, a modified pET-21d(+) plasmid expression vector (Novagen) designed to encode an N-terminal 8-histidine tag followed by a Prescission protease cleavage site, were performed as previously reported [121].

The DNA sequence encoding human CapG (Genbank: NM001747) was similarly used to design primers that would encompass the mature protein sequence, amplified by PCR and cloned into pSY5. The recombinant villin-6M (native villin of three gelsolin-like domains without headpiece) and CapG plasmids were initially propagated in *E. coli* NovaBlue cells (Novagen) and prepared using plasmid preparation kit (QIAGEN) following the manual provided by the manufacturer.
E. coli Rosetta cells, transformed with pGEX vector (containing GST-His*(8) tag) into which a G2-G3 insert (residue Val147 to Lys305) was cloned, were kindly provided by Dr. Robert Robinson. The suspension of cells was prepared by mixing the Luria-Bertani (LB) media culture containing ampicillin (100 µg/mL) and chloramphenicol (34 µg/mL) with 50% glycerol (v/v). These were stored as smaller aliquots (1 mL) at -80 °C.

2.2 Protein preparation

LB medium (50 mL) containing antibiotics was inoculated with a frozen stock culture (100 µL) of the target protein and grown overnight at 37 °C with vigorous shaking. Then, 1L LB medium was inoculated with this non-induced culture so that the OD$_{600}$ was approximately 0.02-0.03. It then was incubated in the shaker at 37 °C until the OD$_{600}$ value reached 0.6-1.0, which usually required 3-5 hours. Over-expression was induced with 0.8 mM (final concentration) isopropyl-β-D-thiogalactopyranoside (IPTG). After an induction period of over 20 hours at 30 °C, the expressed cells were harvested by centrifugation at 5,000 g for 30 min using a Sorvall RC-5B centrifuge and GS-3 rotor, washed with 1/10 bed volume of 50 mM Tris buffer at pH 8.0, and re-suspended in 20 mL of 10 mM imidazole in buffer 1 (10 mM Tris-HCl, 300 mM NaCl, pH 8.0). A cell-disrupter (high pressure homogenizer) (Avestin, Emulsiflex) was used to break the cells. Ruptured cells were subsequently ultracentrifuged at 35,000 rpm for 40-60 min, at
4 °C (Beckman Optima L-90K, 45-Ti rotor) to pellet the cellular debris. Supernatant was collected and proteins with His-tags were shown to be present in this soluble fraction by SDS-PAGE [103]. Taking advantage of the His(*8)-tag, nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography (QIAGEN™) was utilized to purify the recombinant target protein. The centrifugal supernatant was loaded onto a Ni-NTA agarose column that had been pre-equilibrated with buffer 1, at a reduced flow rate (~0.5 mL/min) to allow for efficient binding of His-tagged proteins. The column was washed with the equilibration buffer, followed by 80 mM imidazole in buffer 1 to remove non-specifically bound *E. coli* proteins. Elution was carried out with 300 mM imidazole in buffer 1, followed by 500 mM imidazole in the same buffer to ensure that elution was complete. The elute was collected, buffer-exchanged and concentrated by ultrafiltration to achieve final sample volumes of less than 20 mL. Finally, Villin-6M and CapG were purified by gel filtration (Bio-Rad Sephacryl S300; 90 × 2.5 cm) in 10 mM Tris-HCl, 150 mM NaCl, 5 mM CaCl2, pH 7.5.

To achieve high purity G2-G3, an additional affinity chromatography step (actin-Sepharose column) was required. According to instruction of the manufacturer (Bio-Rad™), the actin-Sepharose column (Bio-Rad™) was pre-equilibrated with buffer 2 (10 mM Tris-HCl, 100 mM NaCl, 0.2 mM ATP, 0.1 mM CaCl2, 1 mM NaN3, pH 7.0), followed by loading the sample onto the column. Then, excess sample and other non-specifically bound impurities were washed off using 300 mM NaCl in buffer 2. Finally, the target protein was eluted with 1-2 M NaCl in buffer 2. To obtain the
His$_8$-tagged G2-G3 fragment, the GST tag was cleaved by thrombin, which was eliminated by utilizing benzamidine-Sepharose beads (Bio-Rad$^{TM}$) (batch procedure), while the GST tag was removed by chromatography on the Ni-NTA agarose again. Final purification of the His$_8$-tagged G2-G3 fragment was by gel filtration following the procedure described above.

Because of the small size of the His-tags and the increased solubility they could provide to the proteins containing them, they were not removed prior to subsequent use in the experiments described in this thesis.

Protein concentrations in solutions of Villin-6M, CapG and G2-G3 were measured spectrophotometrically (Perkin Elmer Lambda 4B UV-Vis Spectrometer) at a wavelength of 280 nm using absorption coefficients of 1.25 mL mg$^{-1}$ cm$^{-1}$, 1.3 mL mg$^{-1}$ cm$^{-1}$, and 1.1 mL mg$^{-1}$ cm$^{-1}$, respectively (calculation based on the amino acid sequence and composition according to ExPASy Proteomic Server). Protein purity was evaluated by SDS-PAGE.

2.3 Purification of actin

2.3.1 Preparation of muscle powder for actin prep

Following our standard laboratory procedure [12], one kilogram of rabbit back and leg muscles (Pel-Freez Biologicals, Rogers, AR) was minced in a blender at 4 $^\circ$C with 3 L of cold 0.3 M KCl, 0.1 M KH$_2$PO$_4$, 0.05 M K$_2$HPO$_4$, adjusted to pH 6.5, with addition
of ATP to 0.2 mM just before use. The debris was centrifuged at 7000 rpm (Beckman Avanti J-20 XPI centrifuge and JLA rotor) for one hour. The supernatant (rich in myosin) was discarded. The sediment was resuspended and washed with 3 to 4 L of cold water for 20 min, and centrifuged again for 30 min. The sediment next was washed with 3 to 4 L of cold 0.01 M NaHCO₃, pH 8.0, for 20 min to remove soluble contaminants such as myoglobin. After an additional centrifugation, the sediment was washed with 3 to 4 L of cold water for 10 to 20 min and centrifuged again. Next, the sediment was twice washed with 3 to 4 L of cold 95% ethanol and centrifuged for 10 min. Finally, the sediment was resuspended and twice washed with 3 to 4 L of acetone and filtered through a double layer of cheesecloth. The solid residue obtained was spread out on a sheet of filter paper in a fume hood at room temperature until completely dry. The residue was packaged in 5 g lots stored at -20 °C.

2.3.2 Actin preparation and actin complexes purification

Actin was purified from rabbit skeletal muscle [12, 104]. 5 g rabbit skeletal muscle powder was extracted with 60 mL of buffer A (2 mM Tris-HCl, 0.2 mM CaCl₂, 0.2 mM ATP, 1 mM DTT, at pH 7.6-7.8) for 45 min on ice. The extract then was filtered and centrifuged at 35,000 rpm (Beckman Optima L-90K, 45-Ti rotor) for one hour at 4 °C. The resulting supernatant was polymerized overnight by addition of KCl (to 50 mM) and MgCl₂ (to 2 mM). The following morning, additional KCl was added to 0.8 M in
order to release proteins such as troponin and tropomyosin from filamentous actin. After stirring for one hour, the solution was centrifuged again at 35,000 rpm, but for 3 hours. The resulting pellet was resuspended in 10 mL buffer A and dialyzed against three changes of 1 L of buffer A prior to centrifugation again at 35,000 rpm for one hour. The resulting supernatant, a solution of essentially pure G-actin, was applied to a Sephacryl (S-300) gel filtration column and eluted with buffer A while monitoring with an ECONO UV monitor (Bio Rad) at 280 nm. The concentration of G-actin was evaluated by UV-Vis spectrophotometry (Perkin Elmer Lamda 4B) at 290 nm with an absorption coefficient of 0.63 mL mg⁻¹ cm⁻¹ [12].

