Structural Studies of Gelsolin and Binding Partners

by

Anantasak Loonchanta

B.Eng., Suranaree University of Technology, 2003
M.Sc., Uppsala University, 2005
M.Sc., Suranaree University of Technology, 2007

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)

August 2012

Abstract

Gelsolin is composed of six similarly folded domains, G1 through G6. It controls actin dynamics through its calcium-sensitive ability to bind to, sever, and cap F-actin filaments. This project was designed to test and extend current models for gelsolin activity by growing crystals of intact gelsolin or large fragments of gelsolin in protein complexes with its main target, actin, as well as with specialized gelsolin-specific antibody fragments, nanobodies. Gelsolin-specific nanobodies (GsnVHH) may serve to lock in either resting or activated forms of gelsolin structure, but are of intrinsic interest on their own in that they have potential therapeutic value.

This thesis describes cloning, expression, purification, activity assays, crystallization details, and solution of the structures of six proteins and protein complexes. 1) *Recombinant human G1-G3 bound to natural source rabbit actin (G1-G3:actin).* The present structure is to a higher resolution than earlier ones, providing increased confidence in positioning of amino acid side chains and bound Ca^{2+} ions. 2) *Isolated, activated gelsolin domain G3.* While not previously reported in the literature, this structure is virtually identical to that observed for activated G3 within the G1-G3:actin complex. 3) *Activated recombinant human G4-G6.* The source material and crystallization conditions are different from those reported elsewhere, but the results confirm that the C-terminal half of gelsolin can be fully activated in the presence of Ca^{2+}, even in the absence of actin. 4) and 5) *Two crystal structures*

of GsnVHH11. This nanobody had not previously been crystallized. The two structures differ in the path of the CDR1 loop that is involved in binding gelsolin. 6) *GsnVHH9*. The results, from different materials and crystallization conditions than previously reported, confirm the validity of the previous structure.

Finally, runs using *in silico* protein-protein docking software suggest possible binding sites for GsnVHH11 and GsnVHH13 on activated G2-G3 and activated G4-G6, respectively.

Preface

I cloned, expressed, purified, and assayed activities of the following gelsolin constructs: FL-gelsolin, Gel-NL, and Gel-CL. In addition, I prepared an acetone powder of minced rabbit skeletal muscle and purified G-actin from the powder. Gelsolin domain G2-G4 was cloned, purified and provided by Hui Wang, a former student in the Burtnick laboratory. Nanobodies GsnVHH9 and GsnVHH11 were provided by a collaborator, Sarah De Clercq of the Gettemans laboratory, VIB Department of Medical Protein Research, Ghent University, Belgium. I purified and performed crystallization screening tests on all protein complexes that involved gelsolin constructs with actin or with nanobodies.

I collected and processed initial X-ray diffraction data for all crystals that grew from my screens. I collected the data either in the X-ray Crystallography Hub at the UBC Centre for Blood Research (CBR), or on site or remotely at the Canadian Macromolecular Crystallography Facility (CMCF), Canadian Light Source (CLS), Saskatoon. Some X-ray diffraction data were collected and processed at the National Synchrotron Radiation Research Center (NSRRC), Taiwan, by collaborators from the laboratory of Dr. R. Robinson, Institute of Molecular and Cell Biology, Singapore. I solved and analyzed all structures presented in this work under the supervision of Drs. Robinson and Burtnick.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABPs – Actin binding proteins

ADP – Adenosine diphosphate

AMPPNP – β,γ-Imidoadenosine-5'-triphosphate

ATP – Adenosine triphosphate

a.u. – Arbitrary unit

cDNA – Complementary deoxyribonucleic acid

CDR – Complementarity determining region

CH – Constant domain of heavy-chain.

CL – Constant domain of light chain

D-loop – DNase I-binding loop

DNA – Deoxyribonucleic acid

DNase I – Deoxyribonuclease I

F-actin – Filamentous actin (actin filament)

Fc – Fragment crystallizable region.

FL-Gelsolin – Full-length gelsolin

G1+ – Gelsolin domain 1 plus a part of the G1-2 linker

G-actin – Globular actin (monomeric actin)

Gel-CL – Gelsolin with part of G3-G4 linker extension to its amino-terminus (residues Asn368 to Ala782)

Gel-NL – Gelsolin with part of G3-G4 linker extension to its carboxyl-terminus (residues Met52 to His428)

GF – Gel filtration

GsnVHH – Gelsolin-specific nanobody

HCAb – Heavy-chain antibody

HEPES – 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffer

Hisx8 – Eight-histidine tag

IPTG – Isopropyl-β-D-1-thiogalactopyranoside

K_d – Dissociation constant

kDa – Kilodalton

MES monohydrate – (2-(N-morpholino)ethanesulfonic acid monohydrate

MIB – Buffer containing malonic acid:imidazole:boric acid at a molar ratio of 2:3:3

Nb – Nanobody

PCR – Polymerase chain reaction

PEG – Polyethylene glycol

PIP2 – Phosphatidylinositol-4,5-bisphosphate

PS – PreScission Protease cleavage site

SAXS – Small-angle X-ray scattering

SD – Subdomain

SDS-PAGE – Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SPG – Buffer containing succinic acid:sodium dihydrogen phosphate:glycine at a molar ratio of 2:7:7

TMR – Tetramethylrhodamine-5-maleimide

VL – Variable light-chain.

VH – Variable heavy-chain

VHH – Variable heavy-chain of heavy chain antibody

Acknowledgements

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

First and foremost, I would to thank my supervisor Dr. Les Burtnick who has supported me throughout my thesis with his help, patience, encouragement, effort, and chemistry knowledge whilst allowing me freedom to work in my own way. Without him this thesis could not have been completed or written. Moreover, Les sent me to data collection summer school, signed up for synchrotron beam time whenever I needed it and bought me a new computer suitable for X-ray crystallography software or whatever I needed for my work, without these factors, the work would not have been done fast and efficiently. One would not wish for a better or friendlier supervisor. Thank you for everything, Les.

I would like to thank Dr. Robert Robinson who is my first teacher of X-ray crystallography. Not only did Bob collect X-ray data for us at the NSRRC, Taiwan, but his knowledge and valuable suggestions were essential factors contributing to the completion of the thesis. Without him, as well, this thesis would never have existed. Thank you very much, Bob.

I would like to thank Dr. Mariena Ketudat-Cairns who is my first molecular biology teacher and was the first who gave me an opportunity to go from chemical engineering to molecular biology. She taught me all I know about molecular biology, a major tool used in this research..

I am thankful to my supervisory committee Dr. Chris Orvig, Dr. Pierre Kennepohl, and Dr. Stephen Withers, who have made me aware of the importance of knowing chemistry. I especially thank Dr. Orvig who also made valuable corrections to my thesis. Also, I am grateful for all the precious comments for the thesis from Dr. Ross MacGillivray, Dr. Don Douglas, and Dr. Joanna Krueger.

I would like to thank my collaborators Dr. Jan Gettemans and Sarah De Clercq for providing nanobodies and for their constructive comments and suggestions. Sarah sent me nanobodies whenever needed and she has also been a link between Jan and me throughout. The thesis would not have been completed without them. Thank you Dr. Jantana Wongsantichon and Dr. Xue (Albert) Bo for all data collection in Taiwan, data transfer, and suggestions in solving structures. I appreciate your help guys! Without you both, all processes would have been delayed. I would like to thank Dr. Sakesit Chumnarnsilpa for help in construct design and for constructive suggestions. Thank you Dr. Hui Wang for the G2-G4 protein that I used in my research.

I thank the UBC Centre for Blood Research for providing excellent research facilities and a friendly environment, not to mention the CBR X-ray Crystallography Hub, the CBR Fermentation Suite, and the CBR Mass Spectrometry Suite.

I would like to thank Alex Diaz for supporting me through his positive and cheerful attitude and always being there for me when I need it. He is part of this achievement. Thank you my friends Aurora, Peter, and Leke for support while I was writing my thesis.

Finally, I would like to thank my dear family: my parents, my sisters, and my aunt who reside in Thailand. Their support and encouragement throughout my studies has pushed me to today's success.

Chapter 1

Introduction

1.1 The cytoskeleton

The cytoskeleton is composed of an interlocking three-dimensional meshwork of three general types of strands (Fig. 1.1): microtubules (~22 nm in diameter), intermediate filaments (8 – 10 nm in diameter), and microfilaments (6 – 7 nm in diameter). It extends throughout the cytoplasm of cells, providing infrastructure to maintain cell shape, as well as to organize, support and transport organelles and vesicles (Alberts *et al*., 2002). The strands differ in size, composition and specific function, but each comprises simple protein subunits that undergo dynamic assembly into linear polymers of uniform thickness, and disassembly back to monomeric subunits (Fig. 1.1).

The three types of cytoskeletal strands interact with various other proteins. These may introduce cross-links with other filaments of the same or of a different type, may influence assembly or disassembly, or may act as motors to move cytoplasmic organelles along filamentous tracks. The cytoskeletal arrangement within cells change dramatically during biological processes such as mitosis, cytokinesis, or cell movement (Lodish *et al*., 2000).

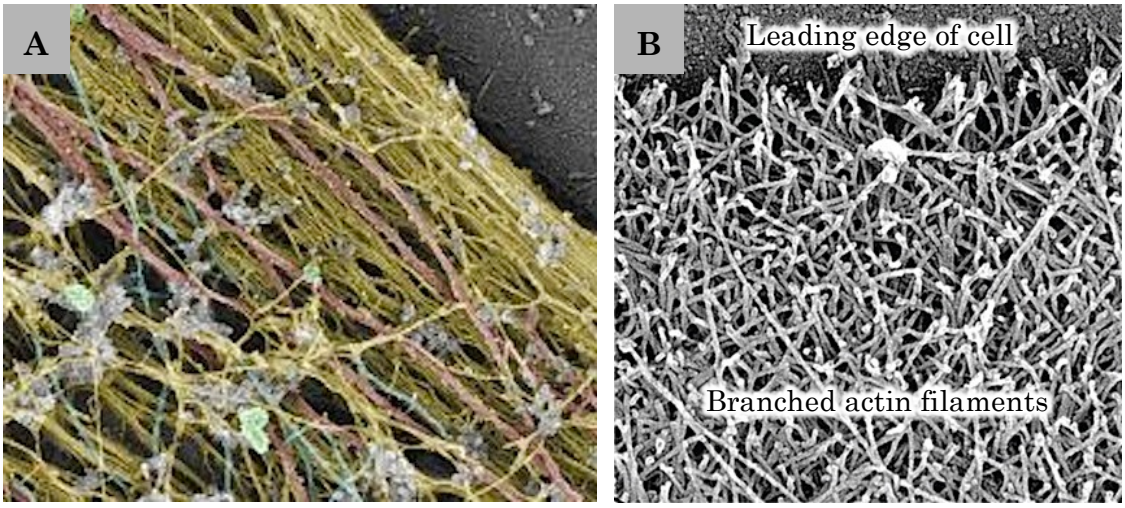

Figure 1.1 Electron micrographs showing cytoskeletal proteins inside a cell. **(A)** The arrangement of microtubules (red), intermediate filaments (green), and actin microfilaments (yellow) (Copyright Tatyana Svitkina, courtesy of the Biology Image Library http://biologyimagelibrary.com/imageID=48810) **(B)** The network of actin filaments at the leading edge of a motile cell. (Reprinted from Pollard and Berro, 2009).

1.2 The actin cytoskeleton

Cells are able to migrate by swimming or crawling (Alberts *et al.*, 2002). Amoeboid or pseudopodal movement is best characterized by dynamic protrusion of the leading edge of a cell to form filopodia or lamellipodia (Fig. 1.1B) (Pollard and Berro, 2009). The central machinery that powers such motion depends on interactions of actin with a number of other proteins in different parts of a cell (Lodish *et al.*, 2000).

1.2.1 Actin and actin dynamics

Actin is a protein of ~43 kDa in mass and 375 amino acids in length. It was discovered in a muscle tissue extract (Straub, 1942; Straub, 1943). Along with myosin, it forms a regular array of filaments (actomyosin) that constitute more than half of the total protein in muscle. Subsequent to the discovery of actin in muscle, it has been found in all other eukaryotic cells. Both the shared and distinctive structural features of the actin machines that generate contractility in muscle and motility in cells make actin the major research focus of many laboratories internationally (Hatano and Oosawa, 1966; Adelman and Taylor, 1969; Dominguez and Holmes, 2011; Popp and Robinson, 2011).

Actin is a highly conserved protein and is one the most abundant proteins in the eukaryotic cells (Pollard and Cooper, 2009). Actin not only provides force to drive cell movement, it participates in a wide range of other cellular functions, such as providing internal mechanical support and tracks for movement of intracellular materials, for sensing environmental forces, for internalizing membrane vesicles, and for dividing the cell in two. Actin is essential for the survival of most cells and thus is highly conserved among species and isoforms.

There are six different actin isoforms in birds and mammals, encoded by six separate genes (Perrin and Ervasti, 2010). The isoforms share 93% identity in amino acid sequence with each other.

Inside cells, actin exists in two forms: G-actin (globular monomer) and F-actin (filamentous polymer) (Wegner, 1976; Korn *et al.*, 1987; Bugyi and Carlier, 2010). Under physiological conditions, G-actin spontaneously polymerizes into long, double-helical, stable F-actin (Fig. 1.2), independently of any actin-binding proteins (Pollard, 2007). The polymerization process starts slowly (nucleation phase), because actin dimers are unstable. But once nuclei of three or four actin subunits have developed, polymerization proceeds rapidly (elongation phase) until a steady state in which the average filament length is stable. During elongation, one end of the filament grows much more rapidly than the other. The fast-growing end is called "barbed" or "+"-end, while the slow-growing end is called "pointed" or "–"-end. In the steady state, dynamic addition and deletion of actin subunits occurs at both ends of the polar filament. Under certain conditions, net depolymerization occurs at the pointed end, with net polymerization at the barbed end, resulting in a phenomenon called "treadmilling" (Fig. 1.2; Pak *et al.*, 2008).

Generally, G-actin that adds onto the end of a filament during elongation contains a bound adenosine triphosphate (ATP) molecule. Soon after incorporation into a filament, hydrolysis of the terminal phosphate (γ-phosphate) on ATP leaves adenosine diphosphate (ADP) as the nucleotide bound to each actin unit (Fig. 1.2). Incorporation into a filament and hydrolysis of its bound ATP results in subtle structural changes in an actin subunit. Regulatory proteins that enhance disassembly of F-actin generally

do so by interacting with sections of the filament consisting of ADP-actin subunits (Fig. 1.2).

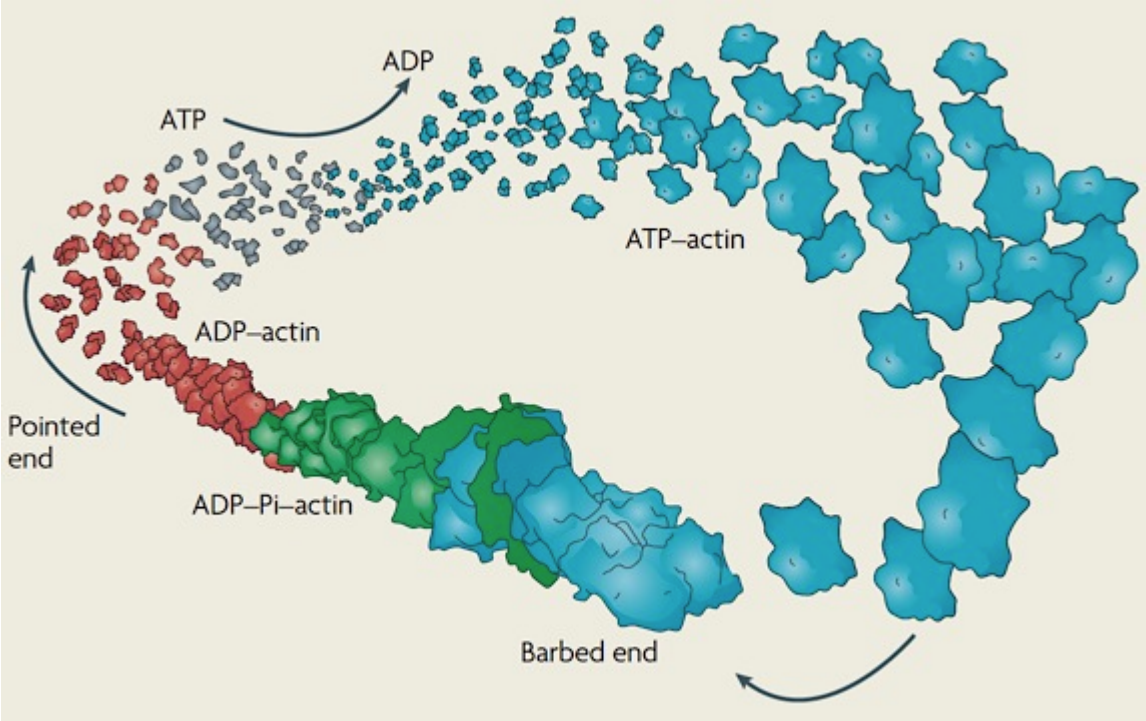

Figure 1.2 Actin dynamics and treadmilling. (From Pak *et al.*, 2008). Reprinted with permission from Macmillan Publishers Ltd.

1.2.2 Structure of globular actin (G-actin)

The first X-ray crystallographic structure of G-actin was determined in the form of a heterodimeric complex with DNase I, which blocks the polymerization of G-actin and enables crystal growth (Kabsch *et al.*, 1990) (Fig. 1.3). Several corroborating crystal structures have since been determined, using a variety of tricks to prevent F-actin formation in the

protein solutions employed (Dominguez and Holmes, 2011). These include complex formation with actin-binding proteins (ABPs) such as profilin (Schutt *et al.*, 1993; Chik *et al.*, 1996), gelsolin (McLaughlin *et al.*, 1993; Robinson *et al.*, 1999; Choe *et al.*, 2002; Burtnick *et al.*, 2004), vitamin D binding protein (DBP) (Otterbein *et al.*, 2002), thymosin-β4 (Irobi *et al.*, 2004), or ciboulot (Hertzog *et al.*, 2004). Also, binding to or reaction with certain small molecules blocks polymerization and has permitted growth of modified G-actin crystals, *e.g.*, a tetramethylrhodamine-5-maleimide (TMR)-modified G-actin with bound ADP (Otterbein *et al.*, 2001) or AMPPNP, a non-hydrolyzable analog of ATP (Graceffa and Dominguez, 2003), an antiparallel actin dimer crystallized with ATP and Latrunculin-A bound in the nucleotide-binding cleft (Bubb *et al.*, 2002), and both G-actin and an antiparallel actin dimer after ATP-ribosylation (Margarit *et al.*, 2006). The structure of an uncomplexed, unmodified G-actin is now available (Wang *et al.*, 2010), where crystals grew in a G-actin solution that included the presence of a chaperone protein, Hsp27.

Despite the different conditions employed to crystallize G-actin, the resulting structures share many common features. The polypeptide chain folds into a right-handed αβα-unit, a multi-stranded β-sheet, and a β-meander and can be visualized in terms of four subdomains, SD1 – SD4, respectively, being held together to in great extent by a bound nucleotide and by salt bridges (Kabsch *et al.*, 1990). SD1 and SD2 together form the small or

outer domain of G-actin, while SD3 and SD4 together form the large or inner domain, named because of their relative sizes and positioning with respect to their proximities to the long axis in models for F-actin filaments. SD1 and SD3 are related structurally and are thought to have evolved by gene duplication, whereas SD2 and SD4 may be regarded as large insertions into SD1 and SD3, respectively (Kabsch *et al.*, 1990).

The overall structure of G-actin is relatively flat, with dimensions 55 Å × 55 Å × 35 Å (Fig. 1.3B). It contains two clefts, the deep upper one between SD2 and SD4 and the shallow lower one between SD1 and SD3. The outer and inner domains are separated by the clefts, but remain joined by a hinge formed by a loop centred at Lys336 and a linker helix spanning from Gln137 through Ser145. Another important linkage between the domains is formed by a nucleotide (ATP or ADP) associated with a divalent cation (Ca^{2+} or, more likely in cells, Mg^{2+}) that sits at the bottom of the upper cleft, which is also known as the nucleotide-binding cleft. The lower cleft mediates longitudinal contacts between actin units in F-actin and is the docking site for most ABPs (Fujii *et al.*, 2010; Oda *et al.*, 2009). Because most of the amino acid side chains that line the cleft are hydrophobic (Tyr143, Ala144, Gly146, Thr148, Gly168, Ile341, Ile345, Leu346, Leu349, Thr351, and Met355), and hydrophobic interactions contribute most strongly to contacts made here, the lower cleft is also known as the hydrophobic cleft.

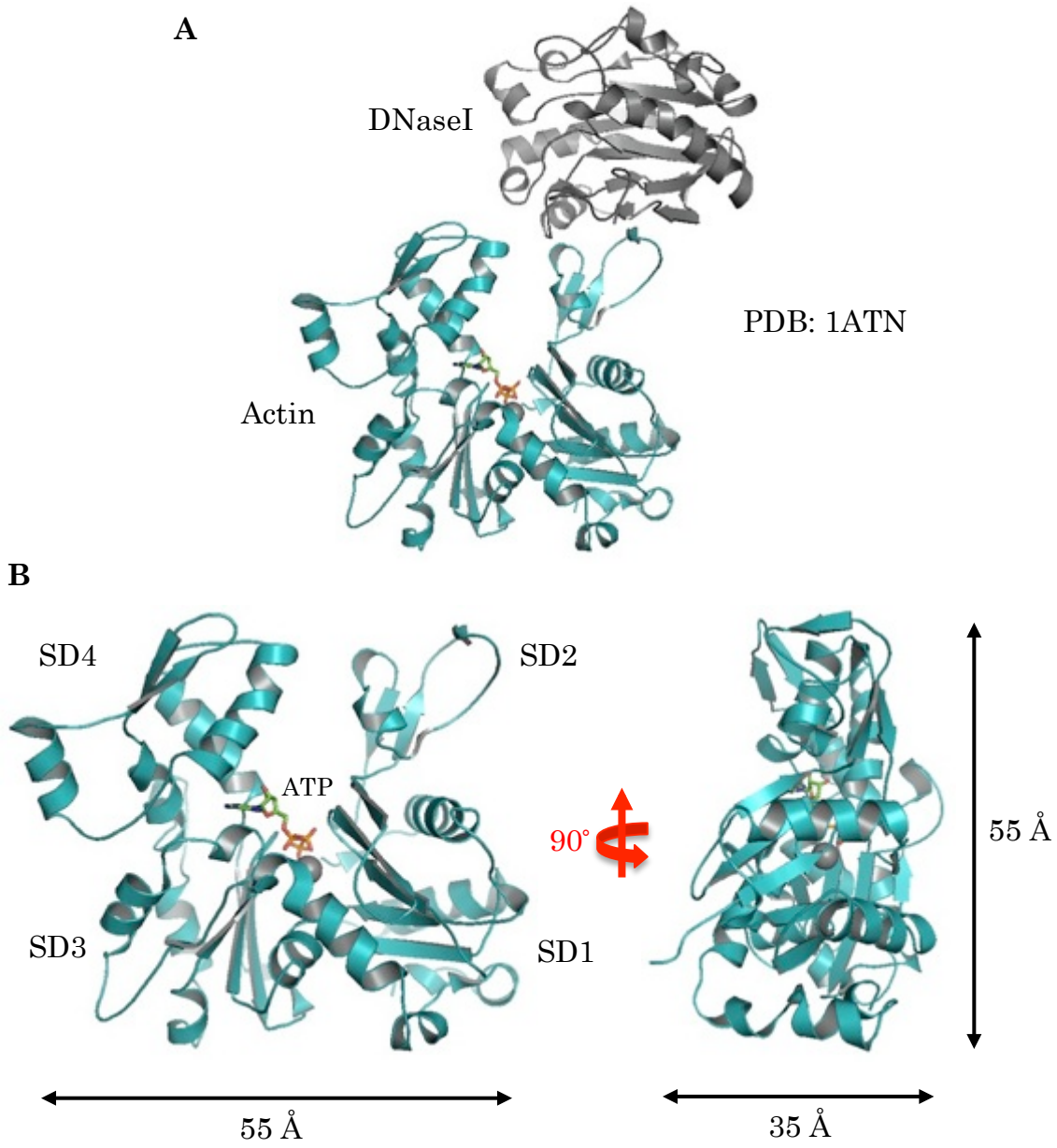

Figure 1.3 The first crystal structure of G-actin. **(A)** The structure of the DNaseI:G-actin complex (PDB code 1ATN; Kabsch *et al.*, 1990). **(B)** The structure of G-actin extracted from (A) and its view after 90°-clockwise rotation about the vertical of the structure on the left. SD1, 2, 3, and 4 refer to actin subdomains 1, 2, 3, and 4, respectively. The grey sphere represents a Ca^{2+} ion. Figures were prepared using PYMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC).

Only minor structural differences exist between the ATP- and ADP-bound states of G-actin, but the identity of the bound nucleotide has a great influence on the binding affinities between actin and ABPs and on the degree of contact between actin units within F-actin (Dominguez, 2004). These differences primarily involve two loops: the Ser14 β-hairpin loop in SD1 and the so-called sensor loop in the C-terminal portion of SD2, which carries the methylated His73 (Graceffa and Dominguez, 2003). The Ser14 loop in SD1 is structurally equivalent to the Asp157-containing loop in SD3. These two loops bind nucleotide phosphates. In the ATP-bound state, Ser14 forms hydrogen bonds with the γ-phosphate of ATP and with the main chain of the sensor loop. But in the ADP-bound state, once the γ-phosphate of ATP has been released, Ser14 rotates to form a hydrogen bond with the β-phosphate of ADP. Simultaneously, the sensor loop "senses" the new state of the bound nucleotide and moves to occupy the space vacated by the γ-phosphate.

The DNase I-binding loop, or D-loop (residues 39-51), of G-actin is located at the top of SD2 (Fig. 1.3B) and, as its name implies, mediates important interactions within the DNase I:G-actin complex (Kabsch *et al.*, 1990). The D-loop is disordered in most crystal structures of G-actin, but can adopt a well-defined conformation, such as a short α-helix in the TMR-modified, ADP-bound state (Otterbein *et al.*, 2001). The D-loop plays a critical role in longitudinal contacts between actin units within F-actin (Fujii *et al.*, 2010). Slight conformational changes in the Ser14 and sensor loops of G-actin

when ATP is hydrolyzed to ADP are thought to propagate to the D-loop in SD2, which could explain the decreased stability of the ADP-state of F-actin (Graceffa and Dominguez, 2003).

1.2.3 Structure of filamentous actin (F-actin)

Initial detailed visualization of the structure of F-actin involved electron microscopy (EM) of negatively stained actin fibers (Fig. 1.4). The images were interpreted to represent F-actin as two beaded chains that turn gradually around each other in a right-handed double-helical arrangement with a half-pitch of 35.9 nm (Hanson and Lowy, 1963). Alternatively, and more in line with the polymerization process as currently understood, thirteen actin units can be interpreted as a repeating section of beads in a single left-handed helical strand, forming almost exactly six left-handed turns in an axial distance of 35.9 nm. The rise per actin unit is 2.76 nm and the twist per molecule is –166.6°. This value is sufficiently close to 180° that the structure on first sight appears to be two long-pitch chains twisted into a right-handed double helix. The overall diameter of the filament is about 80 Å.

It has not yet been possible to crystallize actin filaments to obtain a high-resolution crystallographic structure. The first model for the atomic structure of F-actin was based on 8-Å resolution X-ray fiber diffraction data collected from oriented F-actin gels (Holmes *et al.*, 1990) prepared by polymerization of concentrated solutions of G-actin (Popp, *et al.*, 1987) and

fitted with refinement-modified atomic coordinates from the structure of G-actin (Kabsch *et al.*, 1990). This model for F-actin is widely known as the Holmes model, or as the Heidelberg model, and has served as a reference against which all subsequent models have been compared.

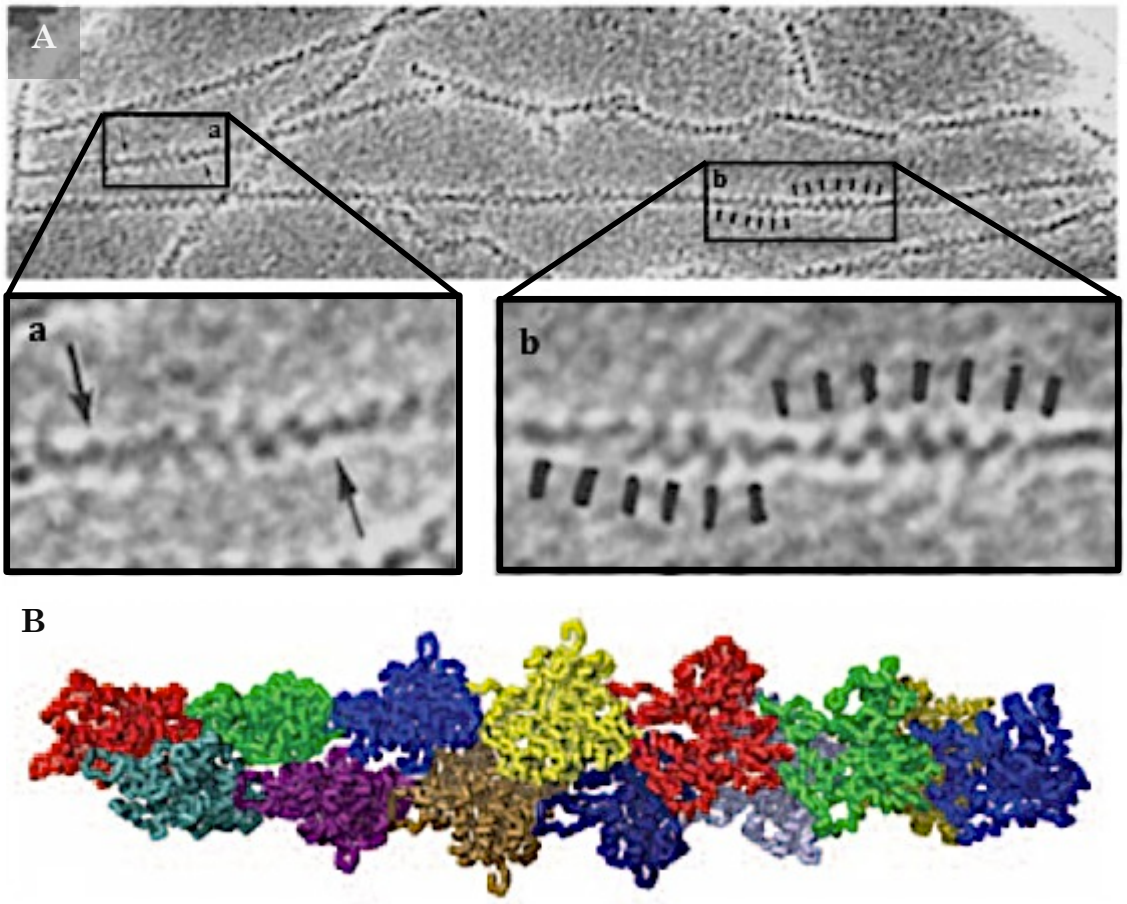

Figure 1.4 Structure of F-actin. **(A)** EM images of F-actin (Hanson and Lowy, 1963), reprinted with permission from Elsevier Publishers. **(B)** Structure of F-actin based on Oda *et al.*, 2009, (from Bugyi and Carlier, 2010, and reproduced with permission of Annual Reviews *via* Copyright Clearance Center).

Recently, higher resolution F-actin structures have been obtained. Oda *et al.*, 2009, reported an F-actin structure with 3.3 Å resolution in the radial direction and 5.6 Å resolution along the equator, using relatively short gelsolin-capped filaments in a sol subjected to a strong magnetic field to maintain a high level of order. The structure of F-actin at 6.6 Å resolution, reported by Fujii *et al.*, 2010, was directly visualized by electron cryomicroscopy (cryoEM) and fitting the atomic coordinates for ADP-G-actin (Otterbein *et al.*, 2001; PDB code 1J6Z) into the cryoEM map.

Conversion between actin monomer structures modelled in F-actin and observed in G-actin involves two types of rotations: a scissors-like opening and closing of the nucleotide cleft within the plane of Figure 1.3B, and a perpendicular propeller-like rotation of the outer domain with respect to the inner one around the hinge connecting them at the base of the nucleotide-binding cleft (Galkin *et al.*, 2002; Oda *et al.*, 2009; Tirion *et al.*, 1995). Gln137, at the start of the hinge helix, is thought to play a critical role in actin-catalysed nucleotide hydrolysis (Iwasa *et al.*, 2008), which is activated on the transition of actin monomer structures from G-type to F-type (Carlier *et al.*, 1994), presumably because interdomain rotation brings the side chain of Gln137 into contact with the γ-phosphate of the ATP (Fujii *et al.*, 2010; Oda *et al.*, 2009).

1.3 Gelsolin

Gelsolin was discovered in macrophage extracts because of its calcium-sensitive ability to activate the gel-sol transformation of actin filaments (Yin and Stossel, 1979). It occurs in a wide range of eukaryotic cells and is best known for its involvement in regulation of the dynamics of cytoskeletal architecture during actin-based cell motility (Kwiatkowski *et al.*, 1989; Yin, 1999). It also is involved in programmed cell death (apoptosis) in certain vertebrate cells (Kothakota, *et al.*, 1997; Kamada, *et al.*, 1998). Gelsolin is expressed in two principal isoforms: cytoplasmic and plasma (secreted). Cytoplasmic gelsolin regulates cell motility and architecture. Plasma gelsolin possesses a 25-amino acid N-terminal extension compared to the cytoplasmic form and functions as part of an actin filament scavenging system in blood (Vasconcellos and Lind, 1993). Both forms control actin dynamics through their calcium-sensitive abilities to bind to and sever actin filaments and to cap the faster growing (barbed) ends of such filaments. In addition, cytoplasmic gelsolin, in response to appropriate uncapping signals, such as phosphatidylinositol-4,5-bisphosphate (PIP2) (Hartwig *et al.*, 1995; Janmey *et al.*, 1987), can nucleate rapid filament growth from a pre-existing pool of gelsolin-capped F-actin oligomers (Yin *et al.*, 1981; Doi and Frieden, 1984).

Gelsolin is composed of six similarly folded domains, G1 through G6, respectively, that are connected by linker polypeptides of varying lengths (Fig. 1.5A) (Kwiatkowski *et al.*, 1986). Sequence analysis shows that G1 is

most closely related to G4, G2 to G5, and G3 to G6, suggestive of a gene duplication event. In addition, all six domains are thought to have originated from a common ancestral domain. Studies involving proteolysis and mutation have identified three domains of gelsolin that are involved in binding actin: G1, a calcium-independent G-actin binding domain, G2, a calcium-independent F-actin binding domain, and G4, a calcium-dependent G-actin binding domain (Bryan, 1988).

At sub-micromolar concentrations of Ca^{2+}, intact gelsolin is in an inactive form that is not able to bind actin. The structure of inactive (Ca^{2+}-free) gelsolin was first solved using crystals of equine plasma gelsolin (Burtnick *et al.*, 1997) (Figs. 1.5B and 1.5C). Recently, the structure of Ca^{2+}-free human gelsolin also has become available (Nag *et al.*, 2009). Both structures confirm the similarity of the six domains, as originally predicted from amino acid sequence analysis (Kwiatkowski *et al.*, 1986) and show that the six domains fold into a compact globular structure in which the F-actin binding site on G2 is masked by interactions between domains G2 and G6, in particular, with a short helical extension at the C-terminus of G6, now called the "helix latch" or "tail latch" (Fig. 1.5). Removal of the latch drastically interferes with the regulation of gelsolin by calcium ions (Kwiatkowski *et al.*, 1989; Way *et al.* 1990).

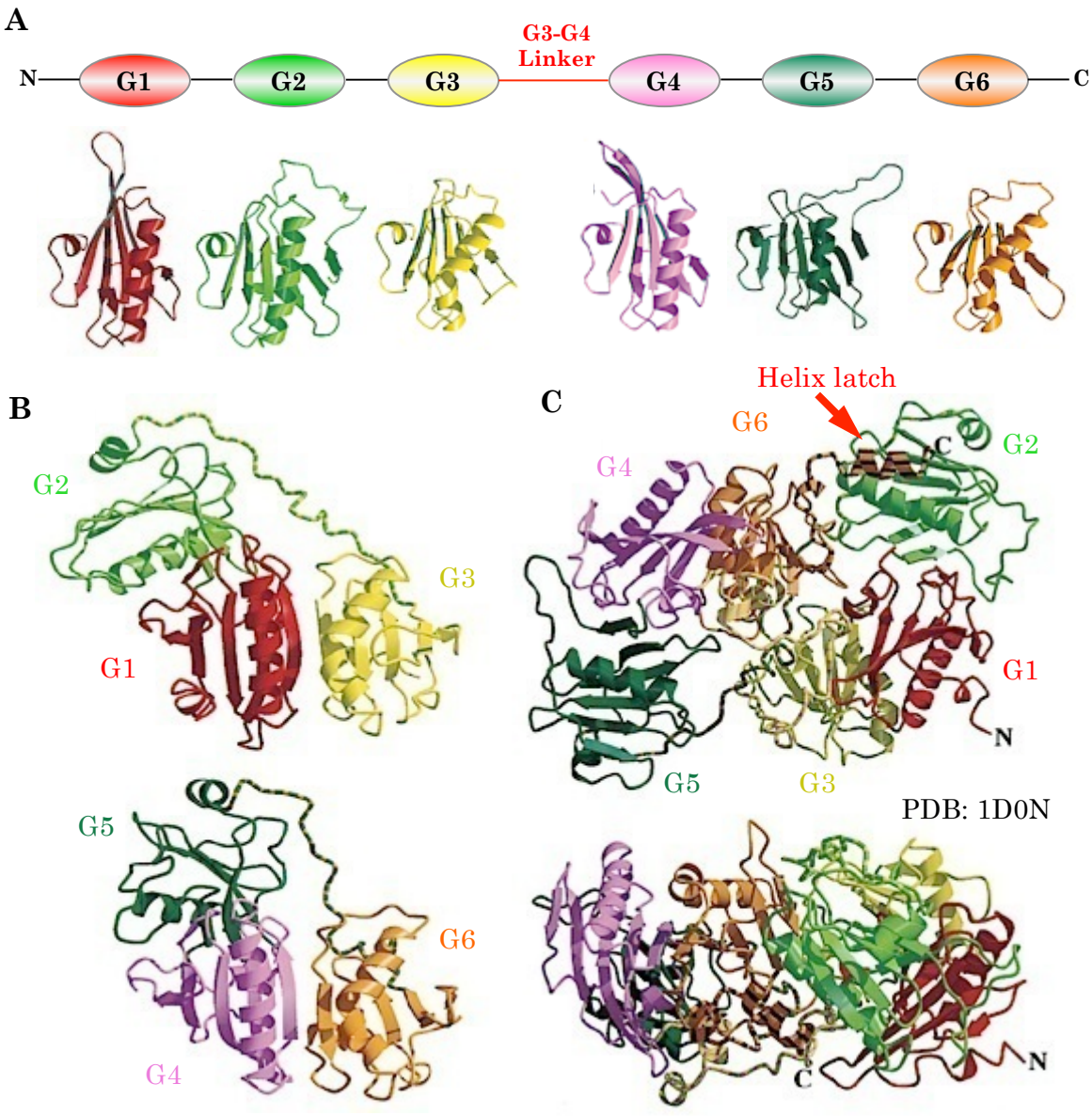

Figure 1.5 Gelsolin domains and the crystal structure of inactive full-length gelsolin. **(A)** Schematic diagram of the six domains of gelsolin highlighting the amino acid residues (top) and crystal structures (extracted from PDB code 1D0N) corresponding to each domain. **(B)** The crystal structure of the N-terminal (G1-G3) (top) and C-terminal halves (G4-G6) (bottom) extracted from the crystal structure of full-length gelsolin in (C). **(C)** Orthogonal views of the crystal structure of full-length gelsolin (PDB code 1D0N) (Burtnick *et al.*, 1997). The red arrow indicates the "helix latch" (a.k.a. "tail latch").

Once activated in the presence of an increased concentration of calcium ions, the two halves of gelsolin come free of each other to expose actin-binding surfaces, but remain joined by an extended 50-residue polypeptide, the G3-G4 linker. Radiolytic footprinting (a method in which solvent-accessible regions of proteins are modified by exposure to synchrotron irradiation) has been used to study the conformational change of gelsolin upon calcium activation (Kiselar *et al.*, 2003). Upon activation, residues in five regions of gelsolin become exposed: G1 (residues 49-72), G1 (residues 121-135), G2 (residues 162-166), G4 residues 431-454), and the "helix latch" (residues 722-748). Simultaneously, residues in domains G3 (276-300) and G6 (652-686) become less accessible, indicating they have moved into protected environments.

A three-stage transition in response to calcium activation was suggested: the first step in activation of gelsolin, which occurs at Ca^{2+} concentrations below 10 nM, involves a slight opening of the compact Ca^{2+}-free structure (Ashish *et al.*, 2007), followed at about 0.1 to 10 μM free Ca^{2+} by the opening of the compact shape to expose the actin binding sites (Kiselar *et al.*, 2003; Pope *et al.*, 1997; Ashish *et al.*, 2007). The last step in activation of gelsolin occurs at above 10 μM free Ca^{2+} where the structure further uncoils to reach a maximally open state at 200 μM free Ca^{2+} (Ashish *et al.*, 2007).

A high resolution structure of intact activated gelsolin is not available, but we have crystal structures for activated forms of the N- and C-terminal

halves of gelsolin (Burtnick *et al*., 2004; Robinson *et al*., 1999; Choe *et al*., 2002), each bound to a molecule of G-actin, and for an isolated Ca^{2+}-bound G4-G6 structure (Narayan *et al*., 2003; Kolappan *et al*., 2003) (Figs. 1.6 and 1.7). The structure of Ca^{2+}-bound G4-G6 emphasizes that activation involves large-scale movement of G6 away from G4, opening the G4-G6 latch to expose an actin-binding site on G4. The long helix in G6 that is kinked in inactive gelsolin becomes straight in the activated form. These observations are repeated in the structure of the G4-G6:actin complex (Fig.1.6), indicating that full activation of G4-G6 can be brought about by binding Ca^{2+} in the absence of actin. Furthermore, the structure of the complex shows the binding interface between G4 and actin to involve the long helix of G4 and the hydrophobic cleft of actin. There are minor contacts between G6 and actin, which had not been reported previously to have direct interactions. G5 does not contact actin, but forms a substantial bridge between G4 and G6.

The structures reported for activated G1-G3 bound to actin (Fig. 1.7) also show large-scale rearrangement of domain packing on activation of the N-terminal half of gelsolin, including the straightening of a kinked long helix on activation. However, these structures, when compared with the one for the G4-G6:actin complex (Fig 1.6B), emphasize how evolution of the two halves of gelsolin has diverged since the gene duplication event that led to their creation.

These structures identified two classes of Ca^{2+}-binding site on gelsolin. Type-I sites are found in G1 and G4, each of which shares coordination of one Ca^{2+} ion with actin (Robinson *et al*., 1999). Type-I Ca^{2+} ions mediate interactions at the interface between actin and an actin monomer-binding domain of gelsolin. One type-II calcium-binding site is entirely contained within each of gelsolin domains G1 through G6 (Choe, *et al*., 2002). These sites, having a wide spectrum of affinities for Ca^{2+}, are thought to induce conformational changes that are part of the overall activation process.

The structure of the N-terminal half of gelsolin, G1-G3, bound to G-actin (Fig. 1.7), has the long helix of G1 fitting into the hydrophobic cleft on actin, as seen for the long helix of G4 in G4-G6:actin (Robinson *et al*., 1999) and in the previously reported structure of isolated G1 bound to actin (McLaughlin *et al*, 1993). From this site, the G1-2 linker tightly associates with and extends up the surface of actin to position G2 to contact actin SD2. Finally, the G2–G3 linker coils up into a helix and positions G3 to bind back to actin SD1. The arrangement of G1–G3 in complex with actin differs strikingly from that observed for G4–G6 (Figs. 1.6 and 1.7).

This set of structures suggests a model for activation of gelsolin to bind to, sever and cap an actin filament. Activation involves opening three latches in gelsolin. The first, the “tail latch”, releases interactions between G6, along with its helical extension, and G2. This exposes the F-actin side-binding surface on G2. Opening the second latch, the “G1-G3 latch”, extends the G1-

G2 linker, allowing new contacts to be made along the surface of actin that direct G1 to its binding site on the actin unit corresponding to that observed in the G1-G3:actin structure (Figure 1.7; Irobi *et al*., 2003, Burtnick *et al*.,2004). Opening the last latch, the "G4-G6 latch", enables G4 to bind to an actin unit across the helix from, and in an analogous manner to, G1. This pincer movement of G1 and G4 severs the filament, leaving a newly generated filament barbed end capped by gelsolin. The substantial length of the G3-G4 linker polypeptide, 50 residues, permits gelsolin to span between two alternative pairs of actin units across the strand from each other in F-actin, one requiring an extension of 100 Å and a second of about 75 Å (McGough *et al*., 2003).

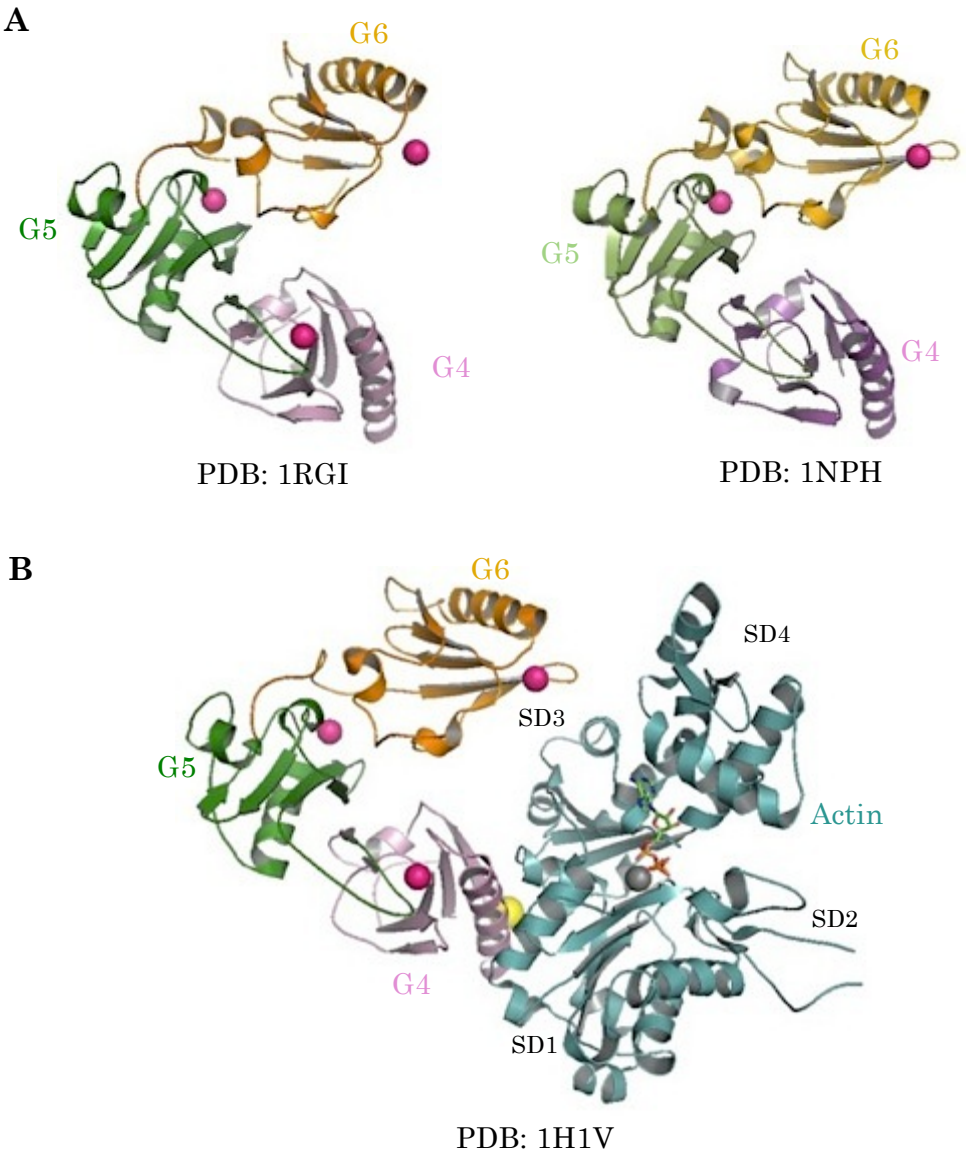

Figure 1.6 Crystal structures of the activated C-terminal half (G4-G6) of gelsolin. **(A)** Activated G4-G6 alone [PDB code 1P8X; Narayan *et al.*, 2003) and PDB code 1NPH; Kolappan *et al.*, 2003)]. **(B)** G4-G6:actin complex (PDB code 1H1V; Robinson *et al.*, 1999; Choe *et al.*, 2002). Pink and yellow spheres represent type-II and type-I Ca^{2+} ions, respectively. The Ca^{2+} ion associated with actin is presented as a grey sphere, with ATP in ball-and-stick format. Figures were prepared by PYMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC).