Villin-6M, CapG and G2-G3 were separately incubated in 10 mM Tris-HCl, 150 mM NaCl and 5 mM CaCl₂, pH 7.5 for 30 min to activate the proteins. Then the protein solutions were added to 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 × 2.5 cm). The fractions containing the protein complex were pooled and concentrated to 10 mg/mL, as determined by UV-Vis absorbance at 280 nm (Perkin Elmer Lamda 4B) using calculated absorption coefficients of 1.4, 1.26, and 1.18 mL mg⁻¹ cm⁻¹, respectively, for actin bound to Villin-6M, CapG and G2-G3 (calculation based on the amino acid sequence and composition according to ExPASy Proteomic Server).
2.4 Crystallization trials

2.4.1 Hanging drop vapor diffusion crystallization method

Protein crystallization is not yet an exact science. Successful crystallization usually begins with a target protein of high purity, with regard both to other protein contaminants and to its folded conformation. To induce crystal growth, it is common to screen dozens, even hundreds, of conditions in which individual parameters, especially solution composition, but also temperature and pH, are varied. Initial successes usually produce small crystals of moderate to low diffraction quality. Such conditions must be optimized to obtain the high quality of crystals required for X-ray diffraction studies. The resulting crystals are orderly three-dimensional arrays of molecules, held together by non-covalent interactions. Crystals of proteins are usually grown by slow, controlled precipitation from aqueous solution under conditions that do not denature the protein. The water-soluble polymer polyethylene glycol (PEG) is widely utilized to induce protein crystallization as it is a powerful precipitant and weak denaturant. To achieve crystallization, the precipitant is added to the aqueous solution in the well of a crystallization tray. In the hanging drop vapor diffusion method, a drop of protein solution is suspended above the precipitant solution (Fig. 11). Water is allowed to evaporate slowly from the protein solution, raising the concentration of protein until precipitation occurs. Whether the protein forms crystals or instead forms a useless amorphous solid depends on many factors, such as temperature, pH, ionic strength, and the presence of various ligands, as
well as on protein concentration and purity. It is usually necessary to screen a variety of protein crystallization conditions over a large range of individual parameters in order to yield crystals. Once potential crystallization conditions have been identified, it generally is necessary to optimize those conditions by finer screening of nearby conditions to obtain the best possible quality crystals for X-ray diffraction analysis [105].

As presented in **Fig. 11**, 1-2 µL of high purity protein solution is mixed with an equal volume of reservoir solution and then suspended as a droplet underneath a glass cover slip, which is sealed onto the top of a well containing 1 mL of reservoir buffer. Vapor diffusion results in net transfer of water from the protein solution to the reservoir, thus increasing the concentration of protein and precipitant in the droplet, which leads to supersaturation of the protein component (**Fig. 12**).

**Figure 11.** Illustration of hanging drop crystallization method (based on Hampton Research catalogue)
Figure 12. Schematic presentation of a two-dimensional phase diagram, plotting the protein concentration against precipitant concentration. The supersaturation area comprises the metastable and nucleation zones (based on [106]).

2.4.2 Crystallization and data collection

Crystals of villin-6M, CapG, and G2-G3 (without actin) were grown at 4°C using the hanging drop vapor diffusion method. The protein solutions were mixed in a 1:1 ratio (v/v) with the reservoir solutions for various crystal screening conditions. Diffraction quality crystals grew using the following precipitant solutions:

1. Activated Villin-6M, 18% (w/v) PEG 2000MME (monomethyl ether), 0.1 M Tris-HCl buffer, 5% (v/v) ethanol, pH 8.0
2. Activated native CapG, 20% (w/v) PEG 10,000, 0.1 M HEPES buffer, pH 7.5
3. G2-G3, 5% (w/v) PEG 8000, 0.1 M Na acetate buffer, pH 6.0, or, alternatively, 20% (w/v) PEG 8000, 0.1 M Tris-HCl buffer, pH 7.5

Prior to X-ray data collection, the crystals were transferred into appropriate
cryoprotectant solutions and flash frozen in liquid nitrogen in order to prevent ice crystals from forming during the process of freezing. The cryoprotectant solutions for freezing protein crystals were as follows:

1. Activated Villin-6M in 15% (v/v) glycerol, 18% (w/v) PEG 2000MME, 0.1 M Tris-HCl buffer, 5% (v/v) ethanol, pH 8.0;

2. Activated native CapG in 25% (v/v) glycerol, 0.1 M HEPES buffer, pH 7.5, 20% (w/v) PEG 10,000;

3. G2-G3 in 20% (v/v) glycerol, 0.1 M Na acetate buffer, 5% PEG (w/v) 8000, pH 6.0; alternatively in 15% (v/v) glycerol, 0.1 M Tris-HCl buffer, 20% PEG 8000, pH 7.5.

The frozen crystals were mounted and maintained at 100 K under a stream of nitrogen to prevent possible radiation damage to the protein crystals during data collection. Diffraction data were recorded with a CCD detector placed behind the crystal. The target crystals were incrementally rotated through an appropriate angle, determined by crystal symmetry, with a diffraction image recorded at each step in order to obtain a complete data set. To avoid generating a “blind spot” in reciprocal space close to the rotation axis, the rotation axis was reset at least once during the data collection process.

Preliminary data sets were collected using a Rigaku RU200 rotating anode source with OSMIC mirrors and a MAR345 image plate detector at the UBC Centre for Blood Research X-ray data collection facility. Higher resolution data were collected on beamline BL13B on an ADSC Quantum 315 CCD detector at the National Synchrotron Radiation Research Center (NSRRC, Taiwan). All collected data were indexed,
integrated and scaled using HKL2000 software [107]. Indexing includes identifying unit cell type and dimensions, and determining the position in reciprocal space of each image peak. Next, the data set of images, each of which may contain thousands of diffraction spots, were integrated into a single file that recorded the Miller indices and intensities of the diffraction spots. Finally, the images in a data set were merged to identify the peaks appearing in more than one image and scaled to ensure that all images had a consistent intensity scale.

2.5 Structure determination and refinement

Since the data collected from an X-ray diffraction experiment give a reciprocal space presentation of the protein crystal lattice, the relative position of each diffraction spot on the detector plate is determined by the size and shape of the unit cell and the extent of symmetry within the crystal. The intensities of reflections correlate with the unit cell contents whereas the dimensions of the reciprocal lattice correlate with the reflection positions. The structure factor \( F(h,k,l) \), a complex number, contains information about both the amplitude and phase of a diffracted wave:

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

where \( f(a) \) is the scattering factor of atom \( a \) and depends on the kind of atom and the
diffraction angle of the corresponding reflection. The intensity is proportional to the square of the structure factor amplitude, \([F(h,k,l)]^2\). The electron density can be calculated based on the following equation after the Fourier transformation of structure factors, given that both amplitude and phase are known.

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

The structure factor amplitudes can be obtained directly from the intensities of the reflections, whereas the phases cannot be easily determined. There are three major indirect techniques to determine the phases, including Molecular Replacement (MR), Multiple Isomorphous Replacement (MIR) and Multi-wavelength/Single-wavelength Anomalous Diffraction (MAD/SAD). MR utilizes previously solved structures of molecules sharing high sequence homology with the target molecule as search models to determine the position and orientation of the target molecule in the unit cell. Generally, there are two steps involved in the MR process, rotation and translation. In the rotation search, the orientation of the search model is determined that produces maximal overlap with the target structure by comparing the Patterson maps for the target and a particular search model in different orientations. In the translation search, the search model with optimized orientation is translated to the best-fit coordinates within the asymmetric unit by moving the model, calculating and obtaining a new Patterson map, and comparing it to the target-derived Patterson map.

All structures presented here in the thesis were solved by MR using single domains of CapG-sev (extracted from PDB ID 1J72) or activated gelsolin (extracted from PDB ID
1RGI or 1H1V) as search models for CapG and Villin, respectively, using MolrRep [108] and Phaser [109], available through the CCP4 software package [110]. Once initial phases are established, a model can be constructed by fitting the known protein amino acid sequence into the initial electron density map by utilizing the interactive computer graphics program, COOT [111]. The obtained model is then subjected to repeated cycles of refinement until the observed structure factors ($F_{\text{obs}}$) agree with the calculated structure factors ($F_{\text{calc}}$) from the model. Evaluation of fit is quantified by the R-factor, expressed as follows:

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

Structural refinement was conducted by rigid body refinement, followed by restrained refinement, with REFMAC 5 software [112]. A refined model with an R-factor in the range of 15% to 20% generally is considered as an acceptable model of the actual protein structure.

### 2.6 Assays for Ca$^{2+}$-induced changes

Ca$^{2+}$-induced conformational changes in G2-G3 were monitored by observation of changes in the intrinsic fluorescence of tryptophan residues in the protein as the free Ca$^{2+}$ concentration was increased [113]. Fluorescence at 340 nm was measured with a Perkin-Elmer LS-55 luminescence spectrometer using an excitation wavelength of 280
nm. G2-G3 was stored initially in a buffer solution including 0.1 M Tris-HCl, 150 mM NaCl, 1 mM EGTA, pH 8.0. The Ca$^{2+}$ titration gradients was determined as described previously [122] using WEB-MAXC (http://maxchelator.stanford.edu/) to calculate the free calcium levels. The free Ca$^{2+}$ ion was increased in steps up to 10 mM by adding 100 mM CaCl$_2$ stock solution.