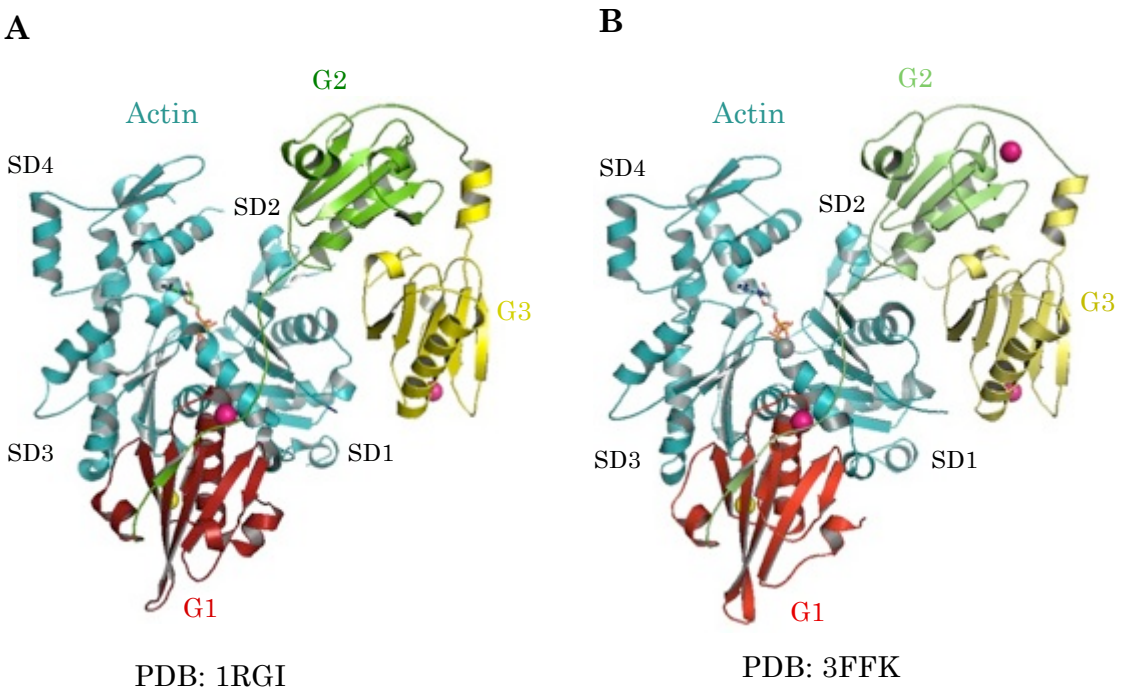

Figure 1.7 Crystal structures of the activated N-terminal half of gelsolin (G1-G3) bound to G-actin. **(A)** Equine G1-G3:actin (PDB code 1RGI; Burtnick *et al.*, 2004). **(B)** Human G1-G3:actin (PDB code 3FFK; Nag *et al.*, 2009). Pink and yellow spheres represent type-II and type-I Ca^{2+} ions, respectively. The Ca^{2+} ion associated with actin is presented as a grey sphere, with ATP in ball-and-stick format. Figures were prepared by PYMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC).

1.4 Nanobodies

1.4.1 What is a nanobody?

Hamers-Casterman *et al*., 1993, first described a unique antibody isotype that exhibits an extensive antigen-binding repertoire in the serum of *Camelus dromedarius*. These antibody molecules lack light chains and comprise only heavy-chain dimers, so are called Heavy-Chain Antibodies, or HCAbs, in contrast to conventional antibodies, which are composed of two identical H-chains and two identical L-chains (Fig. 1.8). An HCAb interacts with its antigen through a single variable domain, VHH (Fig. 1.8). VHH domains can be cloned and expressed easily in bacteria and are strictly monomeric, single-domain antigen-binding entities (Muyldermans and Lauwereys, 1999). In addition, VHHs are more soluble and more stable than fragments of antigen-binding regions of conventional antibodies (Davies and Riechmann, 1996; Arbabi Ghahroudi *et al*., 1997). Because of their prolate shape and dimensions (~2.5 nm in diameter, ~4 nm in length) and the publicity surrounding the subject of nanotechnology, Ablynx – a company developing therapeutic applications of camelid antibodies – has coined "nanobody" (Nb) as a new name for a VHH. Nbs are composed of three hypervariable regions, referred to as complementarity-determining regions CDR1, CDR2, and CDR3, which are surrounded by a framework of four conserved sequence regions. The CDR regions cluster at one end of the Nb in connecting loops between β-strands. The overall structure includes two β-

sheets, one four-stranded and one five-stranded, of an Ig fold (Muyldermans *et al.*, 2001; Padlan, 1994) (Fig. 1.8). Nbs tend to have hypervariable regions that are more extended than in conventional Abs, enabling better access to generally hidden epitopes, such as active sites of enzymes (Conrath *et al.*, 2001; De Genst *et al.*, 2006a; Decanniere *et al.*, 1999; Desmyter *et al.*, 1996). Crystal structures of Nbs show that Cys residues often found in CDR1 and CDR3 form a disulfide bond that is believed to assist in shaping the loop structure (Conrath *et al.*, 2003, De Genst *et al.*, 2006b; Muyldermans *et al.*, 2001). Furthermore, CDR2 regions contain more hydrophilic amino acids than found in conventional VH domains (Fig. 1.8B), which could explain the absence of associated light chains and the relative solubility of Nbs as single-domain entities (Nguyen *et al.*, 2001).

A crystal structure survey of dromedary VHH's in complex with lysozyme indicated that the main contributor in antigen binding of VHH's is the CDR3 loop, providing at least 60-80% of the contacts with antigen (De Genst *et al.*, 2006b). Consequently, the contribution of the CDR1 and CDR2 loops in antigen recognition has to be limited. In addition, because the VHH paratope has a convex shape that often is formed by the protruding long shape of the CDR3 loop (De Genst *et al.*, 2006b; Desmyter *et al.*, 1996), the dromedary VHH's preferably target antigen surfaces with concave topology, such as clefts that form active sites in enzymes (Lauwereys *et al.*, 1998).

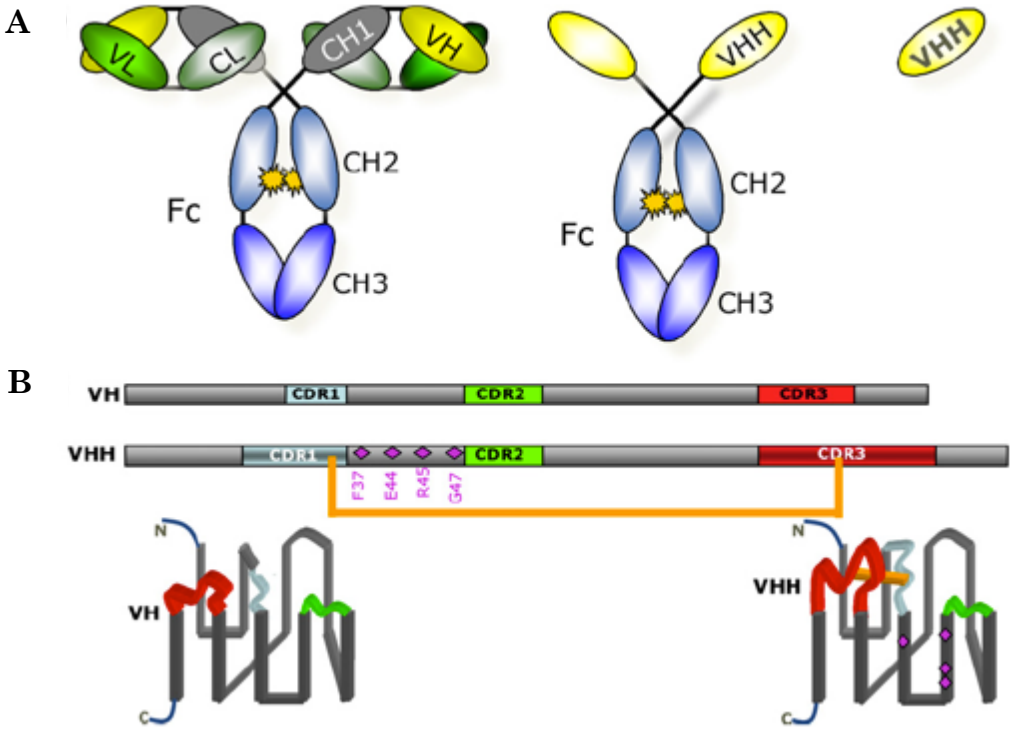

Figure 1.8 Comparison of conventional and heavy chain antibodies, and origin of a VHH domain, now called a nanobody. **(A)** From left to right, a classical antibody, a heavy chain antibody, and an isolated VHH domain or nanobody. **(B)** The sequence organization of VH and VHH domains, indicating the CDRs., below which are represented the folded structures of VH and VHH domains. (from Muyldermans *et al.*, 2009; reprinted with permission from Elsevier). F37 (or Phe37), E44 (or Glu44), R45 (or Arg45), G47 (or Gly47) refer to hydrophilic amino acids commonly found in VHH domains. CDR: complementarity-determining region. CH: constant domain of heavy-chain. CL: constant domain of light chain. Fc: fragment crystallizable region. VL: variable light-chain. VH: variable heavy-chain. VHH: variable heavy-chain of heavy-chain antibody.

1.4.2 Gelsolin-specific nanobodies (GsnVHHs)

Recombinant gelsolin-specific Nbs (GsnVHHs) have been produced and characterized (Van den Abbeele *et al*., 2010). GsnVHH9 and GsnVHH11 bind to G1-G2 or G2-G3 in the N-terminal half of gelsolin in a manner that is independent of Ca^{2+} (K_d ~5 nM), whereas GsnVHH3 and GsnVHH13 bind the C-terminal half (G4-G6) and specifically recognize the Ca^{2+}-activated form of Gelsolin (K_d ~10 nM). Furthermore, GsnVHH3 and GsnVHH9 slow down the degradation of gelsolin by contaminating proteases during up to at least eight days of incubation.

The crystal structures of GsnVHH3, 9, and 13 have been solved (Van den Abbeele *et al*., 2010) and superimposition of the three structures shows great similarity. Some differences do occur in the CDR3 regions (Fig. 1.9) at suspected contact surfaces for gelsolin.

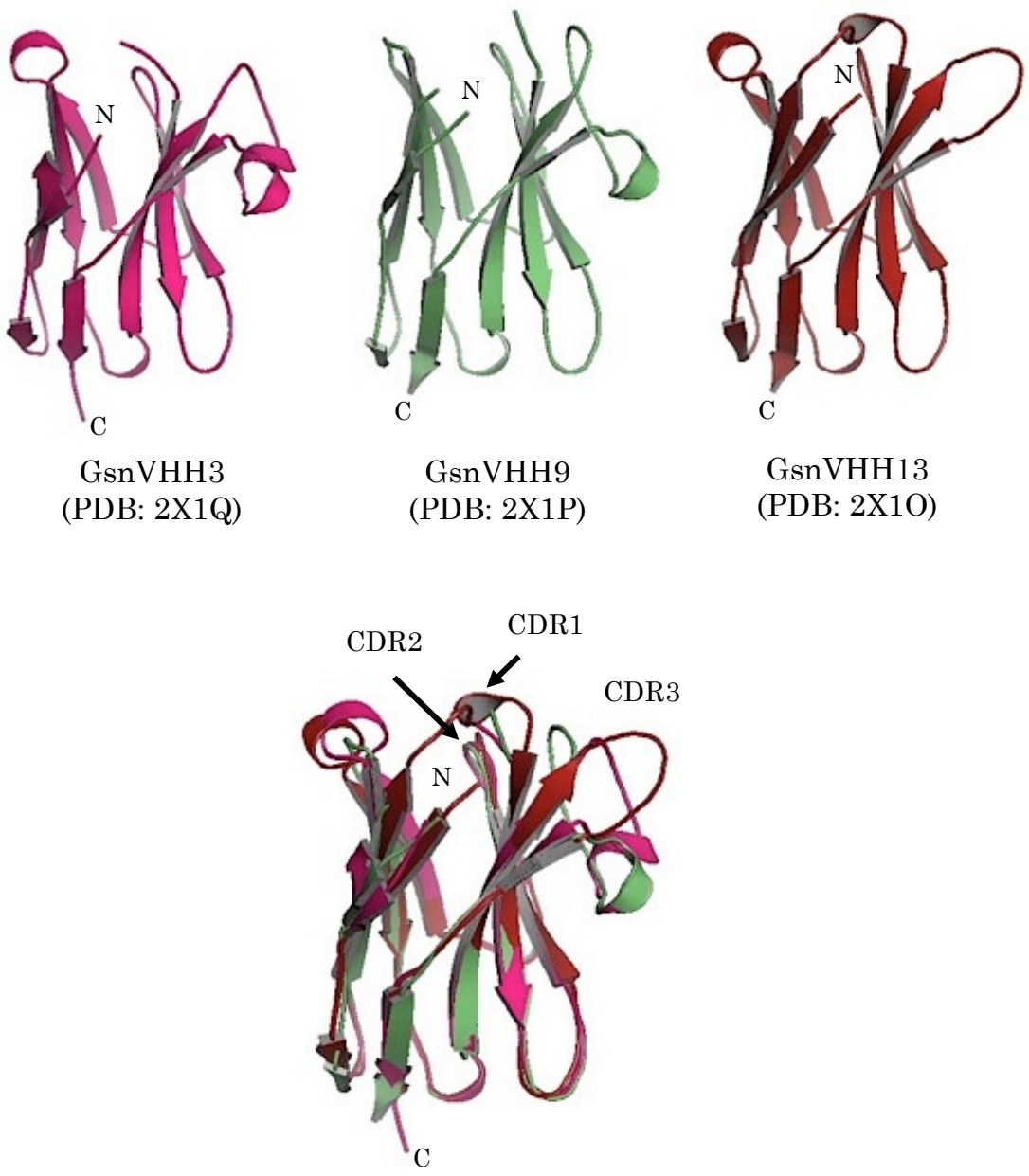

Figure 1.9 Crystal structures of gelsolin nanobodies (GsnVHHs): GsnVHH3 (PDB code 2X1Q), GsnVHH9 (PDB code 2X1P), GsnVHH13 (PDB code 2X1O), and their superimposition.

1.5 Work presented in this thesis

Although crystal structures of the N-terminal half of gelsolin in a complex with actin (G1-G3:actin) involving equine (Burtnick, *et al.*, 2004) and human (Nag, *et al.*, 2009) G1-G3, and of the activated form of the C-terminal half (G4-G6) of human (Narayan et al., 2003) and mouse (Kolappan *et al.*, 2003) gelsolin alone, or with actin bound (Robinson *et al.*, 1999; Choe *et al.*, 2002) have been solved previously, we sought to improve the resolution of the data by use of different crystallization materials and conditions. In addition, while the structures of Gelsolin-specific nanobodies 3, 9, and 13 (GsnVHH 3, 9, and 13) have been determined (Van den Abbeele *et al.*, 2010), that of Gelsolin-specific nanobody 11 (GsnVHH11) has not.

In this thesis, six new crystal structures corresponding to active fragments of human gelsolin and gelsolin-specific nanobodies are presented:

1) a higher resolution structure of the N-terminal half of human gelsolin in a complex with actin (G1-G3:actin),

2) a novel structure of the active form of isolated gelsolin domain 3 (active G3),

3) a structure of the active form of the C-terminal half of gelsolin (activated G4-G6) structure from different starting materials and crystallization conditions than employed previously,

4) and 5) two novel structures of GsnVHH11: GsnVHH11-F1 and GsnVHH-F2, and

6) a structure of GsnVHH9 obtained from different starting materials and crystallization conditions than employed previously.

Chapter 2

Experimental

2.1 Cloning and expression of gelsolin constructs

Three gelsolin constructs were cloned including: 1) full-length human gelsolin (FL-Gelsolin; contains human plasma gelsolin amino acid residues Met52 to Ala782), 2) Gel-NL (Met52 to His428): the N-terminal half of gelsolin (G1-G3) with part of the G3-G4 linker extension at its carboxyl-terminus, and 3) Gel-CL (Asn368 to Ala782): the C-terminal half of gelsolin (G4-G6) with part of the G3-G4 linker extension at its amino-terminus (Fig. 2.1).

The DNA fragments encoding the protein constructs (Fig. 2.1) were amplified by PCR using the primers listed in Table 2.1, and the cDNA encoding full-length human plasma gelsolin (GenBank accession No. CAA28000) was used as a PCR template. The PCR products then were cloned at *SfiI* and *ECoRI* sites into vector pSY5, a modified pET21d plasmid vector [obtained from the Robinson laboratory, Institute of Molecular and Cell Biology (IMCB), Agency for Science, Technology and Research (A*STAR), Singapore], which is designed to include an N-terminal eight-histidine tag (Hisx8 or His-tag) followed by a PreScission Protease cleavage site (PS) (Fig. 2.1).

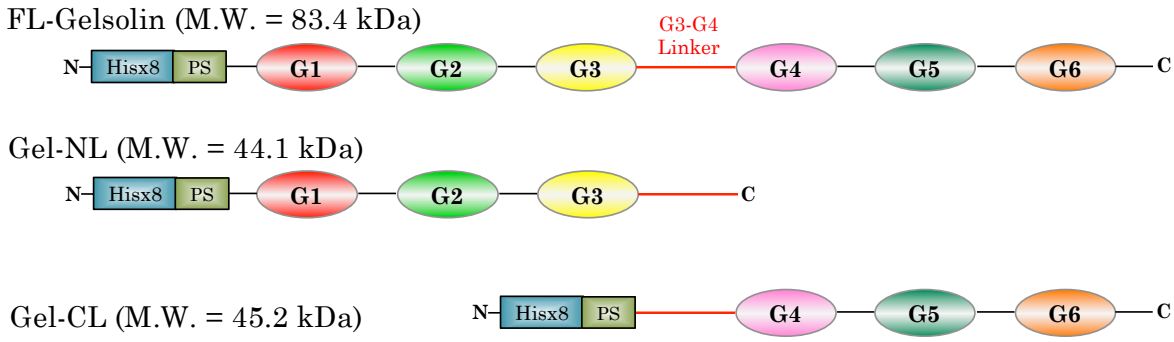

Figure 2.1 Protein constructs of gelsolin. The constructs include: 1) full-length gelsolin (FL-Gelsolin, contains human plasma gelsolin amino acid residues Met52 to Ala782), 2) Gel-NL (Met52 to His428): the N-terminal half of gelsolin (G1-G3) with part of G3-G4 linker extension to its carboxyl-terminus, and 3) Gel-CL (Asn368 to Ala782): the C-terminal of gelsolin (G4-G6) with part of G3-G4 linker extension to its amino-terminus. Hisx8: eight consecutive histidine residues (His-tag). PS: PreScission Protease cleavage site, containing residues LEVLFQGPGRP

The protein constructs were expressed in *E. coli* strains Rosetta (DE3) pLysS (Novagen) and BL21-Codonplus (Stratagene). Starter cells were grown in 250 mL shake-flasks at 37 °C overnight in 50 mL Luria Broth (LB) medium containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol as antibiotics. The starter cells then were inoculated and grown at 37 °C in 5 L

LB medium containing antibiotics in a 7-L fermenter until an OD_{600} of ~2.0 was reached. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) then was added to a final concentration of 1 mM to induce recombinant protein expression under control of the T7 promoter for 18 h at 25 °C. The cells were harvested by centrifugation at 4,500 rpm for 30 min using Beckman polypropylene bottles and an Avanti J-20 XP Beckman centrifuge. Pellets were resuspended in 150 mL of Ni^{2+}-binding buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 20 mM Imidazole) and stored in 5 tubes, each containing 30 mL (corresponding to 1 L of the original culture volume) at -80 °C.

2.2 Purification of gelsolin constructs

Cells containing gelsolin constructs from 1 L of cell culture were thawed and disrupted with a high-pressure homogenizer (Avestin Emulsiflex), followed by centrifugation at 18,000 rpm for 45 min in a Beckman 45 Ti rotor using a Beckman Optima L-90k ultracentrifuge to remove the cell debris. Supernatant, containing the soluble protein in Ni^{2+}-binding buffer, was loaded onto a 1 mL bed-volume Ni-NTA-affinity column and washed with at least 40 mL of Ni^{2+}-binding buffer to remove non-specifically bound protein. The protein then was eluted with 5 mL of Ni^{2+}-elution buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, and 300 mM Imidazole), and the eluted protein solution was applied to a gel filtration (GF) column (Bio-Rad Sephacryl S200; 1.6 x 60 cm) at room temperature and

eluted with GF buffer (50 mM Tris-HCl, pH 8.5 and 150 mM NaCl) at a rate of 1.2 mL/min using a BioLogic LP System (BioRad) equipped with a Kipp & Zonen paper chart recorder and model 2110 fraction collector (BioRad). Protein purity was analyzed by SDS-PAGE. Ten to fifteen μL of the 5-mL protein fractions were mixed with 5 μL of 5-X loading buffer (225 mM Tris-HCl, pH 6.8, 50%(v/v) glycerol, 5%(w/v) SDS, 0.05%(w/v) bromophenol blue, 250 mM DTT) and electrophoresed on a 12% SDS-polyacrylamide gel at 200 volts for 30-45 min. The gel was stained with staining solution [0.025%(w/v) Coomassie Brilliant Blue R250, 40%(v/v) methanol, and 7%(v/v) acetic acid] for 10 min, and destained in destaining solution [40%(v/v) methanol and 7%(v/v) acetic acid] overnight. Concentrations of purified proteins were determined by ultraviolet absorption measurements at 280 nm using a Perkin Elmer Lambda 4B UV-Vis spectrophotometer using extinction coefficients listed in Table 2.1, as calculated with the Peptide Property Calculator (based on Gill and von Hipple, 1989). Purified protein solutions were concentrated to 10 mg/mL in centrifugal concentrators (Amicon). EDTA was added to 2.0 mM to inhibit proteolysis.

Table 2.1 List of primers used in the PCR amplification, and of extinction coefficients at 280 nm used in calculation of concentrations of the gelsolin constructs.

Constructs	**Primer name**	**Primer sequence**	**Extinction coefficient [(mg/mL)$^{-1}$ cm^{-1}]**
Gelsolin	For-Gel1	5'-*GGG CCC GGG CGG CCG* ATG GTG GTG GAA CAC CCC GAG TTC-3'	1.36
	Rev-Gel6	5'-*GGA ATT CC*T TAG GCA GCC AGC TCA GCC ATG GCC CT-3'	
Gel-NL	For-Gel1	5'-*GGG CCC GGG CGG CCG* ATG GTG GTG GAA CAC CCC GAG TTC-3'	1.08
	Rev-Gel3L	5'-*GGA ATT CC*T TAG TGC AGG GTG GCG GCG TCG AAG GG-3'	
Gel-CL	For-Gel4L	5'-*GGG CCC GGG CGG CCG* AAC TGG CGG GAC CCA GAC CAG ACA-3'	1.60
	Rev-Gel6	5'-*GGA ATT CC*T TAG GCA GCC AGC TCA GCC ATG GCC CT-3'	

2.3 Preparation of rabbit muscle acetone powder

A modified method of Pardee and Spudich (1982) was used to prepare an acetone powder of rabbit (*Oryctolagus cuniculus*) skeletal muscle (Wang *et al*., 2010). Frozen young rabbit leg and back muscle tissues were purchased from Pel-Freez Biologicals. One kilogram of thawed rabbit muscle was minced with 3 - 4 L of cold buffer containing 300 mM KCl, 100 mM KH_2PO_4, 50 mM K_2HPO_4, 0.2 mM ATP, pH 6.5, in a metal blender to dissociate and solubilize myosin. The slurry was centrifuged at 7000 rpm using 1-L Beckman polypropylene bottles and an Avanti J-20 XP (Beckman) centrifuge for one hour, and the myosin-rich supernatant was discarded. The sediment, rich in actin and thin filament proteins, was then washed with 3 - 4 L of cold water for 20 min and centrifuged for 30 min and the pellets retained. To remove hemoglobin and myoglobin, the pellets were resuspended and washed with 3 - 4 L of cold 10 mM $NaHCO_3$, pH 8.0, for 20 min, then centrifuged for another 10 min. The sediment was again washed with 3 - 4 L of cold water for 10 to 20 min and centrifuged for 10 min. Initial dehydration of the pellets was accomplished by washing twice with 4 L of cold 95% ethanol, followed by centrifugation for 10 min. More complete dehydration was achieved by washing the residue twice with 4 L acetone and filtration using double-layered cheesecloth. The residue was spread over a sheet of filter paper in a fume hood for a few hours dry. The dried muscle powder was stored in 5 g lots in 50 mL Falcon tubes at -20 °C.

2.4 Purification of actin from acetone powder

Actin was purified from rabbit muscle acetone powder by the method of Spudich and Watt (1971), as subsequently modified (Wang *et al*., 2010).

Day1: Actin was extracted from 5 g of rabbit muscle acetone powder with 100 mL of ice-cold buffer A (2 mM Tris-HCl, 0.2 mM $CaCl_2$, 0.2 mM ATP, 1 mM DTT, pH 7.6, at 4 °C) for 30 min on ice. The extract was filtered through double-layered cheesecloth and filter paper prior to re-extraction with 50 mL of ice-cold buffer A and filtration. The filtrates were combined and clarified by centrifugation at 35,000 rpm in a Beckman 45 Ti rotor in a Beckman Optima L-90k ultracentrifuge for one hour at 4 °C. Then, KCl and $MgCl_2$ were added to 100 mM and 2 mM, respectively, to the pooled supernatant at 4 °C to polymerize actin overnight.

Day 2: To remove F-actin binding proteins such as tropomyosin, additional solid KCl was added up to 800 mM to the polymerized actin (F-actin) solution, and it was then stirred gently for 2 h at 4 °C. Then, the solution was centrifuged at 35,000 rpm in a Beckman 45 Ti rotor for 3 hours to pellet out F-actin from the solution. The pellet, containing F-actin, was homogenized in 10 mL of buffer A using a Potter homogenizer, after which the supernatant was dialyzed against three changes of 1 L of buffer A over a period of three days (*days 3*, *4*, and *5*).

Day 5: The resulting G-actin solution was centrifuged at 35,000 rpm in a Beckman 70.1 Ti rotor in a Beckman Optima L-90k ultracentrifuge at 4 °C

for 3 h to get rid of any remaining F-actin and other undissolved material.

To further purify G-actin, the supernatant was applied to a gel filtration column (Bio-Rad Sephacryl S200; 1.6 x 100 cm) at room temperature and eluted with buffer A at a rate of 1.2 mL/min using **a** BioLogic LP System (BioRad). Actin was eluted as a single peak, but only the second half of the peak was collected for subsequent use, in order to avoid possible contamination by actin dimers and short oligomers. The concentration of G-actin was determined by UV-Vis spectrophotometry (Perkin Elmer Lambda 4B) at 290 nm with an extinction coefficient of 0.63 $(mg/mL)^{-1}$ cm^{-1}.

2.5 Formation and purification of complexes between gelsolin constructs and actin

The concentrated gelsolin constructs at 10 mg/mL in GF buffer (50 mM Tris-HCl, pH 8.5 and 150 mM NaCl) were incubated with 5 mM $CaCl_2$ on ice for 5 min, then mixed to molar ratios of 1:1 and 1:2 with actin and incubated overnight at 4 °C. The protein complex solutions were run through a GF column (Bio-Rad Sephacryl S200; 1.6 x 100 cm) at room temperature, eluted with buffer A at a rate of 1.2 mL/min using a BioLogic LP System (BioRad), as previously described. The purity of each protein complex was characterized by SDS-PAGE as described in section 2.2 above. Protein fractions containing the complex were combined and their concentrations determined

spectrophotometrically at 280 nm using a weight-averaged extinction coefficient calculated for the proteins in the complex. The protein complex solutions then were concentrated to ~10 mg/mL using centrifugal concentrators (Amicon).

2.6 Sources of gelsolin G2-G4 and gelsolin-specific nanobodies: GsnVHH9 and GsnVHH11

The gelsolin domain G2-G4 was cloned and purified by Hui Wang, a former student in Burtnick laboratory (Wang, 2009).

The nanobodies: GsnVHH9 and GsnVHH11 were kindly provided by Sarah De Clercq of the Gettemans laboratory, VIB Department of Medical Protein Research, Ghent University, Belgium (Van den Abbeele *et al*., 2010).

2.7 Crystallization of the protein complexes

Hanging-drop vapour-diffusion crystallization trials (Bergfors, 1999) were conducted using both commercial and homemade screening kits. Commercial kits included: Hampton crystal screen HT, Hampton crystal screen 2, Index HT, Silver bullets and Silver bullets Bio (Hampton Research); Cryos, PACT, and ComPAS suites (Qiagen); and the Fluka Crystallization Kit for Protein Complexes. Homemade screens were based on Stura footprint screens 1 and 2 (Molecular Dimensions), or were of our own design aimed at performing finer screens. In addition, conditions were screened near those

that had previously been used successfully to grow crystals from protein solutions similar to those in the present study, *e.g.*, 2%(w/v) PEG8000, 100 mM sodium acetate, 2 mM $CaCl_2$, pH 4.7, as used for G1-G3:actin, and 10%(w/v) PEG8000, 20%(v/v) glycerol, 100 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer, pH 7.5, as used for G4-G6:actin. One or two microliters of protein solution were mixed with 1 μL of precipitant solution (well buffer) on a cover slip, which then was inverted onto the reservoir containing the well buffer, sealed, and incubated at 4 °C. Below are the details of crystallization screens used for particular protein complexes.

2.7.1 Crystallization of Gel-NL:actin and Gel-NL:2actin complexes

The GF-purified complexes of Gel-NL:actin and Gel-NL:2actin were crystallized using a well buffer of 1.0 - 6.4%(w/v) PEG8000, 100 - 320 mM NaOAc, 1.0 - 6.4 mM $CaCl_2$, pH 4.7 at 4 °C.

2.7.2 Crystallization of GA_2, a complex of one gelsolin with two actins

Crystals grew at 4 °C from a GF-purified protein complex containing one gelsolin plus two actins (GA_2) at 10 mg/ml by mixing it with the following reservoir solution and allowing vapour diffusion to proceed to equilibrium: 25%(w/v) PEG1500 and 100 mM MIB buffer (a buffer containing 25 mM malonic acid, 37.5 mM imidazole, and 37.5 mM boric acid), pH 4.0 – 6.0 or 100 mM SPG buffer (a buffer containing 12.5 mM succinic acid, 43.75 mM sodium dihydrogen phosphate, and 43.75 mM glycine), pH 7.0.

2.7.3 Crystallization of activated G2-G4

The GF-purified G2-G4 was activated by 2 mM $CaCl_2$ for 10 min before being subjected to crystallization against precipitant solution containing 0.1 M MES monohydrate (2-(N-morpholino)ethanesulfonic acid monohydrate) buffer, pH 6.5, 12%(w/v) PEG20,000 at 4 °C.

2.7.4 Crystallization of a complex of Gel-NL with GsnVHH11

A complex between Gel-NL and GsnVHH11, the latter added to a 2-fold molar excess over gelsolin, was prepared using purified proteins, mixed, and incubated at 4 °C overnight. Crystals grew at 4 °C from this protein solution, without further purification, when it was equilibrated against the

following well buffers: 1) 22.5%(w/v) PEG6000, 200 mM imidazole malate buffer, pH 8.5; and 2) 30-40%(w/v) PEG8000, 100 mM HEPES, pH 8.2.

2.7.5 Crystallization of Gelsolin:GsnVHH9 complex

A complex of gelsolin and GsnVHH9 (mole ratio of 1:2) was prepared by mixing the purified proteins and incubation overnight at 4 °C. Crystals were grown from protein solution, used without further purification, at 4 °C using the following well buffer: 33.0%(v/v) PEG600, 200 mM imidazole malate buffer, pH 5.5.

2.8 Actin polymerization and depolymerization assays

2.8.1 Pyrenyl-actin polymerization assay

G-actin (4 μM, 10% pyrene-labelled actin (prepared by reaction of N-(1-pyrenyl)iodoacetamide with actin residue Cys374 based on Kouyama and Mihashi, 1981) was incubated with a test protein (0.04 μM) in buffer A (2 mM Tris-HCl, 0.2 mM ATP, 0.5 mM DTT, 0.2 mM $CaCl_2$, pH 7.6) in a volume of 1.350 mL for 10 min. One hundred and fifty microliters of 10-X polymerization solution (1 M KCl and 20 mM $MgCl_2$) was added to each sample, and the fluorescence intensity at 407 nm was monitored using a Perkin Elmer LS55 Luminescence Spectrometer, with excitation at 365 nm.

2.8.2 Actin polymerization assay by light scattering

G-actin (4 μM) was pre-incubated for 10 min with a test protein (0.04 μM) in 1350 μL buffer A. To each test sample and a control containing only G-actin, 150 μL of 10-X polymerization buffer was added, with light scattering monitored at 90° to the incident beam using a Perkin Elmer LS55 Luminescence Spectrometer having excitation and emission wavelengths set to 360 and 364 nm, respectively, and spectral bandpasses set to 3.5 nm.

2.8.3 Actin depolymerization assay by light scattering

G-actin (2 μM) in buffer A was polymerized by incubation for 1 h after the addition of KCl to 100 mM and $MgCl_2$ to 2 mM. Test proteins then were added at a mole ratio of 1:20 with respect to actin. The final volume of each test solution was 1500 μL. Light scattering was measured as described for the actin polymerization assay.

2.9 X-ray diffraction data collection

Crystals were frozen in liquid nitrogen after being soaked in their respective precipitant solutions supplemented with cryoprotective additives: 15 - 20%(v/v) glycerol (Rodgers, 1994; Garman, 2003) or in some cases, 20%(w/v) trehalose (Pflugrath, 2004). X-ray diffraction data were collected using Macromolecular Crystallography Data Collector software (MxDC) on either of beamlines 08ID1 (Grochulski *et al.*, 2011) or 08B1-1 using Rayonix

MX300 CCD X-ray detectors at the Canadian Macromolecular Crystallography Facility, Canadian Light Source, Saskatoon, Canada. Data were collected regularly at these two beamlines after I was trained to do X-ray diffraction data collection and post-data processing both on-site and remotely during the First Annual CLS Mx Data Collection School held in May, 2011. Some crystals also were sent to our collaborator, Dr. R. Robinson, for X-ray diffraction data collection and processing with Blue-Ice 5 software at beamline BL13B, using an ADSC Quantum 315 CCD detector, at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The wavelength was set to ~1 Å, and the data collected at 105 K. Data were indexed, scaled, and merged in either Autoprocess (formerly known as AutoXDS) (Gonzalez and Tsai, 2010) or HKL2000 (Otwinowski and Minor, 1997).

2.10 Structure determination and refinement

Structural determination was carried out in either the PHENIX software suite (Adams *et al.*, 2010) or the CCP4 suite (Collaborative Computational Project, Number 4, 1994). Molecular replacement methods (Rossmann, 1972; Rossmann, 1990) using appropriate structures as search models were used in either Phenix_AutoMR (Adams *et al.*, 2010) or MOLREP (Vagin and Teplyakov, 1997; Vagin and Teplyakov, 2010). In some cases, AutoBuild (PHENIX suite, Adams *et al.*, 2010) was used to obtain the

initial models. The resulting models were subjected to rounds of rebuilding and refinement using Phenix.refine (Adams *et al.*, 2010) or REFMAC5 (Murshudov *et al.*, 1996; Murshudov *et al.*, 1997; Murshudov *et al.*, 2011). Electron density maps were displayed and model coordinates fitted using the interactive computer graphics program COOT (Emsley and Cowtan, 2004). RMSD values for α-carbon positions in superimposed polypeptide backbone structures were calculated using PYMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC).

The presence of salt bridges in protein structures was determined using ESBRI software (Evaluating the Salt BRIdges in proteins; Costantini *et al.*, 2008) that recognizes when at least one negatively charged side chain carboxyl oxygen atom of amino acid Asp or Glu, *i.e.*, Asp (OD) or Glu (OE), and a positively charged side chain nitrogen atom from one amino acid, Arg, Lys or His, *i.e.,* NH in Arg (NH), NZ in Lys (NZ or NE), are within a distance of 4.0 Å from each other (Kumar and Nussinov, 1999; Kumar and Nussinov, 2002).

2.11 *In silico* protein-protein docking

Protein docking runs were carried out using the Hex protein-docking program (Ritchie, 1998; Ritchie and Kemp, 1999; Ritchie and Kemp, 2000; Ritchie, 2003; Mustard and Ritchie, 2005; Ritchie, 2005; Ritchie *et al.*, 2008; Macindoe *et al.*, 2010; Ritchie and Venkatraman, 2010). Water molecules and

ions were removed from the structures of interest, which then were loaded as receptor and ligand. In orientation control mode, the receptor was held stationary while the ligand was subjected to docking motion searches under automatic control. In docking control mode, the following parameters were set: correlation type = both shape and electrostatic, FFT mode = 3D and post processing = MM minimization. Also, the following default parameters were used: Grid dimension = 0.6 Å, Receptor range = 180 Å, Ligand range = 180 Å, Twist range = 360^{o}, Distance range = 40 Å and scan step of 0.8 Å with steric scan of 18 (N= 18) and final search of 25 (N=25). Comparison and visualization of models were done using PYMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC.) and COOT (Emsley and Cowtan, 2004). Figures were prepared using PYMOL. The receptors and ligands used are described below.

2.11.1 Docking G2-G3 with GsnVHH11-F1 and GsnVHH11-F2

The G2-G3 fragment (residues Arg161 – Phe365) of gelsolin, extracted from the 2.8 Å resolution structure of human G1-G3:actin determined as part of this thesis project (see Chapter 5), was used as the receptor. GsnVHH11 (Mol 1; Fig. 6.2A) or GsnVHH11-F2 (Mol 1; Fig. 6.2B) was set as the ligand.

2.11.2 Docking G4-G6 with GsnVHH13

The activated G4-G6 structure (residues Asp414 to Trp741) presented in this thesis was used as the receptor, while the structure of GsnVHH13 (PDB code 2X1O) was used as the ligand.

Chapter 3

Protein Preparation and Evaluation

3.1 Purification

Gel filtration (GF) was used as a final step in the preparation of proteins suitable for crystallization. Fractions collected during elution were characterized with respect to proteins present and purity by SDS-PAGE.

3.1.1 Purification of rabbit skeletal muscle G-actin

G-actin (prepared from rabbit muscle powder, see sections 2.3 and 2.4) was eluted from a GF column at a rate of 1.2 mL/min with buffer A (2 mM Tris-HCl, 0.2 mM $CaCl_2$, 0.2 mM ATP, 1 mM DTT, pH 7.6, at 4 °C) as a single peak and collected as 5-mL fractions over the elution volume range between 70 and 110 mL. A ten to fifteen microliter sample of each G-actin fraction was subjected to analysis by SDS-PAGE (Fig. 3.1). The concentration of each G-actin fraction was determined spectrophotometrically using an extinction coefficient of 0.63 $(mg/mL)^{-1}$ cm^{-1} at 290 nm. In order to avoid possible contamination by actin dimers and short oligomers, only the actin fractions corresponding to the second half of the peak were pooled for subsequent use.

In Fig. 3.1, the SDS-PAGE gel of G-actin fractions shows two bands. The upper band at ~43 kDa corresponds to intact G-actin. G-actin in solution

is susceptible to digestion by trace amounts of contaminating proteases: e.g., trypsin cuts at residues Arg62 and Lys68, or chymotrypsin at residue Leu67, to yield core protein of ~34kDa that resists further degradation (Burtnick and Chan, 1980; Jacobson and Rosenbusch, 1976). Thus, the lower band often seen in SDS-PAGE analyses of actin is likely to be a proteolytically resistant core fragment.

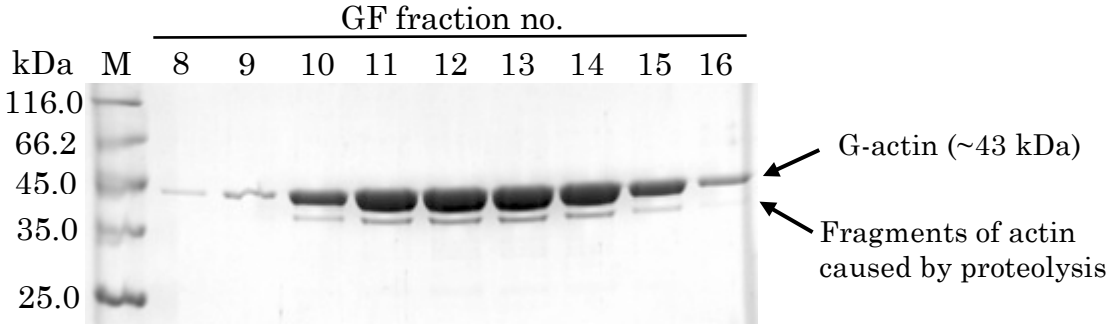

Figure 3.1 Example of an SDS-PAGE gel of G-actin fractions collected during GF purification. The lane labelled M denotes Unstained Protein Molecular Weight Markers (Fermentas), with their respective molar masses (kDa) shown to their left.

3.1.2 Purification of FL-Gelsolin, Gel-NL, and Gel-CL

Final purification of each of the constructs, FL-Gelsolin, Gel-NL and Gel-CL, was by GF chromatography at an elution rate of 1.2 mL/min with 50 mM Tris-HCl, pH 8.5, and 150 mM NaCl at room temperature. The eluate in each case was monitored spectrophotometrically, collected in 5-mL fractions and analyzed by SDS-PAGE (Fig. 3.2).

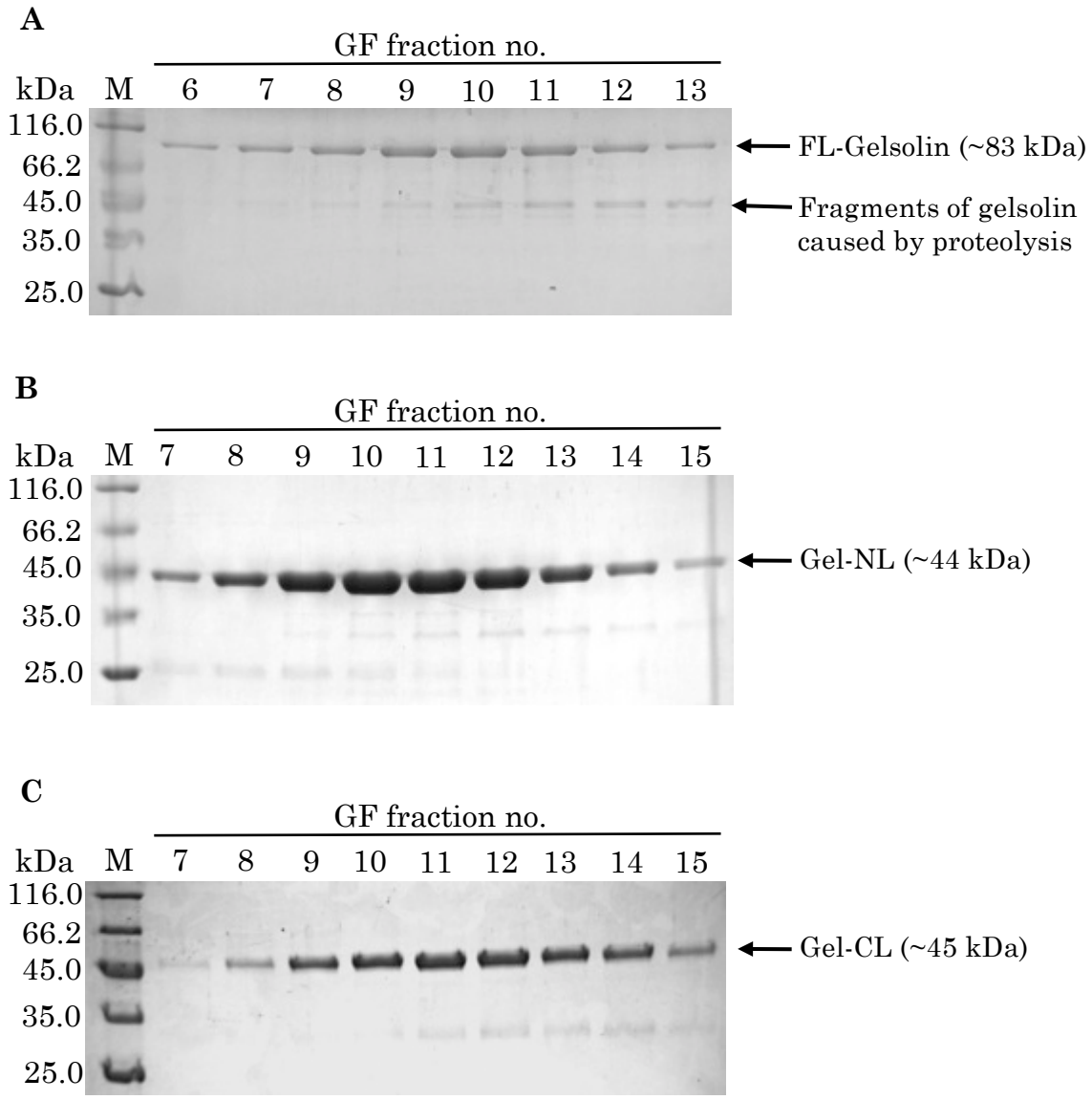

Figure 3.2 Example SDS-PAGE gels of GF-purified FL-Gelsolin, Gel-NL and Gel-CL. **(A)** FL-Gelsolin: fractions 6 - 13. **(B)** Gel-NL: fractions 7 - 15. **(C)** Gel-CL: fractions 7 – 15. M denotes a lane containing Unstained Protein Molecular Weight Markers (Fermentas), with their respective molar masses (kDa) shown to their left.

SDS-PAGE of GF-purified FL-Gelsolin shows two bands (Fig. 3.2A), a mixture of FL-Gelsolin and protease-digested products. The upper band at ~83 kDa corresponds to intact FL-Gelsolin. Gelsolin is susceptible to a variety of proteases that cut it in the G3-G4 linker peptide, which is most exposed in the calcium-activated protein, *e.g.*, the cysteinyl protease Caspase-3 cuts at residue 376 (Kothakota *et al.*, 1997; Geng *et al.*, 1998). The lower band in the gels, at ~44 kDa, corresponds to a half-gelsolin, either the N- or C-terminal half.

The intense bands in Figs. 3.2B and 3.2C correspond to GF-purified Gel-NL and Gel-CL, respectively. The faint bands correspond to contaminating peptides of lower molar mass.

Fractions of the GF-purified individual proteins were pooled and concentrated to 10 mg/mL for subsequent use. EDTA was added to 2.0 mM to inhibit proteolysis.

3.1.3 Purification of an FL-Gelsolin:2actin complex (GA_2)

G-actin in buffer A was added into a solution to a 2-fold molar excess over Ca^{2+}-activated FL-Gelsolin and incubated overnight at 4 °C. The mixture was subjected to gel filtration (Bio-Rad Sephacryl S200; 1.6 x 100 cm) using buffer A as eluent. Fractions (5 mL) were collected and analyzed by SDS-PAGE (Fig. 3.3A).

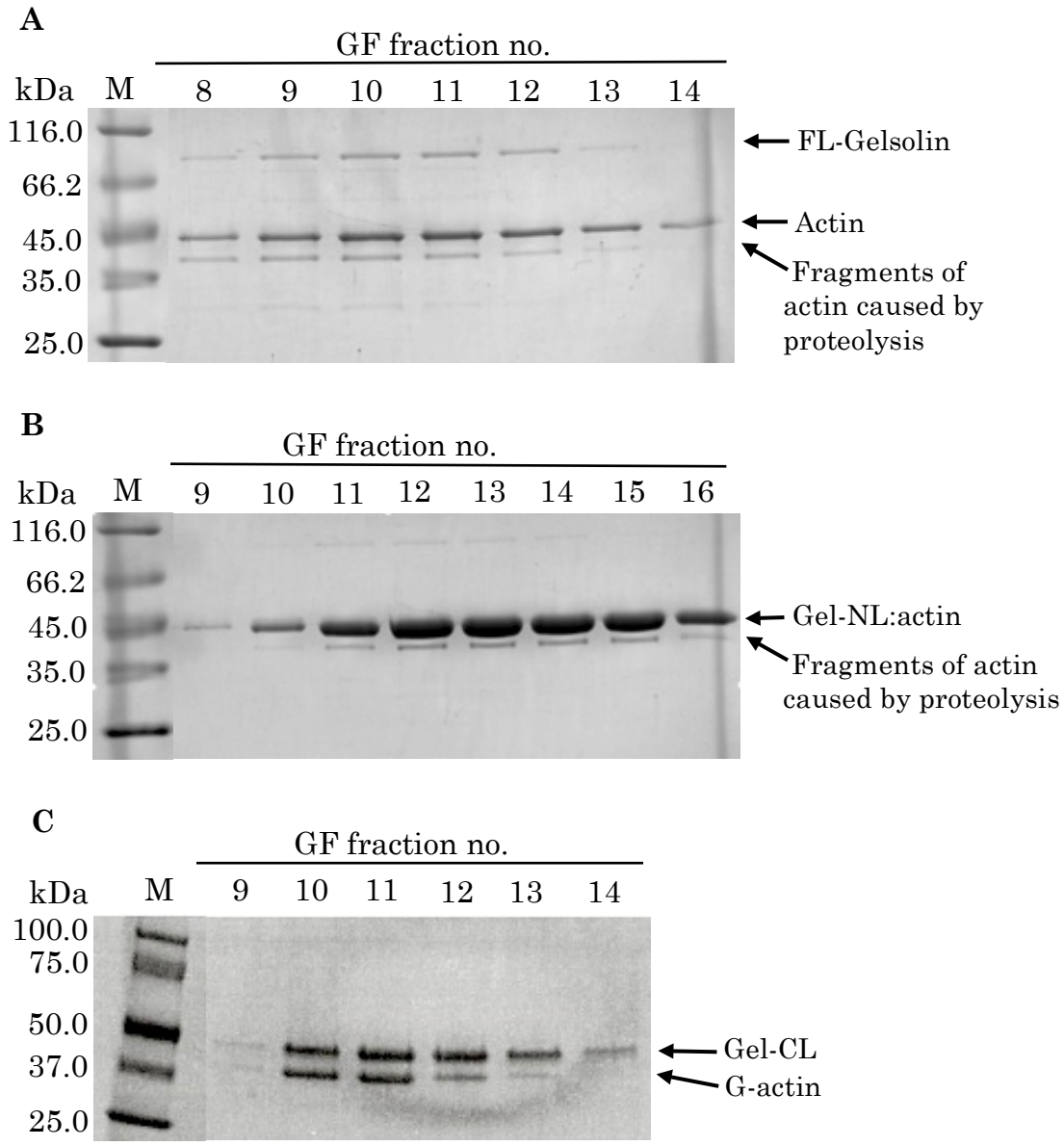

Figure 3.3 Example SDS-PAGE gels of GF-purified GA_2, Gel-NL:actin, and Gel-CL:actin complexes. **(A)** GA_2 complex: fractions 8 – 14. **(B)** Gel-NL:actin complex: fractions 9 – 16. **(C)** Gel-CL:actin complex: fractions 9 – 14. M in (A) and (B) denotes lanes with Unstained Protein Molecular Weight Markers (Fermentas). M in (C) denotes Bio-Rad Precision Plus Protein dual color Standards (Bio-Rad). The respective molar masses (kDa) of the makers are shown to their left.