Studies of solvent-accessibility of the tryptophan residues of G2-G3 in the presence and absence of free Ca$^{2+}$ were performed using a fluorescence quenching method [113]. The results of fluorescence quenching by iodide ions (Fig. 6B) are presented in the form of a plot based on the Stern-Volmer equation for collisional quenching:

$$\frac{F_0}{F} = 1 + K_{sv}[\text{Quencher}] = 1 + k_q\tau_o[\text{Quencher}]$$

where $F_0$ and $F$, respectively, are tryptophan fluorescence intensities (excitation at 290 nm, emission at 340 nm) measured in the absence and presence of iodide ions at specified concentrations. The Stern-Volmer quenching constant $K_{sv}$ is the product of the bimolecular collisional rate constant, $k_q$, for the process and the fluorescence lifetime of the fluorophore in the absence of quencher, $\tau_o$.

To investigate the calcium-dependence of interactions between actin and G2-G3, a solution of purified G2-G3 (10 mg in 40 mL buffer including 10 mM Tris-HCl, 300 mM NaCl, pH 8.0) was divided into two and loaded in separate experiments onto the actin-Sepharose affinity chromatography column described previously. One sample was pre-incubated with 1 mM CaCl$_2$, and the other with 1 mM EDTA for over 2 hours at 4 °C. Both samples were treated as described previously with regard to actin-Sepharose
affinity chromatography, and each fraction was monitored by absorbance at 290 nm during the process. Co-sedimentation experiments [36] were carried out to further investigate the F-actin binding characteristics of G2-G3. Freshly prepared G-actin and G2-G3 in buffer A were mixed together at a 1:1 molar ratio and then subjected to polymerizing conditions at 22 °C by adding KCl to 50 mM and MgCl₂ to 2 mM. After the polymerization process, the mixtures were centrifuged at 35,000 rpm (Beckman model Optima L-90K ultracentrifuge and Beckman 45 Ti rotor). The supernatant was collected and the pellets re-suspended by washing with buffer A. Both were subjected to analysis by SDS-PAGE.

2.7 Actin polymerization assay

G-actin in buffer A was polymerized by addition of KCl and MgCl₂ to 100 mM and 2 mM, respectively, at room temperature. Polymerization was monitored by right angle light scattering in a LS55 luminescence spectrometer with both emission and excitation wavelengths set at 350 nm [114]. To investigate possible G-actin-sequestering and F-actin-severing functions of CapG, polymerized F-actin (25 min after achieving maximal light scattering) was mixed with different concentrations of CapG to yield from 1:1 to 4:1 molar ratios to actin monomer. Effects of these additions on light scattering intensities were recorded.
3. Results

3.1 Protein preparation

To investigate the structures of target proteins, the open reading frame (a DNA sequence that does not contain a stop codon in the specified reading frame) of G2-G3 (23 kDa), native CapG (46 kDa), and Villin-6M (52 kDa) were each cloned into an *E. coli* expression vector pET-21d(+), which introduces a (His)$_8$-tag at the N-terminal of the expressed protein. Each of the three proteins was heterologously overproduced as a His-tagged fusion protein, which enabled us to use a single purification step of Ni$^{2+}$-chelation chromatography to obtain the purified protein.

A representative SDS-PAGE result (Fig. 13A) demonstrates that the expression levels and solubility of the recombinant G2-G3 were acceptable for further purification. Under the employed conditions, 40% of total amount of overproduced G2-G3 in *E. coli* cells (estimated from densitometry of SDS-PAGE gel) was recovered in soluble form with a yield of 20 mg/L of culture after purification. His$_8$-tagged G2-G3, after enzymatic cleavage and further purification, was obtained with a yield of 8-10 mg/L of culture. For Villin-6M and CapG (Fig. 13B), over 95% of the total amount of overexpressed proteins (including soluble and insoluble fractions) in *E. coli* cells were recovered in soluble fractions. The yields of Villin-6M and CapG are 20 mg/L and 30 mg/L of LB culture, respectively.
Figure 13. SDS-PAGE of target proteins. **A:** lane 1 (MK): protein markers, only applies to lane 2 to 4 from the same SDS-PAGE gel; lane 2: induced *E. coli* cells; lane 3: insoluble fraction; lane 4: soluble fraction. Lane 5-6 & 7 are obtained from two different gels, respectively; lane 5: purified GST-His8-tagged G2-G3; lane 6: purified His8-tagged G2-G3; lane 7: G2-G3/actin complex; **B:** lane 1: purified His8-tagged villin-6M; lane 2: insoluble fraction; lane 3: soluble fraction; lane 4: purified His8-tagged CapG; lane 5: insoluble fraction; lane 6: soluble fraction (staining: Coomassie Brilliant Blue R-250) [115, 116]
3.2 Structure investigation by X-ray crystallography

Several hundred crystallization conditions were screened, including our own PEG and ammonium sulfate screens [117] and commercial screens such as the Hampton\textsuperscript{TM} and Qiagen\textsuperscript{TM} crystallography screen kits for macromolecular complexes. None of the conditions tested led to crystals of a Villin-6M/actin complex. Furthermore, crystals of a G2-G3 complex with actin, grown against either 0.1 M HEPES buffer, 10\% (w/v) PEG 6000, 2 mM CaCl\textsubscript{2}, pH 6.5, or, alternatively, 0.1 M NaAc buffer, 5\% (w/v) PEG 20,000, 0.1 M KSCN, pH 5.0, failed to diffract at all, even following attempts at optimization.

Success in growing diffracting crystals was achieved in the following cases:

1. Activated Villin-6M alone, grown against 18\% (w/v) PEG 2000MME, 0.1 M Tris-HCl, 5\% ethanol (v/v), pH 8.0;
2. Activated CapG alone, grown against 20\% (w/v) PEG 10,000, 0.1 M HEPES buffer, pH 7.5, or, alternatively, 18\% (w/v) PEG 8000, 0.1 M NaAc buffer, pH 5.0;
3. G2-G3//actin complex, 12\% PEG 8000 (v/v), 0.1 M NaAc buffer, pH 5.5, or, alternatively, 20\% (w/v) PEG 8000, 0.1 M Tris-HCl buffer, pH 7.5.

Unfortunately, the small crystals of G2-G3 obtained diffracted poorly (resolution poorer than 4 Å) and subsequent attempts to enlarge the crystals and to optimize the quality of diffraction by varying the PEG used, pH, and various salts and other additives or by varying the data collection strategies [118] did not succeed.
3.2.1 Novel structure of domains C2-C3 of CapG

Crystals of CapG alone, supplemented with 5 mM CaCl$_2$ to ensure an activated state, grew using a precipitant solution of 20% (w/v) PEG 10,000, 0.1 M HEPES, pH 7.5 and resulted in a refined structural model in which only domains C2 and C3 were required to account for all of the density in the calculated electron density map (Fig. 14A, Table 1). Domain C1 is not evident in this structure, possibly due to it being disordered in the crystals or due to proteolytic cleavage of the domain 1-2 linker during the period required for crystal growth. The two domains pack together into a compact globular unit that is notably different from that observed in the structure reported for the mutant, CapG-sev [91] (Figure 14B). In native CapG, domains C2 and C3 are joined by an $\alpha$-helical linker peptide and share an inter-domain interface of 728 Å$^2$ (Fig. 14A). Each of domains C2 and C3 exhibits the characteristic gelsolin-type domain fold, consisting of a central five-stranded $\beta$-sheet sandwiched between a long $\alpha$-helix almost parallel to the strands and a short helix perpendicular to the strands. As expected for activated CapG, each of C2 and C3 has a Ca$^{2+}$ bound into its type II metal ion-binding site (Fig. 14A). That Eu$^{3+}$ can be soaked into crystals of CapG-sev suggests that its conformation also corresponds to an activated one (Fig. 14B) [91].
**Table 1.** Data collection and refinement statistics for the domains C2 and C3 in crystals grown from native human CapG.

<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>I4</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a = b = 113.0 Å, c = 47.2 Å, α = β = γ = 90°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>30.0-1.53 (1.58-1.53)&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>45130 (4507)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.1 (5.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.4 (99.8)</td>
</tr>
<tr>
<td>Average I/I&lt;sub&gt;o&lt;/sub&gt;</td>
<td>45.9 (4.93)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.8 (34.6)</td>
</tr>
<tr>
<td>R&lt;sub&gt;factor&lt;/sub&gt; (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>17.4 (19.9)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19.7 (22.8)</td>
</tr>
<tr>
<td>Molecules in asymmetric unit</td>
<td>1</td>
</tr>
<tr>
<td>Residue range</td>
<td>CapG 135-348&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters; Ca&lt;sup&gt;2+&lt;/sup&gt;)</td>
<td>2071 (381; 2)</td>
</tr>
<tr>
<td>Mean derived B-factor (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>25.6</td>
</tr>
<tr>
<td>R.M.S. deviation bonds (Å)</td>
<td>0.006</td>
</tr>
<tr>
<td>R.M.S. deviation angles (deg)</td>
<td>1.01</td>
</tr>
</tbody>
</table>

<sup>a</sup> R<sub>merge</sub> (|I−<I>|/<I>)  
<sup>b</sup> R<sub>factor</sub> (||F<sub>O</sub>−|F<sub>C</sub>||/|F<sub>O</sub>|)  
<sup>c</sup> Based on 5% of the data  
<sup>d</sup> Values in parentheses represent the highest resolution shell  
<sup>e</sup> Break in electron density at Lys204
A disulfide bond, corresponding to the Cys188-Cys201 disulfide bond of G2-G3 [119], exists between Cys165 and Cys178. The conservation of two cysteine residues in domains 2 of gelsolin and CapG that are capable of forming a disulfide bond indicates a role in oxidizing environments, where the formation of the domain 2 disulfide bond will increase stability of the released proteins, and in the case of gelsolin, enhance its activity [119]. Hence, it is possible that cytoplasmic gelsolin and CapG released into the extracellular space upon cell rupture will augment the plasma gelsolin actin-scavenging system.