3.1.4 Purification of Gel-NL:actin and Gel-NL:2actin complexes

Activated purified Gel-NL was used to make complexes with actin at mole ratios of 1 Gel-NL to 1 or 2 actins, denoted Gel-NL:actin and Gel-NL:2actin, respectively (as described in Chapter 2). Figure 3.3B shows an example of SDS-PAGE gel of the Gel-NL:actin complex fractions 9 - 16 after GF-purification. One major band was observed for each GF fraction, likely due to the similarity in molar masses of actin and Gel-NL (~43 and ~44 kDa, respectively), reflecting low resolving power in the gels used. The lighter staining band below the main one results from proteolysis of actin, as described in section 3.1.1. GF-purified fractions 9 – 16 were pooled and concentrated to 10 mg/mL and used in crystallization screens. SDS-PAGE gels of GF-treated Gel-NL:2actin complexes showed similar results.

3.1.5 Purification of Gel-CL:actin complex

The construct Gel-CL, the C-terminal half of gelsolin with a G3-G4 polypeptide linker, was used to form a complex with actin at a mole ratio of 1 to 1, denoted Gel-CL:actin (as described in Chapter 2). The complex was isolated by GF-chromatography as above and confirmed to contain both proteins by SDS-PAGE (Fig. 3.3C).

3.2 Activity assays

Gelsolin regulates actin assembly and disassembly through its abilities to bind, sever, cap, and nucleate actin filaments (Fig. 3.4; McGough *et al.*, 2003). The severing activity of gelsolin requires activation of the resting structure to expose F-actin side-binding sites on G2, which initiates contact with a filament. Next, the G1-G2 linker extends and binds to the surface of actin to position G1 near its binding site at the hydrophobic cleft, where it inserts itself between two longitudinally arranged actin subunits. Meanwhile, the G3-G4 linker wraps around F-actin to position G4 near to its binding site at the hydrophobic cleft of an actin unit across the filament from the one to which the N-terminal half of gelsolin is anchored by G2. Insertion of G4 into its binding site concludes a pincer movement that severs the filament, leaving the newly generated barbed end capped with respect to elongation or annealing with the pointed end of a nearby filament (Fig. 3.4A; Harris and Weeds, 1984; Coue and Korn, 1985; Sun *et al.*, 1994; Robinson *et al.*, 1999; Hans *et al.*, 2003; Burtnick *et al.*, 2004). Interestingly, fragment G1-G3 on its own can sever F-actin, though not quite as efficiently as full-length gelsolin, but in a manner that is independent of Ca^{2+} (Chaponnier *et al.*, 1986; Kwiatkowski *et al.*, 1989; Way *et al.*, 1989; Yin *et al.*, 1990). However, unlike intact gelsolin, G1-G3 is reportedly unable to nucleate actin polymerization (Way *et al.*, 1989), a property shared with G4-G6, the C-terminal half of gelsolin (Bryan and Hwo, 1986; Chaponnier *et al.*, 1986).

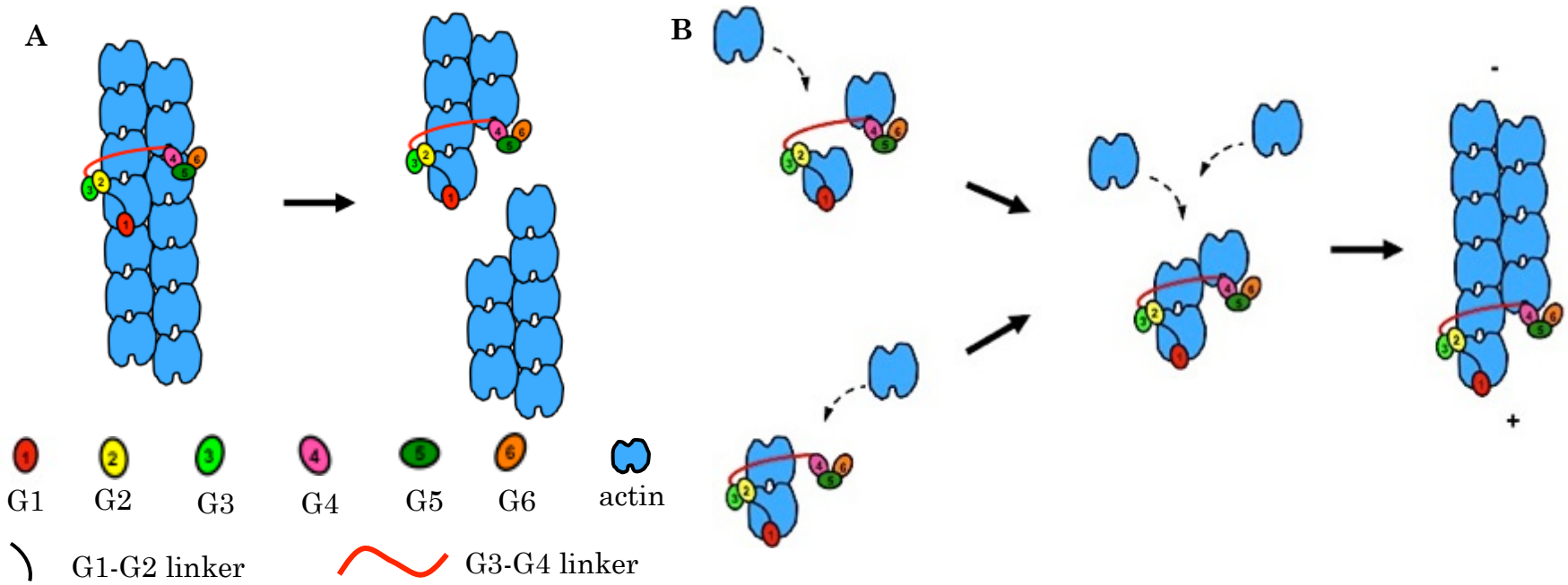

Figure 3.4 Cartoon models for severing, capping and pointed-end nucleation activities of gelsolin. **(A)** Model for severing and capping of F-actin by gelsolin. First, G2 binds to the side of F-actin. Then, the G1-G2 linker directs G1 to its binding site on the same actin monomer, while the G3-G4 linker wraps around F-actin, directing G4 to its binding site across the filament from G2. The binding of G1 and G4 to their respective sites severs the filament, leaving the newly created barbed end capped. **(B)** Model for nucleation of actin pointed-end elongation by full-length gelsolin. Two ways for assembling three actin units by gelsolin into a stable nucleus for pointed-end growth are presented. The + and – signs indicate the barbed (or plus) and pointed (or minus) ends, respectively.

A goal of the present work is to add atomic detail to this model presented above (Fig. 3.4). To ensure that the recombinant proteins that are used in these studies are properly folded and capable of exhibiting gelsolin-like activities, they were tested in well-established assays of actin polymerization and depolymerization. Specifically, a fluorescence assay that incorporates pyrenyl-actin and a light scattering assay that is sensitive to filament length to test FL-Gelsolin, Gel-NL, and Gel-CL, were employed as described in section 2.8 of this thesis.

3.2.1 Actin polymerization assays of the gelsolin constructs

In actin polymerization assays, test proteins (to 0.04 μM) plus 10-X polymerization buffer were added simultaneously to 4 μM G-actin solutions, which then were monitored in fluorescence or light scattering assays. Polymerization resulted in increased signal strength in both assays. In a control where 10-X polymerization buffer alone was added to the G-actin solution, a slow rise in signal reflects the nucleation phase of polymerization. In the presence of a test protein that nucleates polymerization, the signal rises more rapidly (Figs. 3.5A and 3.5B). As expected, FL-gelsolin nucleates actin polymerization, whereas the C-terminal fragment of gelsolin (Gel-CL) does not. In contrast to expectation (Way *et al.*, 1989), the N-terminal fragment (Gel-NL) exhibits nucleating activity.

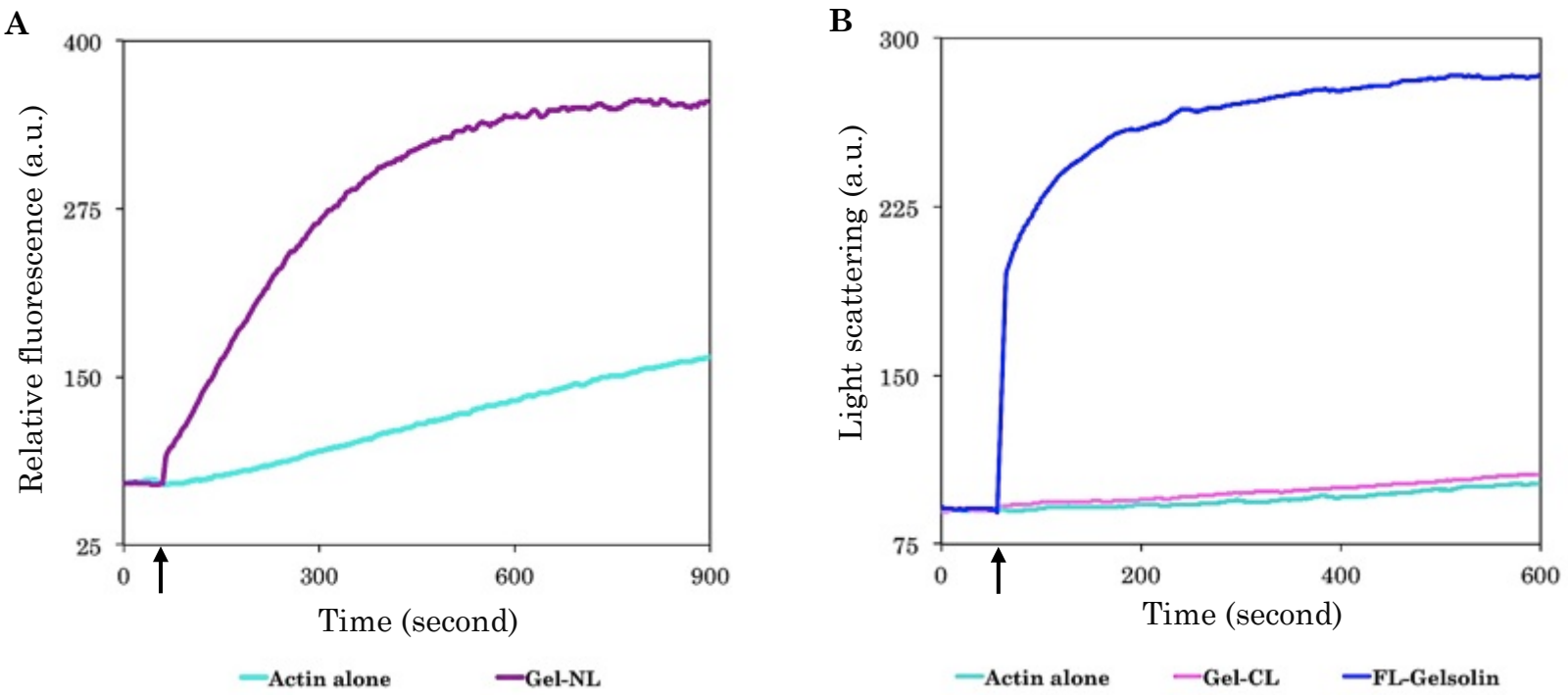

Figure 3.5 Assays of actin polymerization of gelsolin constructs: FL-Gelsolin, Gel-NL, and Gel-CL. **(A)** Fluorescence assays of the effects of Gel-NL on actin polymerization (with pyrenyl-actin at 10 mole% of the total 4 μM actin) **(B)** Light scattering assays of the effects of FL-gelsolin and Gel-CL on polymerization of 4 μM G-actin. In the control run (actin alone), buffer A without additional protein was added. Actin at a final concentration of 4 μM was 100-fold in molar excess over the test proteins used. Arrows indicate the time at which the test solutions and 10-X polymerization buffer simultaneously were added to the G-actin solutions in buffer A.

3.2.2 Actin depolymerization assays of the gelsolin constructs

F-actin solutions were prepared from G-actin at 2.0 μM in buffer A by addition of KCl and $MgCl_2$ to final concentrations of 100 mM and 2 mM, respectively (Fig. 3.6). These provided baseline light scattering intensities. Next, each test protein was activated by the addition of $CaCl_2$ to 2 mM, and then added to a final concentration of 0.1 μM to a sample containing F-actin. Severing activity in a test sample produced a sudden drop in light scattering intensity, the result of F-actin being cut into shorter length polymers and oligomers.

When tested (Fig. 3.6), activated Gel-CL was unable to sever F-actin, as previously reported (Bryan and Hwo, 1986; Chaponnier *et al*., 1986). Also as expected (Kwiatkowski *et al*., 1989; Yin *et al*., 1990; Way *et al*., 1989; Chaponnier *et al*., 1986), FL-gelsolin and Gel-NL were able to induce a sudden drop in light scattering intensity, to nearly the same extent.

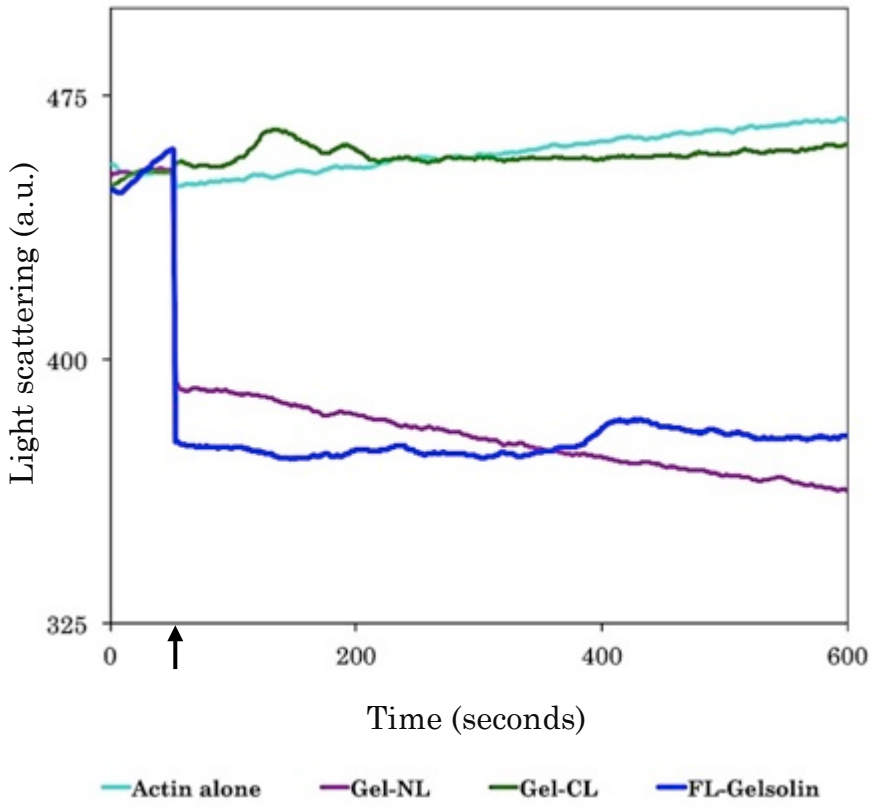

Figure 3.6 Light scattering assays of actin depolymerization by FL-Gelsolin, Gel-NL and Gel-CL. The following additions were made, at a time indicated by the arrow, to a solution of F-actin (2 μM) in polymerization buffer, with test proteins at a mole ratio of 1:20 with respect to actin: Buffer A (actin alone), FL-gelsolin, Gel-NL, and Gel-CL.

3.3 Discussion

Recombinant FL-gelsolin and two fragments, Gel-NL and Gel-CL, were prepared and purified for subsequent crystal growth trials in the presence of actin. Gel-NL exhibits severing activity but Gel-CL does not, which is in agreement with previous reports for the N-terminal half (G1-G3) and the C-terminal half of gelsolin (G4-G6), respectively (Kwiatkowski *et al*., 1989; Yin *et al*., 1990; Way *et al*., 1989; Chaponnier *et al*., 1986). Gel-CL, as expected (Bryan and Hwo, 1986; Chaponnier *et al*., 1986), was unable to nucleate actin polymerization. However, in contrast to expectations (Way *et al*., 1989), Gel-NL demonstrated nucleation. As Gel-NL comprises G1-G3 plus the G3-G4 linking polypeptide, we can suggest that it is the presence of the ~50-residue linker sequence that confers nucleation activity to the N-terminal half of gelsolin.

These constructs, together with natural source rabbit muscle actin or with recombinant anti-gelsolin nanobodies, were used to prepare various protein complexes to subject to crystallization screening trials.

Chapter 4

A New Look at Activated Domain Structures in Gelsolin: The N-terminal Half

4.1 Background

In the inactive states of both equine (Burtnick *et al*., 1997) and human gelsolin (Nag *et al*, 2009), each protein domain adopts a typical gelsolin fold: a five-stranded core β-sheet sandwiched between two α-helices, one long and one short, that lie approximately parallel and perpendicular, respectively, to the strands of the sheet. The six domains are organized into a compact globular structure that is held together by extensive noncovalent interactions. Within this globule, most of the contacts between the pseudosymmetrical halves of intact gelsolin involve domain G2 and domain G6, along with its helical tail extension. In this manner, the primary actin-binding sites in G1, G2, and G4 are buried (Burtnick *et al*., 1997; Nag *et al*., 2009).

Examination of the electrostatic surface of inactive (Ca^{2+}-free) human gelsolin reveals two distinct charged patches (Nag *et al*., 2009). A positively

charged area has been associated with the binding of the triphosphate moiety of ATP and is speculated to be the binding site for other polyphosphate molecules, such as PIP2 (Urosev *et al.*, 2006). A negatively charged region is located at the junction of G2 and G6. It involves two unoccupied type-II Ca^{2+}-binding sites (Nag *et al.*, 2009). Binding of Ca^{2+} to these sites is proposed to be interdependent and cooperative with regard to disrupting the G2–G6 interface that is dependent on electrostatic interactions between residues from both vacant Ca^{2+}-binding sites (Nag *et al.*, 2009). Specifically, in the Ca^{2+}-free state, Arg168 and Arg169, respectively, from G2 interact with Asp669 and Asp670 from the G6 Ca^{2+}-binding site, while Arg168 also binds to Glu209 and Asn206 from the G2 Ca^{2+}-binding site. In addition, interactions of Asp744 and Asp747 with Arg207 tether the helical tail extension of G6 across the F-actin-binding site on G2 (Nag *et al.*, 2009)

During activation of gelsolin, Ca^{2+} binding disrupts interactions between G2 and G6, releases the helical tail latch, and induces large-scale rearrangement of the domains (Choe *et al.*, 2002; Kolappan *et al.*, 2003; Narayan *et al.*, 2003; Burtnick *et al.*, 2004). As part of this process, domains G3 and G6, respectively, undergo rotations of ~90° and translations of ~40 Å relative to G1 and G4 and form new contacts with G2 and G5 (Choe *et al.*, 2002; Burtnick *et al.*, 2004). Activated actin-bound gelsolin has been suggested to contain as many as eight bound Ca^{2+} ions (Choe *et al.*, 2002; Kazmirski *et al.*, 2002; Burtnick *et al.*, 2004; Chumnarnsilpa *et al.*, 2006): two

at type-I Ca^{2+}-binding sites associated with G1 and G4, with Ca^{2+} coordinated by amino acid residues from both actin and gelsolin (Choe *et al.*, 2002), and six at type-I Ca^{2+}-binding sites, one in each of G1 through G6, with Ca^{2+} coordinated by residues originating only from gelsolin (Choe *et al.*, 2002).

Intriguingly, the 3-Å resolution crystal structure of the N-terminal half of equine gelsolin in a complex with actin (equine G1-G3:actin; Burtnick *et al.*, 2004) has a vacant type-II Ca^{2+}-site in G2, while the analogous site in the 3-Å resolution structure of human G1-G3:actin is occupied (Nag *et al.*, 2009). That this site in equine G2 is a valid metal ion-binding site is demonstrated by soaking crystals of equine G1-G3:actin in solutions containing $TbCl_3$ and observing it to be occupied by Tb^{3+} (Chumnarnsilpa *et al.*, 2006).

Comparison of the structure of human G1-G3:actin (Nag *et al.*, 2009) with that of intact inactive gelsolin (Burtnick *et al.*, 1997) shows that binding Ca^{2+} at the type-II Ca^{2+}-binding site in G2 results in Glu209 releasing and moving away from Arg168 towards Asp187, taking Asn206 with it. This results in increased mobility of Arg168 and Arg169, weakening the contacts between G2 and G6. Minor straightening of the G2 long helix, to which Asn206 and Glu209 are attached, shifts the position of Arg-207, which interacts directly with the helical tail in inactive gelsolin. Through this domino effect, Ca^{2+} binding by G2 may drive release of the tail latch during activation.

Inclusion of the carboxylate group of Asp259 in the coordination sphere of the Ca^{2+} bound to G2 induces the previously unordered G2-G3 linker peptide to coil into an α-helix. A helical G2-G3 linker is not compatible with the position of G6 as it exists in the structure of inactive gelsolin, introducing a steric element into the activation process.

In summary, Ca^{2+}-binding by G2 drives the initial stages of activation of gelsolin in three ways: release of contacts to the helical tail latch, disruption of the G2/G6 interface, and promotion of the folding of the G2-G3 linker into a form that is incompatible with association between G2 and G6 (Nag *et al.*, 2009).

This thesis reports on two structures that are relevant to the discussion above. The first is a higher resolution (2.8 Å) structure of the N-terminal half of human gelsolin in a complex with actin that resulted from use of different crystallization materials and conditions from those reported previously (Burtnick *et al.*, 2004; Nag *et al.*, 2009). The second is of an activated form of isolated human gelsolin domain G3, which has not previously been published.

4.2 Higher resolution structure of human G1-G3:actin

The crystal structure of human G1-G3:actin at a resolution of 2.8 Å was achieved using crystals grown from solutions of complexes of one Gel-NL with one or two actins (Gel-NL:actin and Gel-NL:2actin, respectively).

Various crystal forms were obtained and subjected to X-ray diffraction analysis.

4.2.1 Crystallization

Trials with commercial and homemade screens did not result in single crystals suitable for X-ray diffraction analysis. However, single crystals grew from solutions of both Gel-NL:actin and Gel-NL:2actin under conditions similar to that used for crystallization of a complex between equine plasma gelsolin and actin [2%(w/v) PEG8000, 100 mM NaOAc, 2 mM $CaCl_2$, pH 4.7, at 4 ^oC (Burtnick *et al.*, 2004)]. Crystals grown from a solution of Gel-NL:actin that diffracted to 2.8 Å resolution were grown against a well buffer of 6.4%(w/v) PEG8000, 320 mM NaOAc, 6.4 mM $CaCl_2$, pH 4.7, while crystals of Gel-NL:2actin were grown against 2.1%(w/v) PEG8000, 104 mM NaOAc, 2.1 mM $CaCl_2$, pH 4.7. These precipitant solutions are somewhat different from the previously reported crystallization condition for human plasma G1-G3 bound to actin [9%(w/v) PEG4000, 100 mM $Ca(OAc)_2$, 100 mM NaOAc, pH 4.6 (Nag *et al.*, 2009)], possibly because crystals of either equine or human G1-G3 in a complex with actin were grown using full-length gelsolin rather than Gel-NL, which contains only the amino-terminal half of gelsolin.

Crystals grown from solution of either Gel-NL:actin or Gel-NL:2actin grew into a mix of three different forms in the same well (Fig. 4.1) While crystals grown from solutions of Gel-NL:actin were most likely to take on the

crystal form labeled GNL.A-1 (Fig. 4.1), those grown from solutions of Gel-NL:2actin were more likely to take on either of forms GNL.A-2 or GNL.A-3 (Fig 4.1).

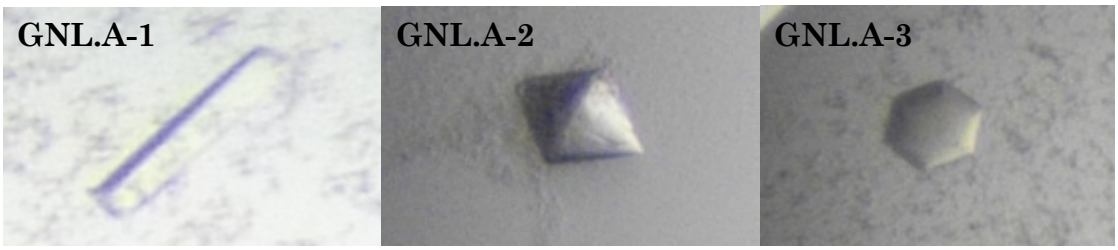

Figure 4.1 Examples of crystals grown from pooled fractions after GF-purification of Gel-NL:actin and Gel-NL:2actin complexes.

4.2.2 Data collection and refinement

Ten X-ray diffraction datasets were collected at different resolutions from ten crystals that were cryoprotected using 20% glycerol (Appendix A: Tables A.1 and A.2). All crystals belonged to space group $P2_12_12_1$ and had nearly identical unit cell parameters a = 101.50 Å, b = 147.95 Å, c = 146.96 Å, and $\alpha = \beta = \gamma = 90.0°$ (values calculated from the two highest resolution datasets, both to 2.8-Å resolution, one for a crystal of type GNL.A-1 that grew in a solution of Gel-NL:actin and the second for a crystal of type GNL.A-2 that grew in a solution of Gel-NL:2actin). These values are similar to those for human G1-G3:actin crystals (Nag *et al.*, 2009). The structures were solved by molecular replacement (Rossmann 1972; Rossmann 1990), using the

structure of human G1-G3:actin (PDB code 3FFK) as a search model in MOLREP (Vagin and Teplyakov 1997; Vagin and Teplyakov 2010) or in Auto_MR (Adams *et al*., 2010). Data collection and refinement statistics in Table 4.1 are for 2.8-Å resolution data using a crystal that grew in a solution of Gel-NL:actin. The crystallographic asymmetric unit in both crystal types GNL.A-1 and GNL.A-2 contained two 1:1 molar complexes of G1-G3 with actin, *i.e.,* G1-G3:actin. The G3-G4 polypeptide linker appended to the C-terminus of G3 in Gel-NL was not evident in the final refined model of either of the two complexes, Mol 1 and Mol 2, in the asymmetric unit (Fig. 4.2). Hence each structure is of a G1-G3:actin complex. Possible reasons include that the linker is flexible due to a lack of interaction with the rest of the structure, or that it has been removed due to proteolysis during the time required for crystal growth.