**Figure 14.** Ribbon representations of the backbone structures of domains C2-C3 of A: activated native CapG with two bound Ca\(^{2+}\), and B: the mutant, CapG-sev with a bound Eu\(^{3+}\) (extracted from PDB ID 1JHW).
3.2.2 Structure of domain V6 of Villin-6M in the activated state

Villin is a member of the gelsolin superfamily that exhibits a bundling function in addition to the usual gelsolin-type characteristics [120]. Diffraction from crystals of Villin-6M grown in 18% (w/v) PEG 2000MME, 0.1 M Tris-HCl, 5% ethanol (v/v), pH 8.0 resulted in a refined structural model in which only domain V6 was required to account for all of the density in the calculated electron density map (Table 2). The unit cell parameters and refinement statistics are very similar to those reported previously by our group for villin domain V6 [121]. The crystallographic asymmetric unit contains two V6 domains, which permits calculation of a solvent content of 60%. Consistent with the previous crystallization results, only domain V6 contributes to diffraction, even though the intact C-terminal villin fragment Villin-6M is present in the protein droplet before crystallization, as confirmed by the SDS-PAGE. It is possible that proteolysis occurred during the incubation period required for crystal nucleation and growth. Alternatively, the V4-V5 portion of Villin-6M, connected to V6 by a flexible linker peptide, may not be constrained to defined positions in the crystal lattice and could have considerable freedom of movement in the solvent portion of the crystals. (Insufficient protein was retrieved from these crystals to check for proteolysis by SDS-PAGE or mass spectrometry.)
Table 2. Data collection and refinement statistics for the villin V6 domain in crystals grown from recombinant native Villin-6M

<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>P3_1</td>
</tr>
<tr>
<td>Unit Cell</td>
<td>a = b = 49.88 Å, c = 98.86 Å, α = β = 90°, γ = 120°</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>25.0-1.86 (1.96-1.86)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>22860 (2276)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.0 (4.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.3 (98.9)</td>
</tr>
<tr>
<td>Average I/I_0</td>
<td>15.5(8.87)</td>
</tr>
<tr>
<td>R_merge (%)^a</td>
<td>5.8 (28.6)</td>
</tr>
<tr>
<td>R_factor (%)^b</td>
<td>18.8 (21.8)</td>
</tr>
<tr>
<td>R_free (%)^c</td>
<td>22.7 (25.8)</td>
</tr>
<tr>
<td>Molecules in asymmetric unit</td>
<td>2</td>
</tr>
<tr>
<td>Villin V6 residue range</td>
<td>635-740</td>
</tr>
<tr>
<td>Non-hydrogen atoms (waters)</td>
<td>1968 (180)</td>
</tr>
<tr>
<td>Mean derived B-factor (Å^2)</td>
<td>41.8 (48.6)</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>1.510</td>
</tr>
</tbody>
</table>

^a R_merge (|I−<I>|/<I>)  ^b R_factor (||F_o||−||F_c||/<F_o>)  ^c Based on 5% of the data

^d Values in parentheses represent the highest resolution shell
Like domain G6 of gelsolin, V6 of Villin-6M possesses a five-stranded \( \beta \)-sheet at its core, sandwiched between two \( \alpha \)-helices, a long helix aligned roughly parallel to the strands of the sheet, and a short one roughly perpendicular to the strands (Fig. 15A). V6 displays appropriate residues positioned to form a type II calcium-binding site (Fig. 15B), but this site is vacant here as it was in the previous one [121]. Comparing with usual kinked long helix in G6 of gelsolin, straight long helix here probably means \( \text{Ca}^{2+} \)-binding is not necessary for “straighten-process”. In addition, \( \text{Ca}^{2+} \) might not be able to break the relative stable interactions within this potential binding site.

**Figure 15.** Schematic representation of the structure of villin domain V6. **A:** Ribbon diagram with N- and C-termini indicated. No bound \( \text{Ca}^{2+} \) was detected. **B:** Key residues that could constitute a type II calcium-binding site.

### 3.3 Investigation of G2-G3 by biochemical assays

#### 3.3.1 Fluorescence investigation of calcium-binding by isolated G2-G3

Structural data suggest that a type II calcium-binding site exists within each of the
six gelsolin domains [45–48]. The affinities of these sites are known to vary significantly and some of the sites act cooperatively [122, 123]. The metamorphosis of gelsolin to an activated structure through the binding of Ca$^{2+}$ ions entails the opening of at least three latches so as to expose actin-binding surfaces on different domains of gelsolin [45, 46]. Conformational changes induced in isolated G2-G3 by the binding of calcium ions were monitored by observing changes in the intrinsic fluorescence of tryptophan residues within G2-G3. Based on the amino acid sequences of G2 and G3 (GenBank ID: AAH26033), each contains two tryptophan residues (Table 3, G2: Trp158 & 178; G3: Trp296 & 347), both of which are located in β-sheets at the cores of the respective domains.

**Table 3.** Amino acid sequences of G2 and G3 of human gelsolin

<table>
<thead>
<tr>
<th>Amino acid sequence of G2 (129-248)</th>
<th>H VVPNEVVVQR LFQVKGRVV RATEVPV$^\text{WE}$ SFNNGDCFIL DLGNIHQ$^\text{WC}$ GSNSNRYERL KATQVSKGIR DNERSGRARV HVSEEGTEPE AMLQVLGPKP ALPAGEDTA KEDAANRKL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino acid sequence of G3 (249-377)</td>
<td>A KLYKVSNGAG TMSVSLVADE NPFAQALKS EDCFILDHGK DGKIFV$^\text{WK}$ KGQANTEERKAA LKTASDFITK MDYPKQTQVS VLPEGGETPL FKQFFKN$^\text{WR}$ PDQTDGLGLS YLSSHIANVE RVPFDAAT</td>
</tr>
</tbody>
</table>
In response to a calcium ion concentration increase up to 10 mM, tryptophan fluorescence from G2-G3 in 0.1 M Tris-HCl buffer, 300 mM NaCl, pH 8.0, displayed a biphasic increase (Fig. 16A), suggesting successive conformational changes that each reduce the solvent exposure of the tryptophans, likely the result of folding the tryptophans more deeply into the protein core. The midpoints of the two transitions indicate $K_d$ values for the relevant binding sites to be approximately 30 nM for the higher affinity site and 6 µM for the lower affinity one. As type I metal ion-binding sites on gelsolin require the participation of actin, which is absent from these experiments, the two binding events are consistent with the proposed existence of a type II binding site in each of the six gelsolin domains [45, 46], here one in each of G2 and G3.