One crystal with a different space group, $P6_522$, was obtained from a solution of Gel-NL:2actin. It diffracted to a resolution of only 7.5 Å. It was of crystal form GNL.A-3, and the unit cell parameters were $a = b = 145.76$ Å, $c = 373.49$ Å, and $\alpha = \beta = 90.0°$, $\gamma = 120.0°$. This set of space group and cell dimensions has been observed previously for similar resolution data for crystals grown from a G1-G3:actin solution (Chumnarnsilpa, *et al*., 2006).

Table 4.1 Data collection and refinement statistics for G1-G3:actin

Statistic	**G1-G3:actin**
Beamline	BL13B1, NSRRC
Wavelength, Å	1.000
Space group	$P2_12_12_1$
Unit cell dimensions	a = 101.5 Å, b = 147.9 Å, c = 147.0 Å α = β = γ = 90°
Resolution range, Å	25.6 – 2.78 (2.88 – 2.78)[d]
Unique reflections	55,708 (2,232)
Redundancy	4.7 (4.8)
Completeness, %	99.1 (91.9)
Average I/σ	29.3 (6.0)
R_{merge}[a], %	8.3 (33.8)
R_{factor}[b], %	18.2 (31.0)
R_{free}[c], %	24.5 (38.8)
Molecules in asymmetric unit	2
Gelsolin residue range	Mol 1: 28-374; Mol 2: 17-372
Actin residue range	Mol 1: 3-46, 49-364 Mol 2: 4-40, 51-368
Non-hydrogen atoms (Ca^{2+}, waters)	11,208 (10, 255)
Mean derived B-factor, Å^2	56.8
RMSD bonds, Å	0.008
RMSD angles, °	1.214

[a] $R_{merge} = (\Sigma | I - \langle I \rangle | / \Sigma \langle I \rangle)$

[b] $R_{factor} = (\Sigma | F_o | - | F_c | / \Sigma | F_o |)$

[c] based on 5% of the data

[d] highest resolution data shell

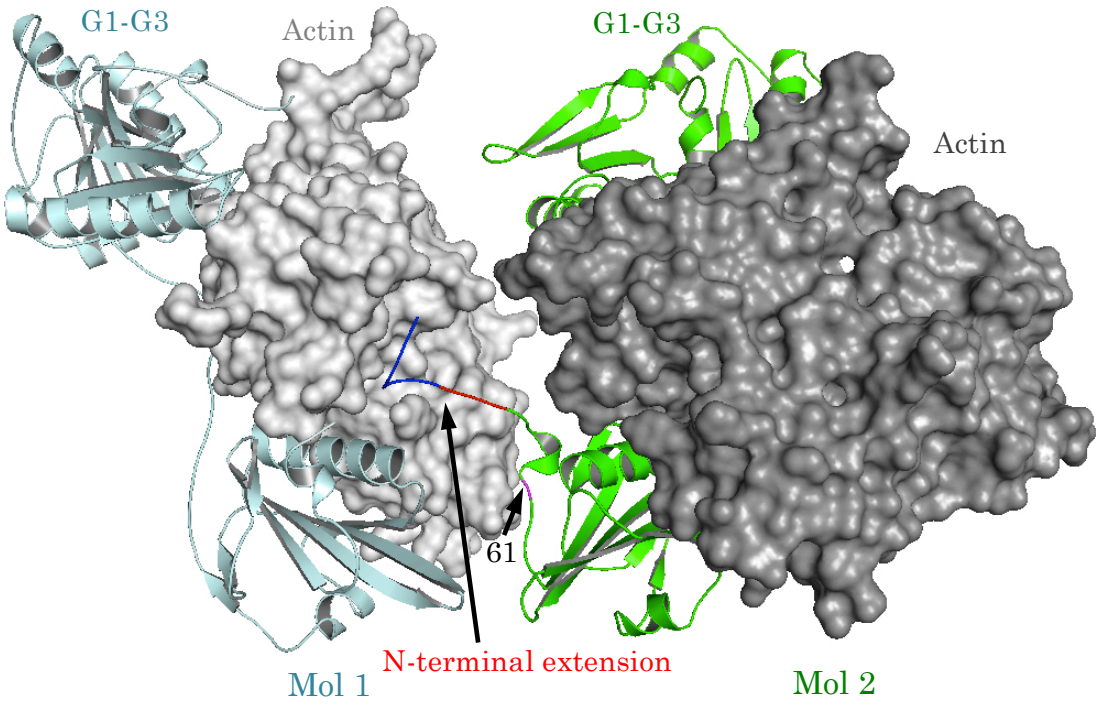

Figure 4.2 Crystal structure of G1-G3 in a complex with one actin at 2.8-Å resolution. Two G1-G3:actin complexes are found in each asymmetric unit. G1-G3 and actin are represented in ribbon and surface modes, respectively. The extra N-terminal extension residues on G1 of Mol 2 are highlighted in blue and red, with corresponding residues number. The blue region contains part of the engineered PreScission Protease cleavage site: Leu17, Phe18, Gln19, Gly20, Pro21, Gly22, Arg23, Pro24, while the red region contains gelsolin residues Met52, Val53, and Val54. These residues bind actin in Mol 1 near its hydrophobic cleft. A number of close contacts are made between the extension residues Pro21 to Arg23 on G1-G3 and actin residues Glu107, Arg116, Tyr133, Val134 and Ile136. Gelsolin residue Lys61 in Mol 2, highlighted in magenta, forms salt-bridge interactions with actin residues Asp286 and Asp288 in Mol1.

4.2.3 Structural analysis and discussion

4.2.3.1 The N-terminal extension of G1-G3:actin

The two structures of G1-G3:actin in the asymmetric unit, Mol 1 and Mol 2 (Fig 4.2), are very similar, with RMSD = 0.341 Å when the α-carbons of the polypeptide backbones are superimposed. However, there are two major differences between the structures: at the D-loop of actin and with regard to the N-terminal extension residues on gelsolin in Mol 2.

Mol 2 is missing residues His40 – Lys50 in the D-loop of actin, while the D-loop in Mol 1 is nearly complete, missing only residues 47 and 48. Actin residues Glu41 – Met44 of Mol 1 are found to form an anti-parallel β-sheet interaction with actin in Mol 2 of a symmetry-related complex. Also, the side chains of Asp51 and Lys50 on actin in Mol 1 form hydrogen-bonding interactions with the side chain and terminal carboxyl group of residue Gln374 at the C-terminus of the gelsolin portion of Mol 1, thus stabilizing the D-loop.

The N-terminal extension residues on G1-G3 in Mol 2 reaches away from the body of Mol 2 to bind back to the regions near the hydrophobic cleft of the neighboring actin in Mol 1 (Fig 4.2). These residues include Leu17, Phe18, Gln19, Gly20, Pro21, Gly22, Arg23, Pro24, Met25, Val26, and Val27, where the sequences Leu-Phe-Gln-Gly-Pro-Gly-Arg-Pro is part of the PreScission Protease cleavage site engineered into the protein construct and Met-Val-Val is part of gelsolin with a corresponding gelsolin residues Met52,

Val53, and Val54. In the present structure, these residues bind actin in Mol 1 near its hydrophobic cleft, adjacent to the previously identified binding site for G1 (McLaughlin *et al.*, 1993). A number of close contacts are made between extension residues Pro21 – Arg23 on G1-G3 and actin residues Glu107, Arg116, Tyr133, Val134, and Ile136. The interactions involve Pro21 (O), Gly22 (O), Arg23 (NH1) and Arg23 (NH2), which are in hydrogen-bond contact with Tyr133 (OH), Ile136 (N), Val134 (O) and Glu107 (OE2), and Arg116 (NH1), respectively. In addition, there is contact made through a salt-bridge interaction between residue Lys61 of G1-G3 in Mol 2 and residues Asp286 and Asp288 of actin in Mol 1 (Fig. 4.2). This interaction involves Lys61 (NZ) with Asp286 (OD1), Asp286 (OD2), and Asp288 (OD2).

4.2.3.2 The structure of human G1-G3:actin

The overall structure of human G1-G3:actin at 2.8-Å resolution shows a standard G-actin conformation decorated with gelsolin domains G1-G3 (Fig. 4.2). G1 is observed to bind to its known site on actin (McLaughlin *et al.*, 1993), the hydrophobic cleft between SD1 and SD3. The G1-G2 polypeptide linker then extends up and tightly binds along the surface of actin to allow G2 to bind on actin SD2. Lastly, the G2-G3 polypeptide linker helps position G3 to bind to the back of actin SD1. A disulfide bond connects residues Cys188 and Cys201. The overall fold of each domain is typical of gelsolin,

consisting of a core β-sheet sandwiched between a long and a sort helix (as described in section 4.1).

Inspection of the contact area between G1-G3 and actin reveals G1 to make contact with actin through its long helix, which is composed largely of hydrophobic amino acids. Analysis of the interface between G2 and actin SD2 shows that residues Asn223 to Arg228 of gelsolin contact actin in the region from Asn92 to Arg95. Both molecules in the asymmetric unit show gelsolin residues Asn223 (O), Ser226 (OG), Arg228 (NH2) to be in hydrogen-bond contact with Asn92 (ND2), Asn92 (OD1), and Glu93 (OE1) on actin, respectively. Furthermore, gelsolin residue Ser226 (O) and actin residue Glu93 (OE2) share the same water molecule through hydrogen bonding. In one of the two complexes in the asymmetric unit, hydrogen-bond contact is made between gelsolin residue Asp222 (OD2) and actin residue Glu57 (OE1). Also, a hydrogen bond between gelsolin residue Asn223 (OD1) and actin residue Arg95 (NH1) was found. Finally, gelsolin domain G3 residues Arg328 (NH1) and Lys319 (NZ) are in hydrogen bond contact with actin residues Glu100 (OE1) and Glu100 (OE2), respectively.

Five calcium ions are found associated with G1-G3:actin: one with actin, bound to ATP in the nucleotide-binding cleft, one type-I Ca^{2+} coordinated by residues from G1 and actin (IG1) (Fig. 4.3F), and one type-II Ca^{2+} coordinated in each of the three gelsolin domains (IIG1, IIG2, and IIG3; Figs. 4.3A, C, and E, respectively). The type-I calcium site in G1 involves

coordination by residues Asp109, Gly114 and Ala116 from gelsolin and residue Glu167 from actin, along with two water molecules (Fig. 4.3F). The type-II calcium site in G1 involves Gly65, Asp66, Glu97, plus a residue from in the G1-G2 linker region, Val145, and two water molecules (Figs. 4.3A and B). The type-II calcium site in G2 involves Gly186, Asp187, Glu209, a residue in the G2-G3 linker region, Asp259, and two water molecules (Figs. 4.3C and D). Residue Thr260 participates in a helix-initiating hydrogen bond in the G2-G3 linker helix, bringing G3 and G2 together (Nag *et al.*, 2009). The type-II calcium site in G3 involves Glu302, Asp303, Glu327, and three water molecules (Fig. 4.3E).

4.2.3.3 Comparison with previous structures of G1-G3:actin

As expected, the structure of human G1-G3:actin at 2.8-Å resolution (Fig. 4.4A) closely resembles that of human G1G3:actin at 3-Å resolution (Fig. 4.4B; PDB code 3FFK; Nag *et al.*, 2009) and the one of equine G1-G3:actin at 3-Å resolution (Fig. 3.5C; PDB code 1RGI; Burtnick et al., 2004). RMSD = 0.377 and 0.314 Å, respectively, when the α-carbons of the polypeptide backbone coordinates of PDB entry 3FFK and the PDB entry 1RGI are superimposed onto the 2.8-Å resolution structure of G1-G3:actin (Fig. 4.4D).

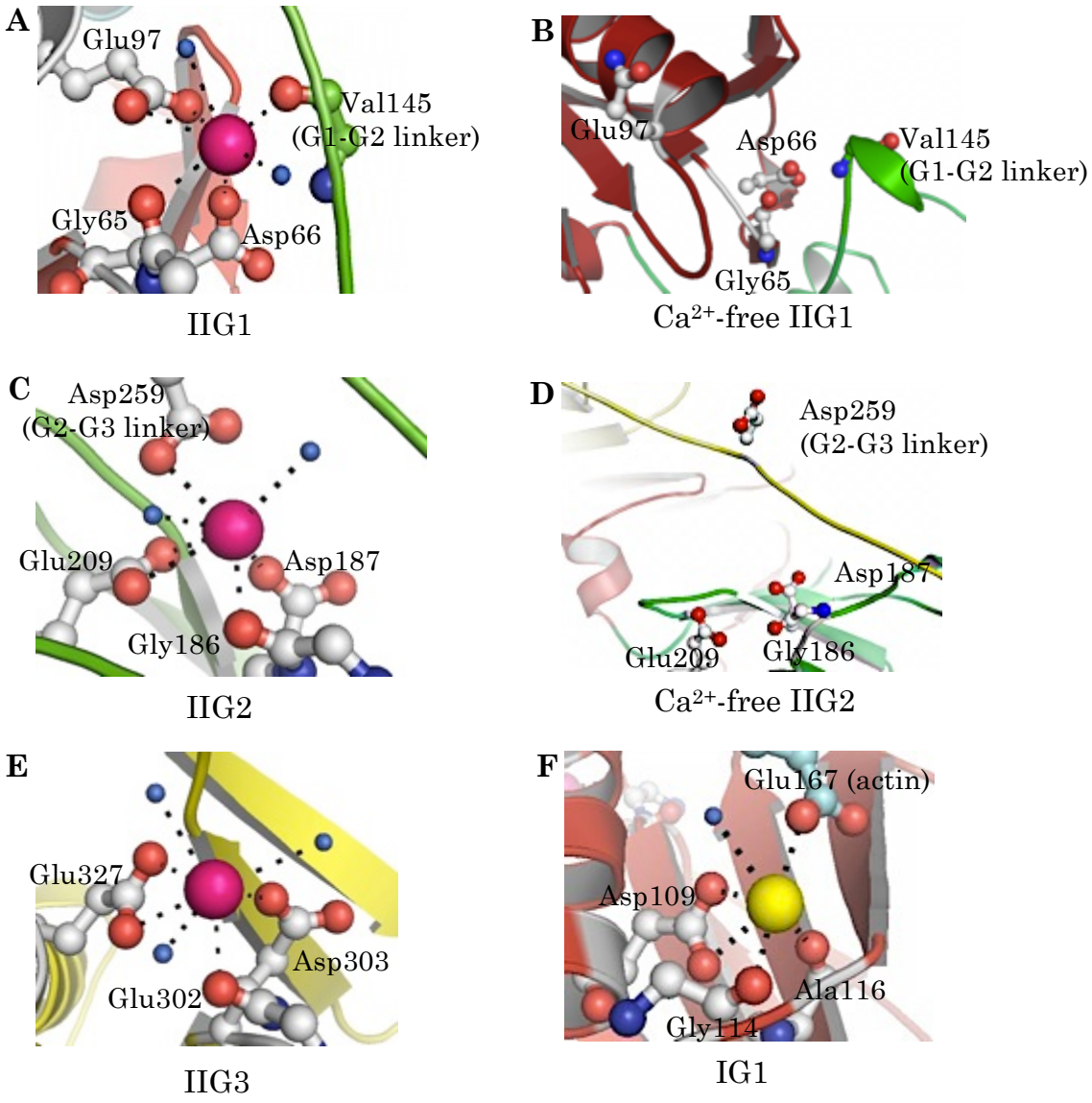

Figure 4.3 Ca^{2+}-binding sites in G1-G3. **(A)**, **(C)**, and **(E)** Occupied type-II Ca^{2+}-binding sites, respectively, in G1 (IIG1), G2 (IIG2) and G3 (IIG3), extracted from the 2.8-Å resolution structure of human G1-G3:actin. **(B)** and **(D)** Vacant type-II Ca^{2+}-binding sites in G1 and G2, respectively, extracted from PDB entry 1D0N. **(F)** Occupied type-I Ca^{2+}-binding site shared by G2 and actin (IG1) extracted from the 2.8-Å resolution structure of human G1-G3:actin. Pink and yellow spheres represent type-II and type-I Ca^{2+} ions, respectively. The Ca^{2+}-coordinating residues are presented in ball and stick mode. Water molecules are presented as blue spheres.

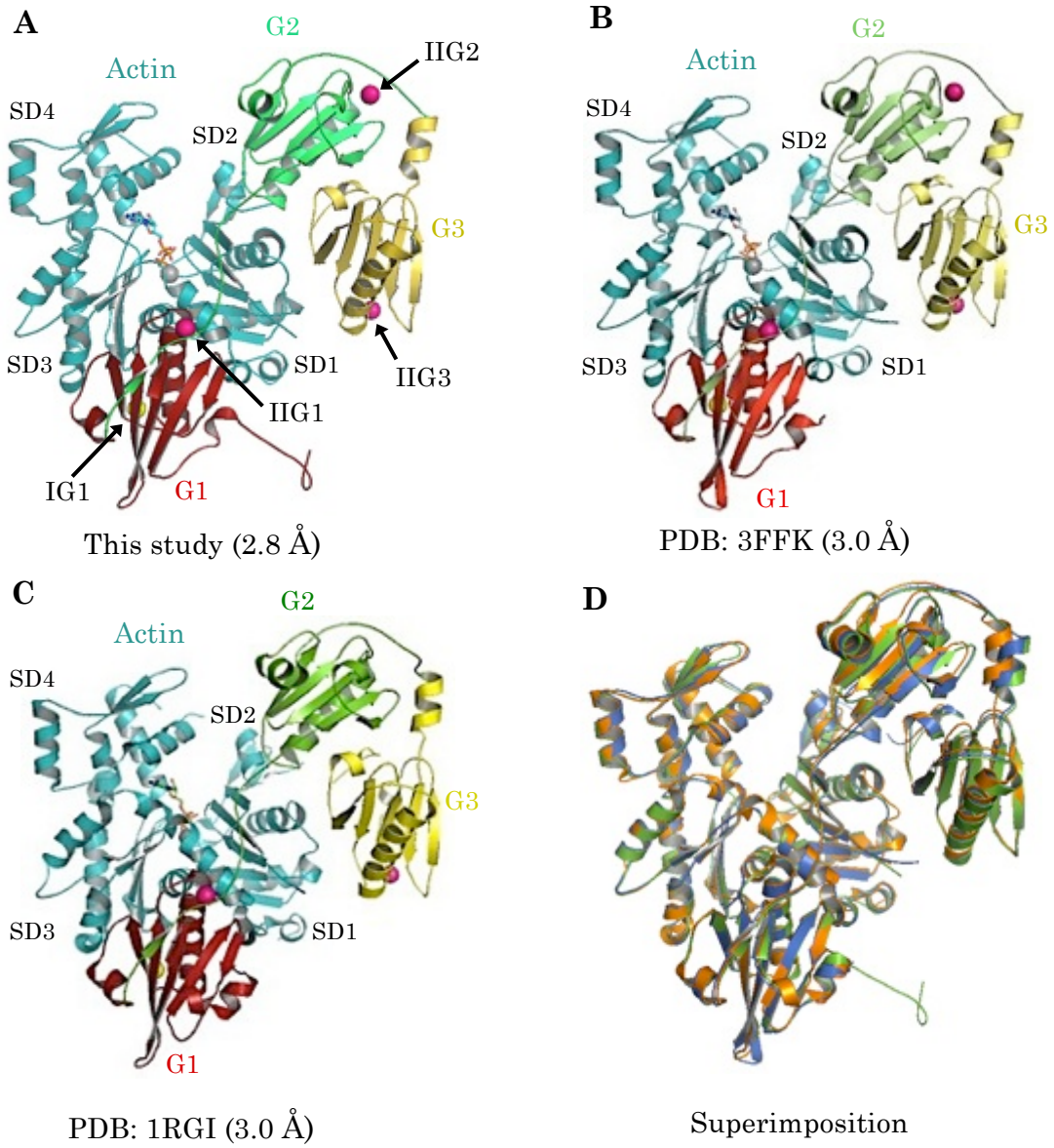

Figure 4.4 Comparison of the 2.8-Å resolution structure of G1-G3:actin with previously reported ones. **(A)** Human G1-G3:actin at 2.8-Å resolution (this study). **(B)** Human G1-G3:actin at 3.0-Å resolution (PDB code 3FFK; Nag *et al.*, 2009). **(C)** Equine G1-G3:actin at 3-Å resolution (PDB code 1RGI; Burtnick *et al.*, 2004). Yellow and pink spheres represent type-I and type-II Ca^{2+} ions, respectively in sites IG1, IIG1, IIG2, and IIG3. The Ca^{2+} ion associated with actin is presented as a grey sphere, with ATP in ball-and-stick format. **(D)** Superimposition of structures in (A) [coloured green], (B) [coloured blue], and (C) [coloured orange]. RMSD = 0.377 and 0.314 Å, respectively, when the α-carbons of the polypeptide backbones of (B) and (C) are superimposed onto those of (A).

Under close inspection at specific locations, some important differences can be noted: 1) A disulfide bond exists between residues Cys188 and Cys201 in both the present and the previous (Nag *et al*., 2009) human G1-G3:actin structures, whereas none is apparent in the structure of equine G1-G3:actin (Burtnick *et al*., 2004). 2) The two human G1-G3-:actin structures each bind five Ca^{2+} ions in the same manner, but the equine structure lacks Ca^{2+} at its IIG2 site. 3) In the type-II Ca^{2+}-binding site in G2 of the two human G1-G3:actin structures, a residue from the G2-G3 linker region, Asp259, participates in coordinating Ca^{2+}, together with Gly186, Asp187, Glu209 and two water molecules. Asp259 plays a similar role in the structure of isolated recombinant human G2 with Cd^{2+} bound to its metal ion-binding site (Kazmirski *et al*., 2002). Residue Thr260 participates in a helix-initiating hydrogen bond in the G2-G3 linker helix, bringing G3 and G2 together. In contrast, Asp259 in equine G1-G3:actin does not bind Ca^{2+}. Instead, it becomes the helix-initiating residue in the G2-G3 linker, which again brings G3 and G2 into close proximity (Burtnick *et al*, 2004; Nag *et al*., 2009).

4.2.3.4 Comparison with Ca^{2+}-free G1-G3 structure

Comparison of the overall structure of domains G1-G3 from the 2.8-Å resolution structure of G1-G3:actin with that of inactive (Ca^{2+}-free) G1-G3 (extracted from PDB code 1D0N; Burtnick *et al*., 1997) shows a large conformational rearrangement the domains relative to each other (Fig. 4.5).

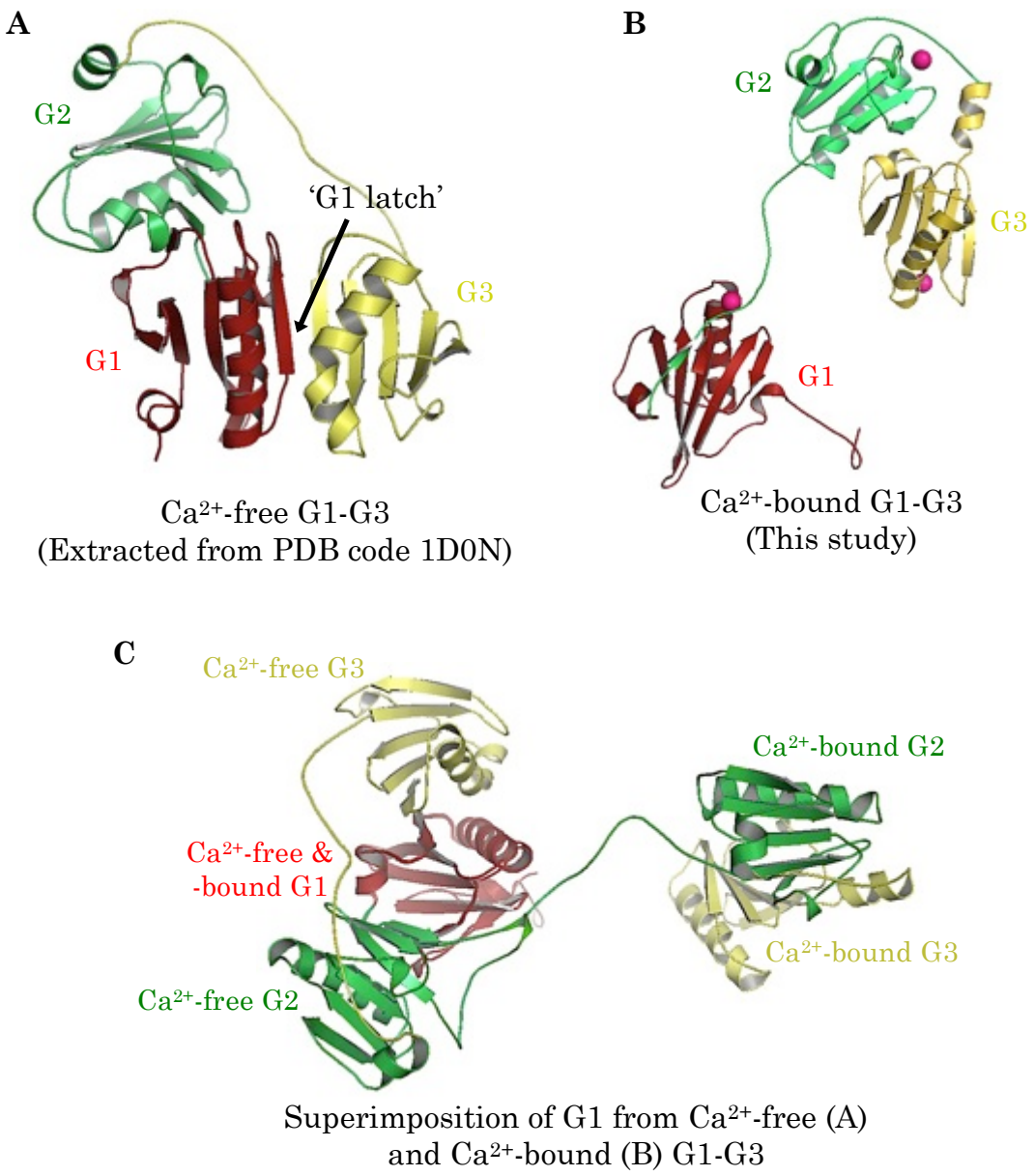

Figure 4.5 Comparison of Ca^{2+}-free and Ca^{2+}-bound structures of G1-G3. **(A)** Structure of Ca^{2+}-free G1-G3 extracted from PDB entry 1D0N (Burtnick *et al.*, 1997). **(B)** Structure of Ca^{2+}-bound G1-G3 extracted from the structure of G1-G3:actin at 2.8 Å resolution. **(C)** Superimposition of domain G1 from (A) and (B).

As a result of the rearrangement of domains, the actin-binding site on G1, which was hindered by G3 in the Ca^{2+}-free structure, becomes exposed (Burtnick *et al.*, 1997; Burtnick *et al.*, 2004). The conformational change primarily involves reorientation of the G1-G2 and G2-G3 linker peptides, without introducing drastic alterations into the individual domain folds. An exception is that the kinked long helix evident in Ca^{2+}-free G3 straightens in the presence of Ca^{2+} (Figs. 4.5A and B; Burtnick *et al.*, 1997; Burtnick *et al.*, 2004; Nag *et al.*, 2009).

Inspection of Ca^{2+}-binding site residues Gly186, Asp187, Glu209, and Asp259 in Ca^{2+}-free G2 shows Asp259 far removed from the others (Fig. 4.3D). Binding Ca^{2+} requires movement of Asp259 to complete the Ca^{2+}-binding site. Similarly, the carbonyl group of residue Val145 points away from the type-II binding site in Ca^{2+}-free (Fig. 4.3B) and reorientation is required for it to become involved in coordination of Ca^{2+} at that site (Fig. 4.3A).

In Ca^{2+}-free G1-G3, the core sheets of G1 and G3 are annealed into a single extended sheet, resulting in formation of the G1 latch (Fig. 4.5A), which disallows contact between G1 and actin (Burtnick *et al.*, 1997). Activation by binding Ca^{2+} unzips the extended sheet, opening the latch that frees G1 and exposes its actin-binding site (Burtnick *et al.*, 2004).

4.3 Structure of activated G3 in isolation

Diffraction data from crystals grown from solutions of activated gelsolin domain G2-G4 yielded the structure of Ca^{2+}-bound gelsolin domain G3. Proteolysis during the time taken to grow crystals is the most likely reason for the electron density map requiring only G3 to satisfy it.

4.3.1 Crystallization

Purified G2-G4 was activated with 2 mM $CaCl_2$ before being subjected to crystallization trials. Two different crystal forms were obtained when G2-G4 solutions were equilibrated at 4 °C against Hampton crystal screen 2, condition no. 22 [12%(w/v) PEG20000, 100 mM MES monohydrate, pH 6.5] (Fig. 4.6A) and homemade Stura footprint screen 1:C3 [33%(v/v) PEG600, 200 mM imidazole malate, pH 5.5] (Fig. 4.6B).

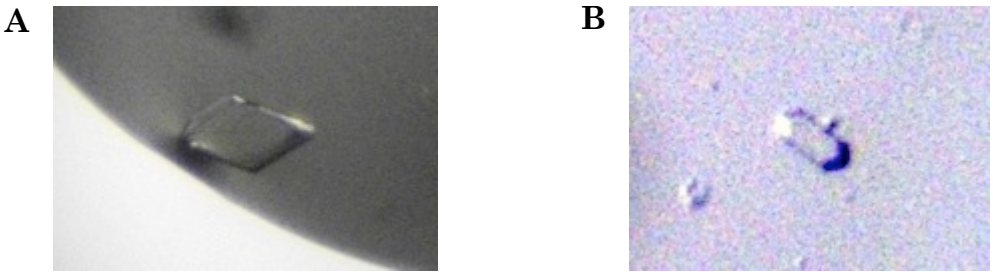

Figure 4.6 Two crystal forms grown using activated gelsolin G2-G4 solutions. **(A)** A crystal grown in Hampton crystal screen 2 [12%(w/v) PEG20000, 100 mM MES monohydrate, pH 6.5]. **(B)** A crystal grown in homemade Stura footprint screen 1:C3 [33%(v/v) PEG600, 200 mM imidazole malate, pH 5.5].

4.3.2 Data collection and refinement

Crystals from Hampton crystal screen 2, no. 22 (Fig. 4.6A), supplemented with 20%(v/v) glycerol, diffracted to a maximum resolution of 2.4 Å, in space group $P2_12_12_1$. The electron density map obtained after molecular replacement analysis of the diffraction data could be explained completely by considering the crystal to be of isolated activated gelsolin domain G3. Again, the missing protein, domains G2 and G4, could be the result of proteolysis during the time required to grow the crystals. The crystallographic asymmetric unit contains one molecule of activated G3. The unit cell parameters turned out to be a = 37.13 Å, b = 42.79 Å, c = 70.33 Å, and $\alpha = \beta = \gamma = 90.0°$ (Table 4.2).

Crystals from homemade Stura footprint screen 1:C3 (Fig 4.6B), supplemented with 20%(v/v) glycerol, diffracted to a maximum resolution of 2.0 Å. The crystal belonged to space group I422 with the unit cell parameters $a = b$ = 120.4 Å, c = 137.4 Å, and $\alpha = \beta = \gamma = 90.0°$. The structures of gelsolin domains G2, G3, and G4 failed to yield solutions when used as search models in molecular replacement analysis. Subsequently, the space group and cell dimensions parameters were used to search against the PDB database (http://www.pdb.org) and found to match the structure of *Escherichia coli* BL21 Gab protein (PDB code 2R6S; Lohkamp and Dobritzsch 2008). The Gab protein from the *E. coli* BL21, which was used as the host cell in the production of G2-G4, must survive the purification process.

Table 4.2 Data collection and refinement statistics for G3.

Statistic	**Active G3**
Beamline	BL13B1, NSRRC
Wavelength, Å	1.000
Space group	$P2_12_12_1$
Unit cell dimensions	a = 37.1 Å, b = 42.8 Å, c = 70.3 Å α = β = γ = 90°
Resolution range, Å	36.6 - 2.37 (2.99-2.37)[d]
Unique reflections	4,769 (2,179)
Redundancy	4.3 (3.9)
Completeness , %	99.8 (97.0)
Average I/σ	18.9 (7.4)
R_{merge}[a], %	8.2 (29.5)
R_{factor}[b], %	19.9 (20.9)
R_{free}[c], %	25.9 (34.7)
Molecules in asymmetric unit	1
Residue range	270-372
Non-hydrogen atoms (Ca^{2+}, waters)	808 (1, 67)
Mean derived B-factor, $Å^2$	33.1
RMSD bonds, Å	0.009
RMSD angles, °	1.228

[a] $R_{merge} = (\Sigma | I - \langle I \rangle | / \Sigma \langle I \rangle)$

[b] $R_{factor} = (\Sigma | F_o | - | F_c | / \Sigma | F_o |)$

[c] Based on 5% of the data.

[d] highest resolution data shell

4.3.3 Structural analysis and discussion

4.3.3.1 The structure of activated G3 in isolation

Activated G3 possesses the usual gelsolin domain fold: a five-stranded β-sheet sandwiched between a long and a short helix, oriented parallel and perpendicular to the sheet, respectively (Burtnick *et al*., 1997). The structure contains one calcium ion occupying the known type-II calcium site (Fig. 4.7A). The coordinating residues are identical to those reported in the published G1-G3:actin structures (Burtnick *et al*., 2004; Nag *et al*., 2009), as well as in the G1-G3:actin structure reported in this thesis. Again, type-II calcium binding in G3 involves Glu302, Asp303, Glu327, and three water molecules (Fig. 4.7B). Superimposition of the current activated G3 structure with that of activated G3 extracted from the structures of human G1-G3:actin (PDB code 3FFK) and equine G1-G3:actin (PDB code 1RGI) shows the structures to be very similar (Fig. 4.7C), with RMSD = 0.514 and 0.516 Å for α-carbon positions, respectively. Of particular note, as observed in structures of G1-G3:actin, the long helix in activated G3 is straight (Fig. 4.7A).

4.3.3.2 Comparison with the structure of inactive G1-G3

In the structure of inactive (Ca^{2+}-free) gelsolin (PDB code 1D0N), the kink in the long helix of G3 appears to exist in order to avoid collision between it and the long helix of G1 (Fig. 4.8A).

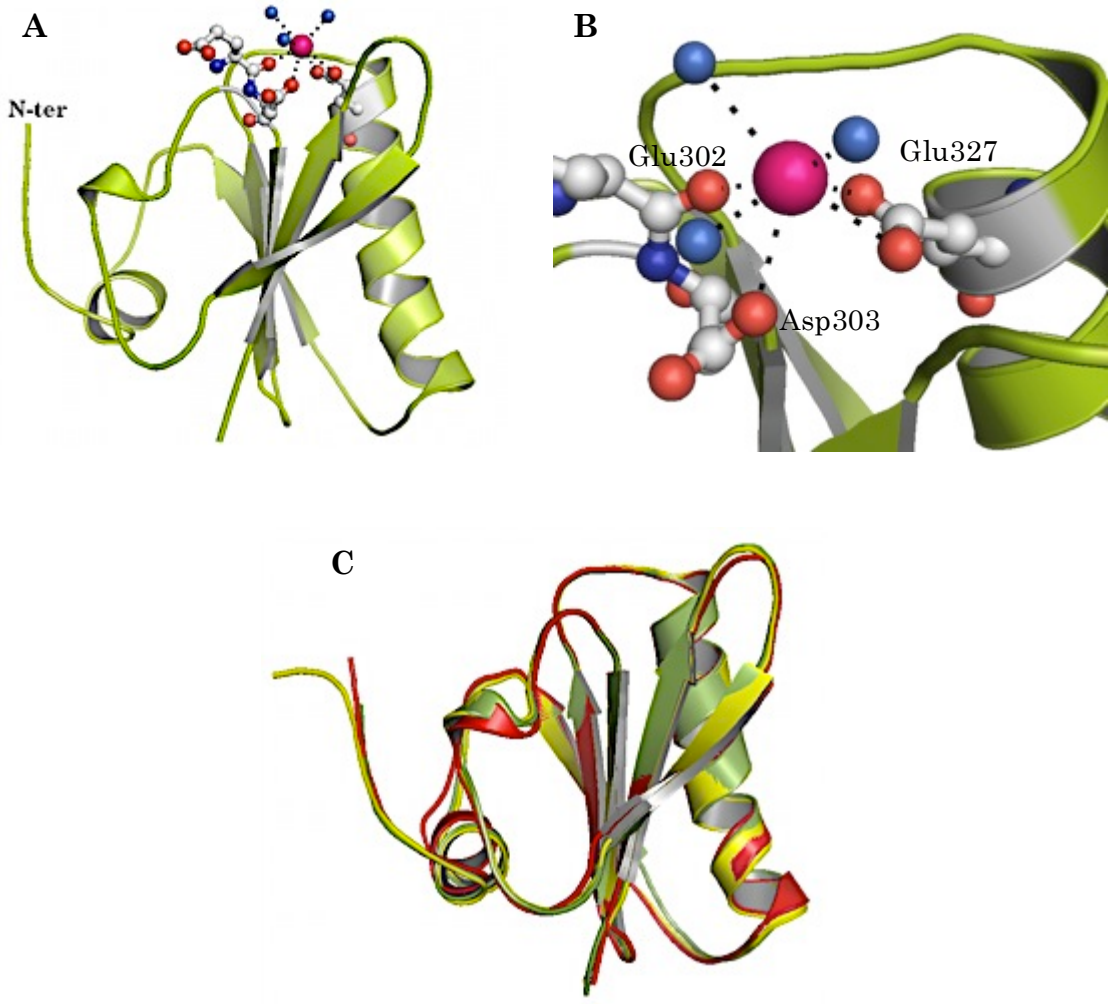

Figure 4.7 The crystal structure of isolated activated human G3. **(A)** A cartoon representation of isolated activated G3. Calcium and coordinating residues are shown in ball-and-stick format. **(B)** A close-up picture of the coordinating residues of the type-II calcium-binding site of G3 in (A), which is identical to that of the reported in the published G1-G3:actin structures (Burtnick *et al.*, 2004; Nag *et al.*, 2009). **(C)** Superimposition of the current structure (green) with those of G3 extracted from human G1-G3:actin (PDB code 3FFK; red) and equine G1-G3:actin (PDB code 1RGI; yellow).

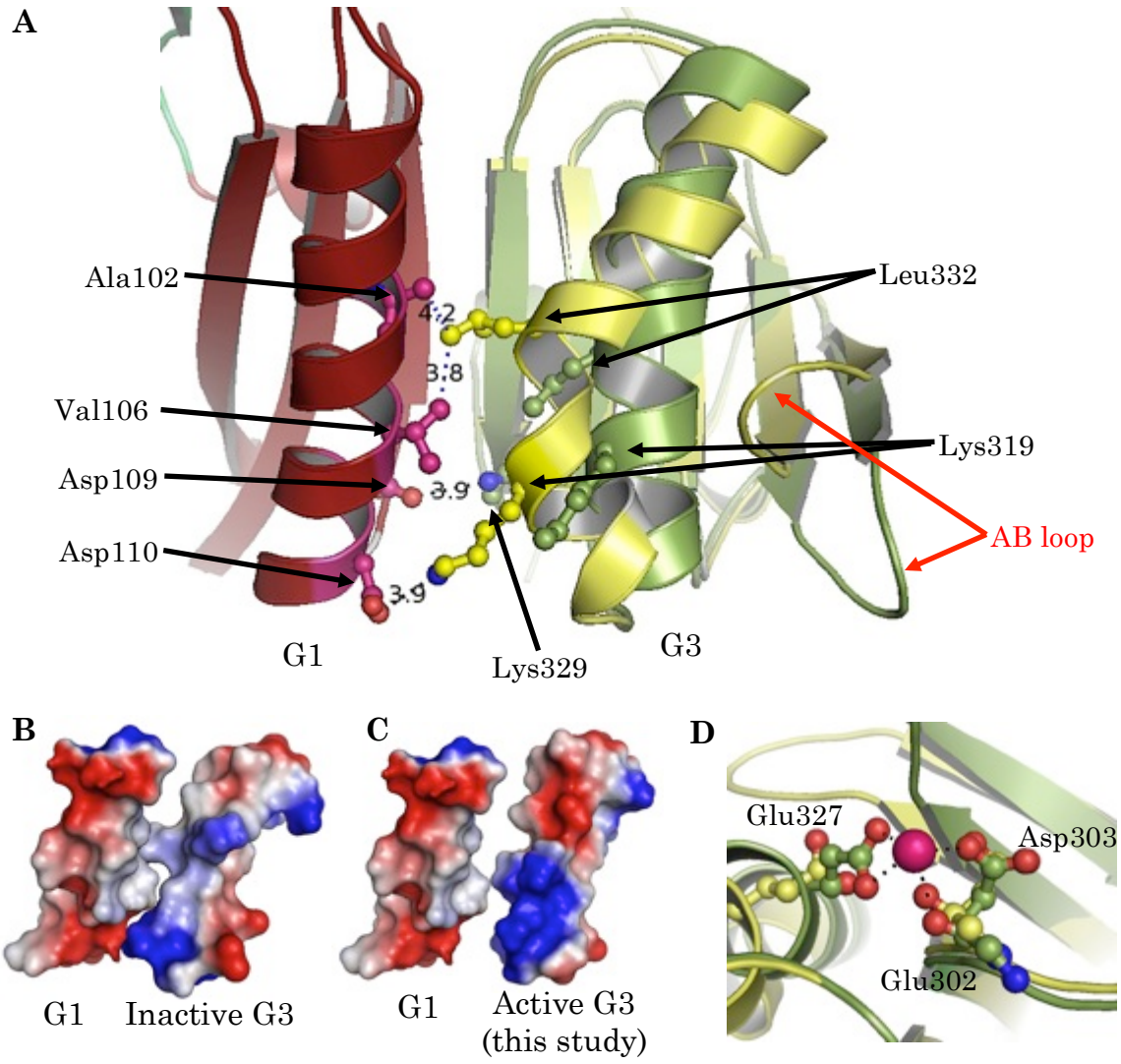

Figure 4.8 Superposition of activated G3 onto inactive G1-G3. **(A)** The structure of activated G3 (green; this study) superimposed onto inactive G1-G3 (G1 is red; G3 is yellow; extracted from PDB code 1D0N), in which the distances in Å between interacting residues are shown. **(B)** Electrostatic surface representations of the long helices of G1 and G3 from inactive G1-G3, oriented as in (A). Red and blue are negative and positive areas, respectively. **(C)** Electrostatic surface representations as in (B) but with the long helix of activated G3 replacing that of inactive G3. **(D)** Comparison of the type-II Ca^{2+}-coordinating residues in inactive (yellow) and activated (green) G1.

It appears that more than just the seam between the two core sheets holds the G1 latch closed in the inactive conformation of G1-G3. Residues Ala102, Thr105, and Val106 from the helix of G1 are situated within 5 Å of Leu332 in the long helix of G3, sufficiently close to identify this grouping as a hydrophobic interaction (Fig. 4.8A). Furthermore, a search for salt-bridge interactions using ESBRI software (Costantini *et al*., 2008) identified Asp109 (OD2), Asp110 (OD1), and Asp110 (OD2) in the G1 long helix to be paired with Lys319 (NZ), Lys329 (NZ), and Lys329 (NZ), respectively, in the long helix of G3 at separations of 3.93, 3.20 and 3.93 Å (Figs. 4.8A and B). Superimposing the helix from activated G3 on to that of inactive G3 moves the Lys residues sufficiently far away from the Asp residues that salt-bridges would not form (Fig. 4.8C).

Next to differences between their long helices, the most apparent distinction between the activated and inactive forms of G3 is in the positioning of the loop between and the A and B strands of their core sheets (the AB loop in Fig. 48A). The AB loop of inactive G3 approaches the helix, whereas that of activated G3 reaches away from it. Straightening of the helix places residue Glu327 into a better position with regard to coordination of Ca^{2+} in the type-II site in activated G3 (Fig. 4.8D).

4.4 Summary

The higher resolution structure of human G1-G3:actin confirms the differences reported previously between the structures of human G1-G3:actin (Nag *et al.*, 2009) and equine G1-G3:actin (Burtnick *et al.*, 2004). Firstly, disulfide bridges between residues Cys188 and Cys201 were detected in the structures of human G1-G3:actin. This disulfide bond in G2 of human plasma gelsolin is thought to influence the Ca^{2+}-activation process by facilitating communication between Ca^{2+}-binding sites in the C-terminus and active sites in the N-terminus (Allen, 1997). Secondly, both human G1-G3:actin structures have a Ca^{2+} ion securely coordinated into the type-II metal ion-binding site in G2, whereas the site is vacant in equine G1-G3:actin. Ca^{2+}-binding at this site in G2 results in rearrangement and stabilization of interactions between domains G2 and G3. Loss of ability to bind Ca^{2+} at this site as a consequence of a naturally occurring mutation of Asp187, a Ca^{2+}-binding residue, to a Tyr or an Asn renders the structure susceptible to furin digestion, which is the first in a cascade of events that leads to amyloid deposits in the inherited condition known as Familial Amyloidosis, Finnish type (FAF; Maury, 1991; Kazmirski *et al.*, 2002).

The structure of isolated activated G3 shows that straightening of the long helix in G3 is required to allow residue Glu327 to complete the Ca^{2+}-coordination sphere in the type-II metal ion-binding site of G3. Ca^{2+}-binding

here then stabilizes the domain structure of G3 relative to that of the inactive form, helping to drive the overall activation process.

Chapter 5

A New Look at Activated Domain Structures in Gelsolin: The C-terminal Half

5.1 Background

Activation by Ca^{2+} of gelsolin releases the tail latch, allowing the N- and C-terminal halves to come free from each other (Burtnick *et al*., 2004). In the structure of the C-terminal half of human gelsolin bound to actin, there are four Ca^{2+}-binding sites associated with gelsolin (PDB code 1H1V; Robinson *et al*., 1999; Choe *et al*., 2002). One is a type-I site, formed between G4 and actin. The remainder consists of one type-II site wholly contained within each of the three gelsolin domains.

Crystal structures of the activated C-terminal half of gelsolin alone, free of actin, from both human (PDB code 1P8X; Narayan *et al*., 2003) and mouse (PDB code 1NPH; Kolappan *et al*., 2003) reveal that activation of the C-terminal half of gelsolin can occur without actin, yielding a conformation that can simply be docked to actin.

Although only two Ca^{2+} ions were originally claimed to be seen in the structure of activated murine G4-G6 (PDB code 1NPH; Kolappan *et al.*, 2003), in G5 and G6, respectively, the Fo-Fc electron density difference map calculated from the deposited coordinates and structure factors for PDB entry 1NPH shows a 6.5σ peak at the expected position of the G4 type-II Ca^{2+} site, indicating a significant level of occupancy of this third type-II site (Nayaran *et al.*, 2003). This would then be consistent with observation of all three type-II sites being occupied in the structure of activated human G4-G6 (PDB code 1P8X; Narayan *et al.*, 2003). In neither structure is a type-I Ca^{2+} ion observed, as expected due to the absence of actin.

The G3-G4 linker polypeptide between the N- and C-terminal halves of is to span between two actin units across the filament from each other when gelsolin binds to F-actin (McGough *et al.*, 2003). Two alternative paths for the linker can be modeled, one requiring an extension of 100 Å and a second of about 75 Å.

This thesis presents the structure of an activated form of the C-terminal half of human gelsolin in the absence of actin (activated human G4-G6) determined using different starting materials and crystallization conditions from those previously reported. This structure confirms the previously one (Narayan *et al.*, 2003).

Attempts to crystallize Gel-CL (G4-G6 with the G3-G4 linker at its N-terminus) in a complex with actin (Gel-CL:actin) failed to yield single crystals suitable for X-ray diffraction analysis.

5.2 Structure of the activated C-terminal half of gelsolin

5.2.1 Crystallization

Crystals grew in solutions of a 1:2 molar ratio complex of FL-Gelsolin and actin (GA_2) using the hanging drop vapor diffusion method at 4 °C when equilibrated against well buffers corresponding to five different conditions from the PACT suite (Qiagen), including conditions no. 4 [25%(w/v) PEG1500, 100 mM SPG buffer, pH 7.0], no. 13 [25%(w/v) PEG1500, 100 mM MIB buffer, pH 4.0], no. 15 [25%(w/v) PEG1500, 100 mM MIB buffer, pH 6.0], no. 76 [20%(w/v) PEG3350, 200 mM potassium thiocyanate, 100 mM Bis-Tris propane, pH 7.5], and no. 87 [20%(w/v) PEG3350, 200 mM sodium iodide, 100 mM Bis-Tris propane, pH 8.5].

Crystals grown against conditions no. 6 and 15 had similar appearances (Fig. 5.1A), whereas those grown against condition no. 13 were of a somewhat different form (Fig. 5.1B). Crystals grown against conditions 76 and 87 also resembled each other (Fig. 5.1C), but differed from the others.