The accessibility of tryptophan residues in Ca$^{2+}$-free and Ca$^{2+}$-bound forms of G2-G3 was monitored by quenching experiments using potassium iodide (Fig. 16B). Two experiments were conducted in 0.1 M Tris-HCl buffer, 300 mM NaCl, pH 8.0, with 1 mM ethyleneglycol-bis-(β-aminoethylether)-N,N,N′,N′-tetraacetic acid (EGTA) in the Ca$^{2+}$-free sample and 1 mM CaCl$_2$ in the Ca$^{2+}$-bound one. The lower slope of the quenching plot in the presence of Ca$^{2+}$ than when calcium ions are chelated with EGTA is indicative of less efficient quenching of tryptophan residues, likely the result of their being more deeply buried in the protein’s core. Figure 16 displays the replicate results after optimization and due to only qualitative purpose, mean data and standard deviation calculation are absent.
Figure 16. Fluorescence studies of tryptophan exposure to solvent in G2-G3 as a function of free Ca\(^{2+}\) concentration. **A:** Ca\(^{2+}\) titration of Trp fluorescence of G2-G3 (5 µM) (pCa = -log[Ca\(^{2+}\)]\(_{\text{free}}\); **B:** Stern-Volmer plot of quenching of Trp fluorescence of G2-G3 by iodide ions (5 µM G2-G3 in 0.1 M Tris-HCl buffer at pH=8.0, [KI] + [KCl] = 50 µM, 1 µM CaCl\(_2\) and no further components (triangles), or with added EGTA (2 µM, squares). The K\(_SV\) value (12.9 ± 1.0 M\(^{-1}\)) in the presence of Ca\(^{2+}\) is less than that in its absence (K\(_SV\) = 21.9 ± 1.7 M\(^{-1}\)).
3.3.2 Actin-binding by G2-G3

Two samples of G2-G3 in 0.1 M Tris-HCl buffer with 300 mM NaCl at pH 8.0 were treated with 1 mM CaCl₂ and 2 mM EGTA, respectively, and separately subjected to actin-Sepharose affinity chromatography. Both G2-G3 samples were retained partially by non-covalent interaction with actin covalently bound to the column matrix, although the EGTA-containing sample was less extensively retained (the estimated protein weight ratio, “Elute”/”Sample”, of Fig. 17A is ~45% whereas the “Elute”/”Sample” ratio in Fig. 17B is ~75%) as determined by densitometry of SDS-PAGE gels.

Cosedimentation experiments indicate that G2-G3 binds to F-actin (Fig. 18). Since G2-G3 and G-actin alone are both in the supernatant before and after the polymerization process, whereas the polymerized F-actin forms the pellet after high-speed centrifugation, the appearance of G2-G3 in the pellet with actin can be only explained as the result of binding interactions between F-actin and G2-G3. Figure 17 and 18 display the replicate results after optimization and due to only qualitative purpose, mean data and standard deviation calculation are absent.
**Figure 17.** SDS-PAGE results for actin-Sepharose affinity chromatography of G2-G3. 

**A.** (with EGTA) G2-G3 in EGTA (0.1 M Tris-HCl buffer with 300 mM NaCl at pH 8.0 and 2 mM EGTA). 

**B.** (with Ca\(^{2+}\)) G2-G3 in Ca\(^{2+}\) (0.1 M Tris-HCl buffer with 300 mM NaCl at pH 8.0 and 1 mM CaCl\(_2\) ).

**Sample:** Purified G2-G3; **Load:** collected solution after sample going-through column (same volume solution collected as G2-G3 solution); **Wash:** collected solution from washing the column with 0.1M Tris-HCl buffer with 300 mM NaCl at pH 8.0 (same volume of washing buffer used as G2-G3 solution, washed columns until no further protein eluted, here only the first time going-through column wash buffer collected); Elute: collected solution from eluting the column with 0.1 M Tris-HCl buffer with 2 M NaCl at pH 8.0 after finishing washing; **MK:** Markers (three bands from top to bottom represent 26, 17, 10 kDa markers); **Arrow:** G2-G3 band. Two lower bands are evident in the SDS-PAGE with Ca\(^{2+}\) while absent in the one with EGTA, possibly due to the presence of proteases with Ca\(^{2+}\)-dependent activation in the buffer.
Figure 18. SDS-PAGE for cosedimentation assay of F-actin binding by G2-G3. Samples containing an equimolar mixture of G2-G3 and G-actin, each at 5 µM, were polymerized and centrifuged. **A (Controls):** MK, markers (from top to bottom, the four bands represent 50, 34, 25, and 17 kDa marker proteins); **Lane 1,** G-actin solution in buffer A before polymerization; **Lane 2,** supernatant of centrifuged G-actin solution in buffer A without adding KCl/MgCl2; no pellet detected; **Lane 3,** G2-G3 in 0.1 M Tris-HCl buffer, 300 mM NaCl at pH 8.0; **Lane 4,** supernatant of G2-G3 solution in the same buffer after centrifugation; no pellet detected; **Lane 5,** G-actin mixed with G2-G3, each at 5 µM, in buffer A, before polymerization; **Lane 6,** supernatant of G-actin mixed with G2-G3 solution after centrifugation, without adding salts to induce polymerization; no pellet detected. **B:** MK, markers (from top to bottom, four bands represent 50, 34, 25, and 17 kDa marker proteins); **Lane 1,** sample prior to polymerization; **Lane 2,** supernatant after polymerization and centrifugation; **Lane 3,** resuspended pellet after polymerization and centrifugation.
3.4 Actin Polymerization in the presence of CapG

Polymerization of G-actin to F-actin may be followed by the increase in light scattering associated with filament development. Curves showing the time course of the increase in scattering intensity are sigmoidal, indicative of a nucleation phase (lag phase), followed by rapid elongation, and culminating in a steady state value once polymerization has been maximized (Fig. 19, curve 2). When CapG is present together with G-actin in the starting samples, the lag phase, corresponding to the nucleation of filament growth, can be eliminated (Fig. 19, curves 3-5). Figure 19 & 20 displays the replicate results after optimization and due to qualitative purpose only, mean data and standard deviation calculation are absent.

Figure 19. Actin polymerization assays with CapG (initial G-actin concentration for all measurements was 5 µM). 1: G-actin control (no additions), 2-5: Actin polymerization in the presence of different concentrations of CapG. 2: 0 µM; 3: 5 µM; 4: 10 µM; 5: 20 µM. Actin polymerization was induced by KCl and MgCl₂ as previously described.
To test whether CapG is able to sever F-actin, the effect of CapG on steady-state samples of F-actin was examined. CapG was added to samples of F-actin and the light scattering intensity at 350 nm was monitored with time (Fig. 20). CapG decreases the degree of light scattering, indicating a reduction in the amount of F-actin present. However, it is not clear from this result whether the reduction is a result of severing in the middle of filaments, which has not previously been reported for CapG, or of sequestration of actin monomers at the ends of the filaments, leading to a net depolymerization of the filaments from the ends inward. Our collaborators in the Robinson laboratory in Singapore are able to answer this question by recording video data from total internal reflectance fluorescence (TIRF) microscopy of fluorescently labeled F-actin filaments (Fig. 21). As a solution of CapG is washed over the filaments, they quickly break up in a uniform manner, not progressively inward from the filament ends. Thus, for the first time, we have evidence that CapG can indeed sever F-actin. Severing is much less efficient that for gelsolin and other identified severers, in that equimolar ratios of CapG to actin are required, whereas severing is evident at a mole ratio of gelsolin to actin of less than 1:20. On the other hand, reduced gelsolin is reported to sever less effectively than oxidized gelsolin as described in [119]. It is very likely that reduced CapG severing activity may have been underestimated on account of the presence of DTT in the buffer creating the reducing environment of CapG, which might explain the absence of CapG severing activity before.
Figure 20. Effects of CapG on polymerized actin. 1: Control F-actin (5 µM) without CapG; 2: actin:CapG mole ratio of 1:1; 3: actin:CapG mole ratio of 1:2; 4: actin:CapG mole ratio of 1:4. Arrow: the time at which CapG was added into the F-actin sample.
Figure 21. TIRF microscopy evidence for severing of F-actin by CapG. a) Severing by CapG. b) Control wash with polymerization buffer in the presence of 1 mM Ca\textsuperscript{2+}. c) Control wash with a known actin-sequesterer, DNaseI. d) Control wash with a known actin filament-capping protein, CapZ. None of the control washes, b) through d), led to measurable severing. (Micrographs courtesy of R. Robinson.)
4. Discussion

4.1 Comparison of CapG, mutant CapG, and gelsolin/actin complex structures

Fragment C2-C3 from recombinant human CapG (Fig. 14A) adopts an Active 1 conformation, as defined for G2-G3 within the first reported structure of the activated G1-G3/actin complex (Fig. 22A) [46]. The root-mean-square deviation (RMSD) of α–carbon positions calculated over 206 homologous positions is 1.56 Å when the Active 1 conformation of G2-G3 (extracted from PDB ID 3FFK) is superimposed upon the novel C2-C3 structure (Fig. 23A).

Figure 22. Schematic representations of alternative activated forms for the N-terminal half of gelsolin bound to a single G-actin. A: Active 1. B: Active 2. In both, the two halves of actin are shown in a space-filling representation, with subdomains 1 and 2 in pink and subdomains 3 and 4 in steel blue. ATP is indicated in the nucleotide cleft. Calcium ions are shown as black or gray spheres. The C-terminal α-carbon on G3 is marked as a cyan sphere.
Figure 23. Superimposition of the α–carbon positions of the active states of G2-G3 and C2-C3. **A:** G2-G3 (colored: G2 green, G3 yellow) from the Active 1 structure for G1-G3/actin (PDB ID 3FFK) superimposed onto the novel C2-C3 structure (colored: C2 pink, C3 cyan). **B:** G2-G3 (colored: G2 green, G3 yellow) from the novel Active 2 structure for G1-G3/actin [13] superimposed onto C2-C3 (colored: C2 pink, C3 cyan) from CapG-sev (PDB ID 1J72). The jagged line represents the furin cleavage site in G2. The spheres denote calcium ions, grey (bound to gelsolin) and purple (bound to CapG).

In contrast, the relative orientation of C2 and C3 observed in crystals of the CapG-sev mutant protein (Fig. 14B) [91] is distinctly different, adopting an Active 2 conformation instead, as defined by G2-G3 within a recently discovered alternative form of an activated G1-G3/actin complex (Fig. 22B) [124]. The RMSD over 196 α–carbon positions when the Active 2 form of G2-G3 is overlaid upon C2-C3 from CapG-sev is 2.68 Å (Fig. 23B).
Crystals of the mutant protein, CapG-sev, grown from 3.0 - 4.0 M ammonium formate, revealed all three CapG domains, C1 through C3, in the form of a domain-swapped dimer [91] (one CapG molecule within such a dimer is shown in Fig. 24D). Domains C2 and C3 from CapG-sev closely resemble the activated versions, rather than the inactive forms, of the equivalent gelsolin domains (Fig. 24), further supported by C3 reportedly accepting Eu$^{3+}$ into its type II metal ion-binding site without major structural change [91] (Fig. 14B). The interface between C2 and C3 covers 407 Å$^2$ on the surface of each domain, while C1 is translated 36 Å away from C2-C3 by extension of the C1-C2 linker. This is reminiscent of the observed extension of the G1-G2 linker across the face of actin for both active conformations of G1-G3, which permits separation of the binding sites on actin for G1 and G2-G3 (Fig. 22; Fig. 24).

The two alternative structures for activated G1-G3/actin confirm that gelsolin undergoes large conformational changes in the course of its function (Fig. 22). Common to achieving both activated states is that the G1/G3 latch in inactive gelsolin (Fig. 24A) must be sprung to expose the actin-binding site on G1 (Fig. 22). As a consequence, the kinked long helix observed in calcium-free G3 becomes straight and the G1-G2 linker extends to separate G1 from G2 allowing G2 and G3 to form new interfaces, Active 1 (Fig. 24B) and Active 2 (Fig. 24C). Interconversion between the two active structures would require dissociation of these G2-G3 interfaces, rotation of G3 relative to G2 and contraction or lengthening of the G2-G3 linker (transition between Fig. 24B and Fig. 24C; Fig 25).
Figure 24. Conformational states of G1-G3. A: Inactive G1-G3 taken from the Ca-free structure of whole gelsolin [PDB ID 3FFN]. The G1/G3 latch refers to the common β-sheet between G1 and G3. B: G1-G3 from the previously known G1-G3/actin structure [PDB ID 3FFK]. C: G1-G3 extracted from the novel G1-G3/actin structure [124]. The jagged lines in B and C represent the furin cleavage sites in G2. Arrows indicate domain movements that are required to transition between these conformations. D: CapG-sev [PDB ID 1J72] in the absence of actin. CapG-sev domains are colored orange (C1), pink (C2), and cyan (C3).

One of the most noticeable distinctions between the novel native CapG structure (Active 1) and that of CapG-sev (Active 2) is that the linker peptide between C2 and C3 adopts an α-helical structure in the former whereas it forms an extended loop in the latter (Fig. 14). This mirrors the states of the G2-G3 linker in the Active 1 and Active 2 conformations, respectively, of gelsolin (Fig. 22; Fig. 24).

The residues that form the interfaces between the second and third domains of CapG and gelsolin are homologous (Fig. 25), indicating that Active 1 and Active 2 forms have been maintained throughout the evolutionary history of both proteins, possibly pointing to a common functional role for each conformation. Similarly, conservation of the
actin-binding residues between CapG and gelsolin is high for the Active 1 conformation, and a conserved core is evident for Active 2, indicating that C2-C3 and G2-G3 should interact with actin in similar ways (Fig. 25).

**Figure 25.** Conservation of the alternate domain 2/3 interfaces between gelsolin and CapG. **A:** The Active 1 and Active 2 interaction surfaces on G2 for G3 are colored orange and cyan, respectively. **B:** and **C:** Two views of the Active 1 and Active 2 interaction surfaces on G3 for G2, colored orange and cyan, respectively. Residues in these interfaces that do not show homology between human gelsolin and human CapG are colored red and blue, respectively. **D:** and **E:** Ribbon/surface hybrid representations to allow comparison with active structures of the homologous domains of CapG in Fig. 14.
4.2 Common CapG and gelsolin features suggest common actin-related functions

Taken together, the data presented above suggest that Active 1 and Active 2 conformations can be achieved both by CapG and gelsolin, whether actin is present or not. Interconversion between the two states is unlikely to be pH-induced as the Active 2 conformation of CapG-sev crystallized under basic conditions [91], while that of G2-G3 crystallized at pH 4.6 [124], similar to the pH at which Active 1 conformations of both C2-C3-containing and G2-G3-containing crystals developed.

Having two crystal forms of a protein in which the interdomain interactions are as different as those observed for domains 2 and 3 of either gelsolin or CapG is uncommon and begs the question as to possible physiological roles for each. Detailed analysis reveals that the residues that form the interfaces between domains 2 and 3 in both the Active 1 and Active 2 forms are conserved between CapG and gelsolin, indicating that each conformation has been maintained through evolution (Fig. 25). Hence, the functions of each distinct activated state are likely to be shared by G1-G3 and CapG, perhaps in protein activation or deactivation, filament recognition, filament capping or uncapping, or even, as suggested by our TIRF results, filament severing.

In one locale, the trans-Golgi compartment of the endoplasmic reticulum, the Active 1 conformation would be expected to be the dominant activated form of gelsolin because in that conformation the known furin-sensitive cleavage site is protected [46], whereas the site is fully exposed in the Active 2 conformation (Figs. 24B and 24C). CapG, which
is not exported to the exterior of cells, would not be expected to be found in the
trans-Golgi compartment.

The actin-binding sites of CapG and G1-G3 comprise highly conserved residues in
the Active 1 form and display a conserved core for the Active 2 form, suggesting that
G2-G3 and C2-C3 will interact with actin in the same manner (Fig. 26).

Superimposition of the Active 1 and Active 2 forms of G1-G3 onto the Oda model
for F-actin [29] demonstrates that both forms will be able to contact the actin above that
to which G1 binds (Fig. 27). The Active 2 conformation directly contacts the upper actin
while the Active 1 conformation would require a small rotation to allow G2 to traverse
the 6-9 Å distance to the upper actin. Surface charge analysis reveals that charge
complementarity of the actin filament with the Active 1 conformation of G2-G3 is
significantly greater than with the Active 2 conformation (Fig. 27), suggesting that initial
recognition of the side of an actin filament occurs when the binding protein (gelsolin or
CapG) is in the Active 1 conformation. In low calcium conditions, loss of the stabilizing
calcium ion bound to the G3 type II metal-ion binding site located at the G3-actin
interface (Active 2; Figs. 26C and 26E) may augment this effect.
**Figure 26.** Active 1 and Active 2 binding sites on actin. A, D and F: Interactions of the Active 1 form of G2-G3 with actin. Color coding on actin: interactions with G2 (green), G3 (yellow); non-interacting regions of subdomains 1 and 2 (brown); non-interacting regions of subdomains 3 and 4 (light blue); and ATP (royal blue). Color coding on G2-G3: interactions with actin subdomains 1 and 2 (brown). B, C and E: Interactions of the Active 2 form of G2-G3 with actin. Color coding on actin: interactions with G2 (green), G3 (yellow); non-interacting regions of subdomains 1 and 2 (pink); non-interacting regions of subdomains 3 and 4 (light blue); and ATP (royal blue). Color coding on G2-G3: interactions with actin subdomains 1 and 2 (pink); interactions with actin subdomain 4 (steel blue). In C and D, residues of G2-G3 that interact with actin in both the active states are colored orange. In E and F, residues of G2-G3 that contact actin but do not show homology between human gelsolin and CapG are colored red. In C through F, surface exposed calcium ions are colored cyan.
Figure 27. Charge complementarity between G2-G3 and the actin filament. A and D: G1-G3 (outlined) from the Active 2 and Active 1 conformations, respectively, superimposed on the central of three actin protomers from the Oda model of the actin filament (surface charge representation, blue is positive, red is negative; [29]). B and C: respectively, show surface charge depictions of G1-G3 on detachment from the actin filament and rotation by 180° around the helix axes of A and D. Charged areas for comparison are circled on the respective actin and gelsolin surfaces. Active 1 (C and D) displays a greater charge complementarity for interaction between G2-G3 and actin. The essentially masked PIP2-binding site on G2 in the Active 1 conformation that interacts with actin is labelled.

The putative site on G2 for binding PIP2 [125] forms part of the F-actin-binding site in the Active 1 conformation and is partially buried at its interfaces with actin and G3 (Fig. 27C; Fig. 28D). In contrast, this PIP2 motif forms an exposed and contiguous surface on the Active 2 structure that lies on the same face of G1-G3 as the G1 PIP2-binding region (Fig. 28A), thus presenting two highly positively charged patches ideal for binding PIP2 (Fig. 28B). It, therefore, is plausible to propose that the Active 1
conformation, with compatible binding surfaces for the filament and a buried PIP2 site, represents the actin filament-recognition state, while the Active 2 conformation, having less favorable electrostatic interactions with the filament but heightened possibilities for interaction with PIP2, could represent a previously unanticipated state of the capped barbed end of a filament that is primed for uncapping. An equilibration between the states at the barbed end of an actin filament would have the consequence of accentuating any uncapping signal by the mass action effect of having the favorable binding of PIP2 and release of the Active 2 state shift the equilibrium away from the Active 1 state, leading to rapid and complete uncapping of the filament. Conversion from the side-binding Active 1 conformation to the Active 2 conformation, primed for uncapping, in the presence of actin (Figs. 22 and 24), may be possible by first dissociating the G2-G3 interface, followed by a movement of G2 across the surface of actin (Fig. 29). The options for encircling the filament by the forty crystallographically unaccounted for amino acid residues of the G3-G4 linker are restricted by the binding position of G4. Hence, any interconversion between the active forms of G2-G3 should occur either before G4-G6 has bound or after it has been released from actin. In contrast, the three-domain protein, CapG, may switch between the two states through a similar mechanism, but without the restrictions imposed by additional C-terminal domains.
Figure 28. Exposure of PIP2-binding sites. A and D: Surface representations of the Active 2 [124] and the Active 1 (PDB ID 1RGI) G1-G3/actin conformations, respectively. Putative PIP2-binding sites are colored blue. B and C: Surface charge representations of G1-G3 from A and D, respectively. The PIP2-binding site (residues 161-172) on G2 shows greater surface exposure in the Active 2 conformation.
We have demonstrated that domains two and three from CapG and gelsolin can adopt two active conformations. Surface analysis suggests that the Active 1 conformations will favor recognition at the side of a filament, a required precondition for severing. Subsequent transformation at the newly generated capped barbed end of the filament would lead to Active 2 conformations for the capping proteins, and these would be primed for uncapping by the binding of PIP2.

While CapG is generally an intracellular protein and would not be expected to contain disulfide bonds, the structural data for native C2-C3 crystals, grown in oxidizing conditions, indicate a disulfide bond between Cys165 and Cys178, corresponding to the Cys188-Cys201 disulfide bond of G2-G3 in extracellular gelsolin. The ability to form such bonds in oxidizing conditions, as found in extracellular fluids, emphasizes the
similarity in folding of the two proteins. It also suggests that in addition to possessing intracellular functions that overlap those of gelsolin, CapG that spills into the extracellular space as a result of cell lysis might support the actin-scavenging activity of oxidized extracellular gelsolin. Given the relative concentrated locations of CapG and wide-spread locations of actin, it is reasonably believed that the molar ratio of CapG/actin in different cellular compartments can vary within a broad range, facilitating the alternative physiological function stages of CapG.

4.3 Comparison of new V6 and previous V6 and G6 structures

The V6 structure from activated villin-6M, refined to a resolution of 1.8 Å (Fig. 15A), is highly similar to a 2.0 Å resolution structure of V6 previously reported from this laboratory (Fig. 30A) [121]. It should be noted that the recent structure resulted from crystals grown from pure Villin-6M in a high Ca²⁺ concentration environment to ensure a fully activated state, whereas the earlier structure was obtained from crystals that grew from a solution of a villin/actin complex using a precipitant solution that contained no added calcium salt. In neither structure can a calcium ion be identified. Another feature common to both structures of V6 is that their long helices are in an extended state, as in the structure reported for calcium-bound gelsolin domain G6 (Fig. 30B), and unlike the case for calcium-free G6, where the helix is disrupted by a definite kink (Fig. 30C). At least when in an isolated state, V6 does not require the binding of a calcium ion to induce
straightening of its long helix, in contrast to extension of the long helix of G6 in intact
gelsolin, which correlates with Ca\(^{2+}\)-induced activation of that protein.

**Figure 30.** Structural comparisons of domains V6 of Villin-6M and G6 of gelsolin. The key residues of a putative type II calcium-binding site in each domain are identified. **A:** V6 (PDB ID 3FG7); **B:** G6 with calcium ion bound (extracted from PDB ID 1P8X); **C:** G6 in calcium-free form (extracted from PDB ID 1D0N).
The amino acid sequence of V6 of Villin-6M is ~50% identical to that of G6 of gelsolin [120], which invites direct comparison of both their folding and their functions. When compared with all gelsolin domain structures available, the structure of domain V6 of Villin-6M most closely resembles that of G6 in its Ca$^{2+}$-bound form. The RMSD of $\alpha$-carbon positions for 99 residues between V6 and G6 (Ca$^{2+}$-bound) is 3.0 Å, even though V6 does not possess a bound calcium ion.

Because of the high degree of sequence homology and shared biological activities of gelsolin and villin (minus its HP domain), it is reasonable to assume that Ca$^{2+}$-free intact villin assumes a compact folded structure similar to that for Ca$^{2+}$-free intact gelsolin (Fig. 3). If so, the long helix of the V6 domain should distort in order to avoid steric clashes with the long helix of V4. This apparent contradiction of our observation of a straight helix in V6 devoid of bound calcium can be resolved by noting that, in the predicted structure for intact villin, the V4-V6 latch would be closed and constraints imposed by the proximity of the remaining two domains, V4 and V5, let alone the N-terminal triad, V1-V3, could indeed result in an overall requirement of having a kinked long helix in V6. As part of the activation process induced by Ca$^{2+}$-binding, the V4-V6 latch is opened to expose the actin-binding site on V4. This would remove conformationally imposed restraints on V6, enabling its long helix to assume a less distorted form, even without direct binding of Ca$^{2+}$ to V6. In fact, molecular dynamics studies of gelsolin G6 suggest that if it were isolated from the rest of the protein, the long helix would preferentially assume an uninked conformation [121].
4.4 Fragment G2-G3 retains important properties of intact gelsolin

The reported structures of domains G2 and G3 reveal each to contain a complete type II Ca$^{2+}$-binding site [46, 47]. This is consistent with our observation of the Ca$^{2+}$-enhancement of tryptophan fluorescence of the G2-G3 protein fragment displaying a two-step correlation with increased Ca$^{2+}$ concentration. The suggestion that Ca$^{2+}$-binding to gelsolin domain G2 is a crucial first event in opening the tail-latch of inactive gelsolin [47, 122] suggests that the G2 site is the higher affinity site of the two, and that occupying this site leads to the reported effects of sub-micromolar Ca$^{2+}$ on the biological activities of gelsolin (as discussed in [47]). However, as demonstrated in the previous section, one must be careful in interpreting results from experiments that involve fragments of whole proteins. Contradictions, apparent or real, can emerge.

Enhancement of tryptophan fluorescence suggests that conformational changes are induced that increase the degree of burial of the aromatic side chains in the hydrophobic core of a protein. This interpretation, in the case of G2-G3, is supported by our fluorescence quenching studies on Ca$^{2+}$-bound and Ca$^{2+}$-free G2-G3, using iodide as a fluorescence quenching agent. In the presence of sufficient calcium ions to saturate the metal ion-binding sites on G2-G3, the tryptophan residues of that protein fragment are shielded from iodide ions in the solvent, consistent with burial of the tryptophan side chains in the interior of the protein.

The Ca$^{2+}$-titration and iodide quenching results are in agreement with circular
dichroism studies of isolated G2 in solution [123] that determined isolated G2 to be in a less stable, more open structure in the absence of bound Ca\(^{2+}\). Furthermore, the lack of interaction between G2 and G3 in the reported structure of Ca-free intact gelsolin [45] suggests that isolated G2-G3 may have a relatively open structure in which the two domains are joined by a relatively flexible and extended segment of polypeptide chain, the G2-G3 linker. In contrast, isolated G2 with a divalent metal ion bound is stable in solution [123], and there is a significant contact interface between G2 and G3 in both the Active 1 and Active 2 forms of G1-G3 bound to G-actin (Fig. 6). Specifically, Trp347 in G3 is located at the centre of the interface of G2 with G3 in the Active 1 conformation [46] and would experience a decrease in solvent accessibility relative to its degree of exposure in Ca\(^{2+}\)-free G2-G3.

It is commonly accepted that major structural changes occur when gelsolin binds calcium ions. The fact that a general reorganization of gelsolin from a compact intact globular protein to a more open conformation is supported by the evidence that gelsolin changes its susceptibility to proteolysis at various sites [126], decreases its sedimentation coefficient [127], and doubles its hydrodynamic volume [128]. The major purpose of the reorganization is to unmask actin-binding sites, enabling gelsolin to bind, sever, and cap actin filaments [45, 46, 48, 129]. Achievement of maximal actin-binding functionality correlates with gelsolin reaching full activation, which requires calcium levels in the millimolar region. Therefore, not only high affinity calcium-binding sites, but also the low affinity ones must participate. The release of G2 from G6 (tail latch) requires only
sub-micromolar calcium concentrations and likely involves a high affinity site within G2 [130]. The lower affinity sites within G3 may be one of those reported in previous studies [131–134] to assist in conformational rearrangements that include opening of additional latches observed within the inactive structure, and which are in accord with previously detected subtle effects of calcium on actin filament-severing by the isolated N-terminal half of gelsolin.

At present, a number of structures are available for various fragments of gelsolin, including both the N- [46] and C-terminal halves in the presence of calcium ions [135] and actin [47, 48], as well as for intact gelsolin in its inactive form [45]. Still lacking is structural data for activated intact gelsolin. The ability of G2-G3 to bind to actin in affinity chromatography and cosedimentation studies indicates that the G2-G3 fragment of gelsolin inherently includes the actin-binding ability reported for both activated intact gelsolin and its isolated N-terminal half. As for the latter case, G2-G3 retains actin-binding ability in the absence of free calcium ions, as observed in the affinity chromatography study, a result of relevance to apoptosis (programmed cell death) [136]. Caspase-3 cleaves gelsolin at a site in the G3-G4 linker that connects the N- and C-terminal halves of the protein, releasing G1-G3 as a Ca\(^{2+}\)-insensitive microfilament severing agent [136].
5. Conclusions and Future Work

5.1 Conclusions

Fragment G2-G3 of gelsolin is responsible for initial contact and association with filaments of F-actin. Without this initial binding event, subsequent filament-severing and capping activities of gelsolin in its role as a modulator of actin dynamics would not be possible. In our efforts to understand how the structure of G2-G3 dictates its function, we attempted to crystallize this fragment using a wide array of commercially available screening conditions, as well as conditions that have proved successful in crystallization of gelsolin, gelsolin fragments, and complexes between actin and gelsolin fragments. While we were able to express, purify and study G2-G3 in solution, we were unable to grow crystals suitable for crystallographic analysis of the structure of the particular fragment of interest.

However, using recombinant fragments of two members of the gelsolin superfamily of actin regulatory proteins, we were able to obtain two crystallographic structures that provided atomic-level information on how G2-G3 would be expected to be activated by binding Ca$^{2+}$ ions, and how it might interact with actin. Specifically, we obtained a 1.8 Å resolution structure for villin domain V6 and a 2.0 Å resolution structure for fragment C2-C3 of CapG, both in the presence of activating levels of Ca$^{2+}$.

The structure of V6 validates one obtained earlier in our laboratory under different crystallization conditions. Both structures lack a bound Ca$^{2+}$ in spite of clear evidence for
the presence of a type II metal ion-binding site. While it was speculated that the earlier structure represented one in activating conditions, the absence of added calcium salts in the precipitant solution used to induce crystallization left some room for doubt. The precipitant used in the current work included 0.1 M calcium acetate, more than sufficient to fully activate any member of the gelsolin family. Further evidence for the activated state in both structures of V6 is the lack of a significant kink in the single long helix within each. This parallels the observation that activated gelsolin domain G6 possesses a correspondingly straight long helix, in contrast to the kinked state of the helix within inactive, Ca\(^{2+}\)-free gelsolin [45]. Because the first three domains within both gelsolin and villin are related to the next three domains as a result of a gene duplication event in their shared evolutionary history, the structure of the sixth domain of each is likely to closely resemble that of the third [45]. That is, the structure of V6 from villin, while being most like that of G6 in gelsolin, is also very close to that of G3 in gelsolin. The conclusion that the straightening of the long helix in G6 contributes energetically to the activation process [121], therefore, extrapolates to a similar role for the straightening of the long helix of G3.

The structure of fragment C2-C3 from CapG, the analogue of fragment G2-G3 from gelsolin, yielded completely novel and unexpected results, as a previously reported structure for a mutant, CapG-sev [91], showed C2-C3 in that structure to be arranged in a way that markedly differed from the published arrangements of G2-G3 in either intact inactive gelsolin [45] or in activated G1-G3/actin (which subsequently was used to define
the Active 1 conformation of G2-G3) [46, 137]. In fact, C2-C3 in the CapG-sev structure turned out to resemble G2-G3 in a recently grown and studied crystal in our laboratory [124], which yielded a novel (Active 2) conformation for G2-G3 in G1-G3/actin. In contrast, the structure reported here for activated C2-C3 from native CapG is in the Active 1 conformation. Hence, within at least this pair of gelsolin superfamily members, two active forms exist. Active 1, which shows considerable surface charge complementarities to the surface of the side of an actin filament, is postulated to be the form in which activated gelsolin exists when it encounters and binds to the side of a filament to initiate the severing and capping processes. Active 2, which exposes a PIP2-binding site in the second domain (G2 or C2), likely exists preferentially at the capped barbed end of an actin filament. Such a capped filament would be primed for uncapping in response to the binding of PIP2, and lead to rapid filament growth upon release of the cap. The conformational transformation between Active 1 and Active 2 states can be modeled to take place in a sterically feasible manner at the barbed end of a capped filament. Thus, our work has led to discovery of a previously unpredicted step in the actions of gelsolin-like proteins in their regulation of actin filament dynamics, the interconversion of such proteins between two different activated states, each with its own specific preferred activity. It also led to a re-examination of the reported lack of severing activity of CapG in light of the Active 1 conformation being the one most suited to side-binding on a filament, and demonstration that relatively elevated CapG levels (~1:1 with actin) can sever filaments.
Our solution studies with the G2-G3 fragment of gelsolin confirm that this fragment retains several important aspects of the activity of the intact protein. It contains two Ca\(^{2+}\)-binding sites, corresponding to a type II site on each of G2 and G3, with the higher affinity site likely being located on G2. Binding Ca\(^{2+}\) buries tryptophans, as confirmed by fluorescence quenching experiments. G2-G3 binds to and co-sediments with F-actin, and is able both to nucleate G-actin polymerization and to shorten pre-existing F-actin filaments in light scattering studies.

5.2 Future work

Although we have advanced the understanding of the structures and biochemical characteristics of gelsolin superfamily members, many questions remain. Firstly, the structures of activated intact gelsolin and of its G2-G3 fragment remain elusive in spite of years of trying in several laboratories around the world. Aside from directly varying the composition and temperature of the precipitant solution employed in crystallization trials, methods that have been successful in other difficult cases should be applied to these systems. Covalent cross-linking of peptide chains, either within individual proteins or between proteins in complexes such as gelsolin forms with actin, could lead to structures that are simpler to crystallize in that they are less likely to degrade by either chemical or enzymatic means. Sequential dehydration and re-equilibration steps subsequent to initial crystallization trial times may lead to better quality crystals that
include a lower percentage of mobile loops and side-chains.

The current CapG structure reveals only the C2-C3 fragment of the protein. We must continue crystallization trials to obtain structures of intact CapG, both in the presence and absence of divalent metal ions and actin, respectively.

The activity of gelsolin is known to be influenced by interaction with a number of ligands, including PIP2, among other polyphosphoinositides, and such messenger molecules as lysophospholipids, and toxins such as lipopolysaccharides. While G2 has been implicated in the binding of PIP2, the locations of the other sites are unknown. By soaking such molecules into existing crystals, such as those of C2-C3 or gelsolin or fragments of gelsolin, we could learn how they affect the normal actions of the proteins to which they bind.

Structural techniques other than X-ray crystallography could be brought to bear on these proteins to get information at resolutions higher than achievable by standard biochemical spectroscopy methods conducted on protein solutions. Multinuclear, multidimensional nuclear magnetic resonance (NMR) methods can yield resolutions comparable to crystallographic methods in determination of proteins of the size of a gelsolin-like domain. Such structures are available, for example, for villin domain V1 [138] and severin domain S2 [139], but have not yet successfully been applied to large fragments of this family of proteins. At somewhat lower resolution, small angle X-ray scattering (SAXS) experiments have been used, taking advantage of published crystallographic structures, to follow shape changes in intact gelsolin on binding Ca$^{2+}$ [70]
or in response to changes in pH [140]. Additionally, advances in electron microscopy that have led to the latest models for F-actin have the resolving power to identify and structurally characterize proteins that decorate the sides of F-actin filaments [21, 22, 141].
References