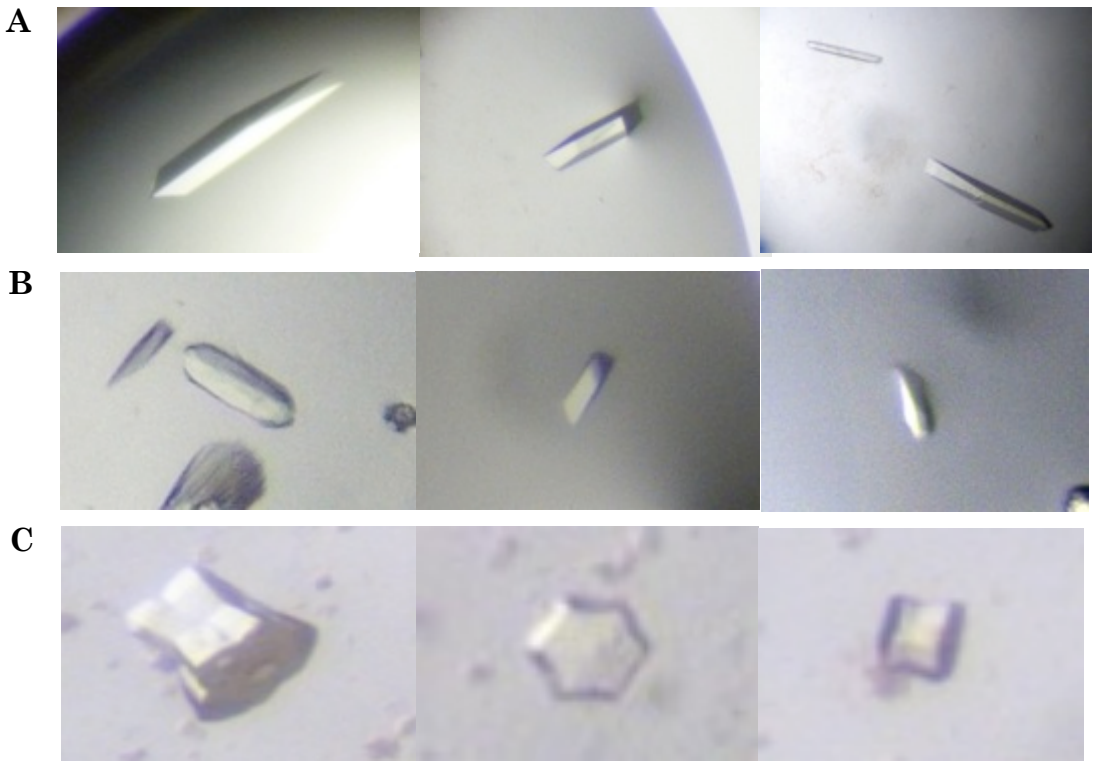

Figure 5.1 Examples of crystals grown from pooled fractions of GF-purified GA_2 complex. Five different conditions of the PACT suite (Qiagen) yielded crystals. **(A)** Crystals grown in conditions no. 6 and 15. **(B)** Crystal grown condition no. 13. **(C)** Crystal grown in conditions no. 76 and 87.

5.2.2 Data collection and refinement

Four data sets were collected from crystals grown in condition no. 76 at resolutions of 2.8, 2.9, 3.4 and 3.5 Å, and one data set of resolution of 4.0 Å was collected from a crystal grown in condition 87. Some of these crystals belonged to different space groups and had unique cell dimensions (Table A.5, Appendix A). Molecular replacement searches using the structures of both inactive and activated forms of gelsolin and gelsolin fragments failed to yield a solution to the data. Using actin as a search model also was not successful. The space groups and cell dimensions parameters (Table A.5, Appendix A) then were used to search against the PDB database (http://www.pdb.org) to find appropriate search models for Molecular replacement analysis. The search results did not reveal any structure with a matching space group and cell dimensions. These data sets remain to be solved.

One crystal from each of conditions no. 6 and 13 yielded X-ray diffraction data sets at resolutions of 2.4 and 2.7 Å, respectively, while three data sets at resolutions of 2.3, 2.4 and 2.6 Å were collected from crystals grown in condition no. 15 (Table A.3, Appendix A). These crystals all belonged to space group $P2_12_12_1$ and had common unit cell parameters: (based on the 2.3-Å data set) a = 85.0 Å, b = 88.0 Å, c = 156.0 Å, and $\alpha = \beta = \gamma = 90.0°$ (Table A.3, Appendix A). The 2.3-Å resolution data set was subjected to molecular replacement analysis (Rossmann 1972; Rossmann 1990) in Phenix_AutoMR (PHENIX suite) using several known structures involving domains of gelsolin

with or without actin. A successful solution was achieved using the structure of activated human G4-G6 (PDB code 1P8X), from which the calcium ions were removed, as the search model. The resulting solution revealed three molecules of activated G4-G6 in each crystallographic asymmetric unit, with the same molecular arrangement as reported previously for activated human G4-G6 (PDB code 1P8X). Neither the N-terminal half of gelsolin nor actin were required to satisfy the electron density map. Possibly, proteolysis occurred during the time required for the crystals to nucleate. The space group and unit cell dimensions (Table 5.1) of these crystals are similar to those reported for human G4-G6 (Narayan *et al*, 2003; PDB code 1P8X), but different from those for mouse G4-G6 (Kolapan *et al*., 2003; PDB code 1NPH), which were $P4_12_12$ and $a = b = 122.52$ Å, $c = 82.72$ Å, and $\alpha = \beta = \gamma = 90.0°$.

Previously, crystals grown from solutions of GA_2 complexes prepared using natural source equine plasma gelsolin (Burtnick *et al*, 2004) or recombinant human plasma gelsolin (Nag *et al*., 2009) with natural source rabbit muscle actin yielded structures for G1-G3:actin. (The second actin and the C-terminal half of gelsolin were not evident.) The present result is the first in which a crystal grown from a solution of GA_2 yielded the structure of the isolated C-terminal half of gelsolin, without the N-terminal half of gelsolin or any actin.

Table 5.1 Data collection and refinement statistics for activated G4-G6

Statistic	**Activated G4-G6**
Beamline	CMCF-08B1-1, CLS (Remote)
Wavelength, Å	0.977
Space group	$P2_12_12_1$
Unit cell dimensions	a = 84.9 Å, b = 88.3 Å, c = 156.2 Å α = β = γ = 90°
Resolution range, Å	48.2 - 2.25 (2.33 -2.25)[d]
Unique reflections	56,375 (5,292)
Redundancy	7.4 (7.2)
Completeness, %	99.9 (99.8)
Average I/σ	17.8 (6.2)
R_{merge}[a], %	11.0 (34.8)
R_{factor}[b], %	20.5 (40.8)
R_{free}[c], %	27.7 (47.2)
Molecules in asymmetric unit	3
Residue range	Mol 1: 413-635, 638-707, 710-741 Mol 2: 414-453, 459-741 Mol 3: 414-455, 460-741
Non-hydrogen atoms (Ca^{2+}, waters)	7,556 (9, 344)
Mean derived B-factor, $Å^2$	53.2
RMSD bonds, Å	0.008
RMSD angles, °	1.190

[a] $R_{merge} = (\Sigma | I - \langle I \rangle | / \Sigma \langle I \rangle)$

[b] $R_{factor} = (\Sigma | F_o | - | F_c | / \Sigma | F_o |)$

[c] based on 5% of the data

[d] highest resolution data shell

5.2.3 Structural analysis and discussion

5.2.3.1 The structure of human G4-G6:actin

The structure of active recombinant murine gelsolin G4-G6 (PDB code 1NPH; Kolappan, *et al*., 2003) was obtained from crystals grown in 25 mM Tris-HCl, pH 8.0, 15%(w/v) PEG8000, and 0.1 mM $CaCl_2$ at 20 °C. Later the same year, the structure of active recombinant human G4-G6 (PDB code 1P8X; Narayan *et al*., 2003) became available from crystals grown in 15% (w/v) PEG8000, 100mM Tris-HCl, pH 7.5, at 4 °C. In the present structure, the set of crystallization conditions and source and identity of the gelsolin fragment are significantly different from those reported previously.

In the present structure of activated G4-G6, each domain was found to assume the typical gelsolin fold (Burtnick *et al*., 1997). The overall structures of the three molecules of activated G4-G6 in an asymmetric unit are very similar (Figs. 5.2A and C). The RMSD = 0.537 Å or 0.699 Å when the α-carbons of Mol 1 are superimposed onto those of Mol 2, or those of Mol 3 onto those of Mol 1, respectively. The RSMD = 0.235 Å, 0.363 Å, or 0.673 Å when the α-carbons of Mol 1, Mol 2, and Mol 3 from the new active G4-G6 structure are superimposed (Fig 4.3B) onto their corresponding atoms in the earlier structure for active human G4-G6 (PDB code 1P8X; Narayan *et al*., 2003). The loop containing residues Ser46 to Glu65 is missing in Mol 1 of the structure from PDB code 1P8X, but is present in the new structure.

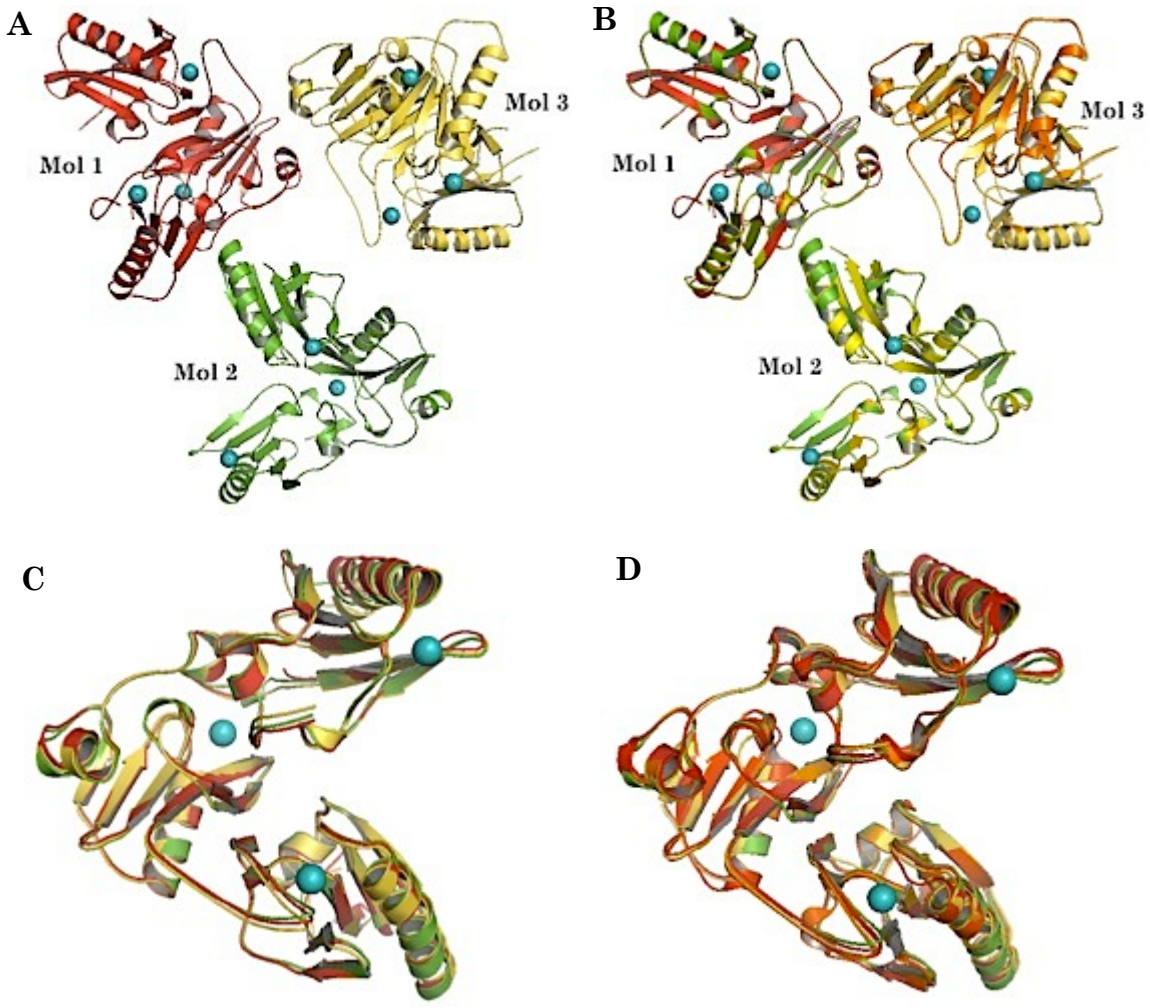

Figure 5.2 Structures of the three molecules of activated human G4-G6 in each crystallographic asymmetric unit and their superimposition. **(A)** The three molecules of activated G4-G6, denoted Mol 1, Mol 2, and Mol 3, that appear in each crystallographic asymmetric unit of the new structure. **(B)** Superimposition of the structure in (A) onto the earlier structure of activated human G4-G6 (PDB code 1P8X). **(C)** Superimposition of all three molecules of activated G4-G6 from (A) onto each other. **(D)** Superimposition of all three molecules of activated G4-G6 from PDB entry 1P8X onto each other. In all parts of this figure, calcium ions are shown as cyan spheres.

5.2.3.2 Comparison with Ca^{2+}-free G4-G6 structure

In inactive gelsolin the G4 actin-binding site is hindered by the core β-sheet latch formed between domains G4 and G6. To attain the structure of activated G4-G6 (Fig 5.2) requires unzipping of the extended sheet shared by G4 and G6 in inactive gelsolin, together with large-scale movement of G6 relative to G4. Moreover, the kinked helix of G6 observed in the inactive structure must straighten to achieve the activated conformation.

5.2.3.3 Ca^{2+}-binding sites

Inspection of the Ca^{2+}-binding sites of the new activated G4-G6 structure reveals three Ca^{2+} ions, one in each of the three type-II sites, confirming the existence a Ca^{2+} in the type-II Ca^{2+} site of G4 (Narayan *et al*, 2004). The fourth Ca^{2+}-binding site (type-I), located in G4, is vacant (Fig. 5.3A), as expected in that actin is not present to complete the site, and as reported for the previous structures (Kolappan *et al*., 2003; Narayan *et al*., 2003). Residues Gly444, Asp445, Glu475, and Thr524, plus two water molecules, coordinate the type-II Ca^{2+} in G4. Coordination of the type-II Ca^{2+} in G5 involves residues Asn564, Asp565, and Glu587, along with two water molecules. Coordination of the type-II Ca^{2+} in G6 involves Asp669, Asp670, Glu692, and three water molecules.

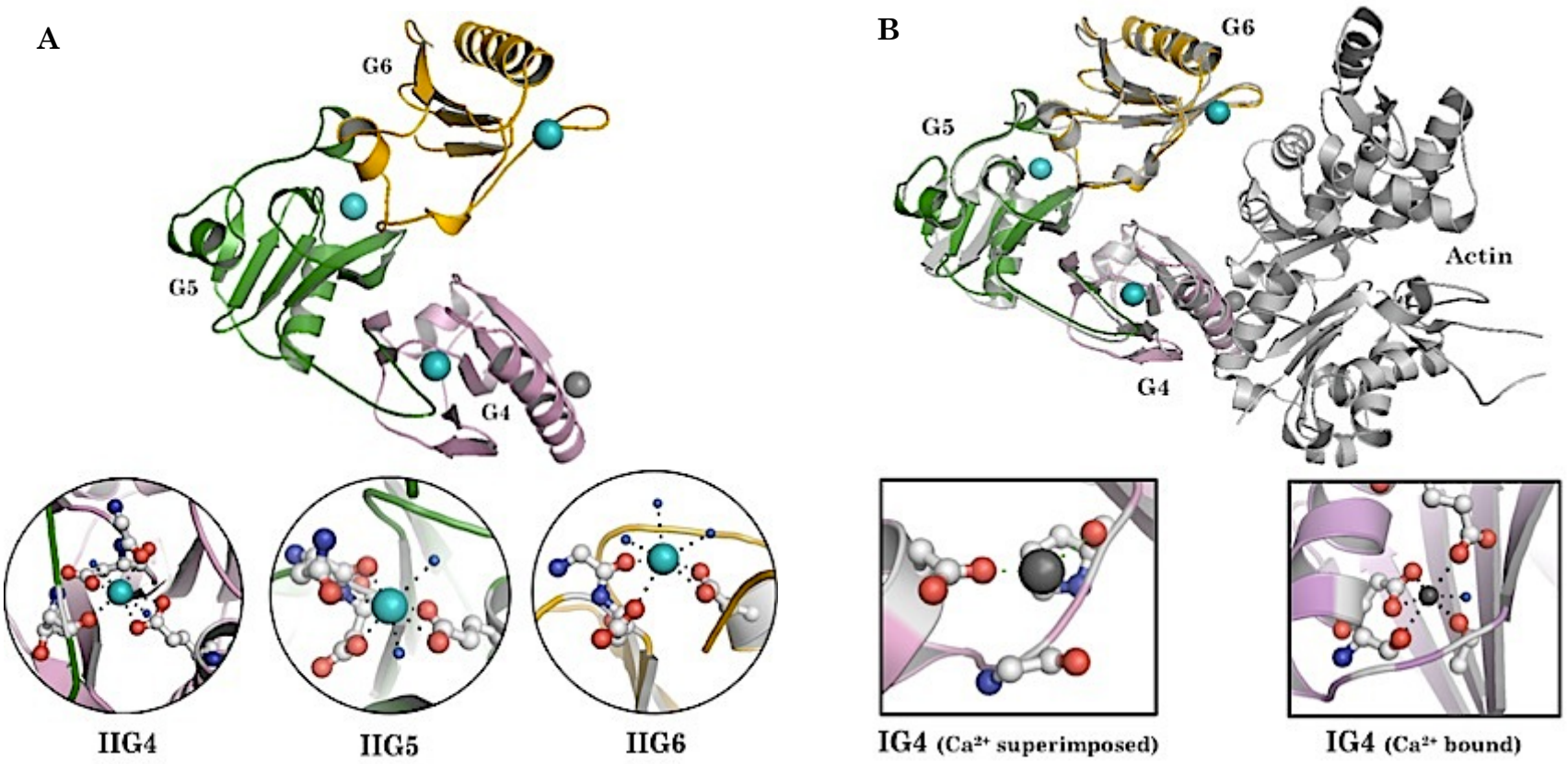

Figure 5.3 The structure of the activated C-terminal half of gelsolin (G4-G6). **(A)** Coordination of type-II Ca^{2+} ions (cyan) within activated G4-G6. **(B)** Superimposition of activated G4-G6 onto G4-G6:actin (PDB code 1H1V), emphasizing the type-I calcium-binding site. The type-I Ca^{2+} is coloured grey.

When superimposed (Fig. 5.3B), the new activated G4-G6 structure and the previous one for activated G4-G6:actin (PDB code 1H1V; Choe *et al*, 2002) show that the isolated activated G4-G6 is ready to bind actin (Narayan *et al*., 2003). A type-I Ca^{2+} ion binds in the G4-G6:actin structure, coordinated by residues Asp187, Gly492, and Pro494 from G4, and Glu167 from actin. From the superimposed structures, we see that the type-I Ca^{2+}-binding site of isolated activated G4-G6 is properly arranged with respect to coordination of a Ca^{2+} ion by Asp487 and Pro494, but the coordination sphere is incomplete in the absence of actin.

5.3 Crystallization of Gel-CL:actin

Neither commercial screens nor our homemade screens resulted in crystals suitable for X-ray diffraction analysis. However, needle crystals from solutions of Gel-CL:actin were obtained (Fig. 5.4A) when the protein solution was equilibrated at 4 °C against precipitant solutions that consisted of 20%(v/v) MPEG550, 100 mM Tris-HCl, pH 8.0 (Fluka crystallization kit for protein complexes, precipitant no. 2). In addition, screens near the conditions successfully used for G4-G6:actin [10%(w/v) PEG8000, 20%(v/v) glycerol, 100 mM HEPES buffer, pH 7.5] also resulted in needle crystals (Fig. 5.4B).

Moreover, an attempt to obtained single crystals by using seeding techniques (Bergfors, 1999) and methods requiring paraffin or silicone oil to

be layered on top of the reservoir (Chayen, 1997) failed to yield single crystals, although needles were observed (Fig 5.4C).

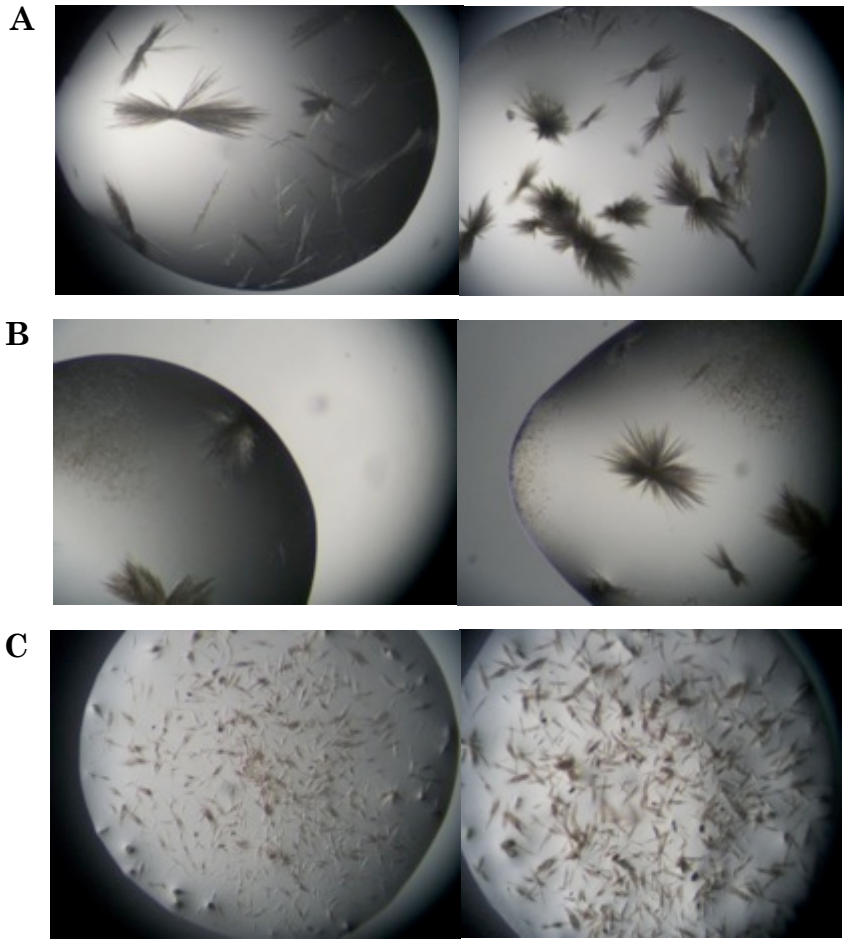

Figure 5.4 Examples of needle crystals grown from solutions of Gel-CL:actin. **(A)** Crystals obtained from 20%(v/v) MPEG550, 100 mM Tris-HCl, pH 8.0 (Fluka crystallization kit for protein complexes, precipitant no. 2). **(B)** Crystals obtained from screens near the conditions successfully used for G4-G6:actin [10%(w/v) PEG8000, 20%(v/v) glycerol, 100 mM HEPES buffer, pH 7.5]. **(C)** Crystals obtained after the seeding (Bergfors 1999) or the use of paraffin and silicone oil techniques (Chayen 1997).

5.4 Summary and discussion

Previous crystal data from preparations of GA_2 did not yield complete structures for one gelsolin bound to two actins, but produced only images of the G1-G3:actin portion of GA_2 (Wang, 2009). The structure of activated human G4-G6 reported here is the first to come from attempts to crystallize GA_2. This constitutes further evidence that the G4-G6 can be activated and stabilized without actin, yielding a structure essentially identical to that reported previously, but originating from different crystallization materials and conditions (PDB code 1P8X; Narayan *et al*., 2003).

Attempts to obtain crystals of Gel-CL:actin yielded only needle clusters not suitable for X-ray diffraction analysis. To date, use of seeding and alternative techniques have not improved the results.

Chapter 6

Gelsolin-specific Nanobodies

6.1 Background

The literature on gelsolin-specific nanobodies (GsnVHHs) suggests distinctive recognition of free *versus* actin-bound gelsolin (Van den Abbeele *et al.*, 2010). GsnVHH9 and GsnVHH11 bind to domains G1-G2 or G2-G3 in the N-terminal half of gelsolin (K_d ~5 nM) in a manner that is independent of Ca^{2+}, whereas GsnVHH3 and GsnVHH13 bind the C-terminal half (G4-G6) and specifically recognize the Ca^{2+}-activated form of gelsolin (K_d ~10 nM). In addition, GsnVHH11 blocks binding of the N-terminal half of gelsolin to actin, while GsnVHH13 does not have an effect on the interaction of the C-terminal half of gelsolin with actin. Moreover, GsnVHH11 protects FAF mutant gelsolin against proteolysis by furin. Lastly, GsnVHH3 and GsnVHH9 slow down degradation of gelsolin by contaminating proteases during up to at least eight days of incubation.

Crystal structures for gelsolin-specific nanobodies GsnVHH3, 9, and 13 (Fig. 1.9; PDB codes 2X1Q, 2X1P, and 2X1O, respectively) are available (Van den Abbeele *et al.*, 2010), but the structure of GsnVHH11, either alone or in a complex with its antigen, gelsolin, has not been solved. This thesis presents two crystal structures of GsnVHH11 (GsnVHH-F1 and GsnVHH-F2) in which

one of the loops critical for antigen binding (CDR1) adopts different conformations. Furthermore, to gain insight into how nanobodies interact with gelsolin, *in silico* protein-protein docking runs were performed, the results of which suggest possible modes of interaction. Finally, this thesis reports a new structure of GsnVHH9 that closely resembles a previously reported one obtained from different starting materials and crystallization conditions (Van den Abbeele *et al*., 2010).

6.2 Structures of GsnVHH11

6.2.1 Crystallization

Crystals of GsnVHH11 were obtained from protein solutions that contained Gel-NL and GsnVHH11 at a mole ratio of 1:2, respectively. Similar crystal forms grew when equilibrated against three different well buffers (Fig. 6.1): (1) 22.5%(w/v) PEG6000, 200 mM imidazole malate buffer, pH 8.5; (2) 30%(w/v) PEG8000, 100 mM HEPES, pH 8.2; and (3) 40%(w/v) PEG8000, 100 mM HEPES, pH 8.2. However, those grown against well buffer (3) were rod-shaped (Fig. 6.1C) and thicker than the ones grown against well buffers (1) and (2), which grew as thin plates (Figs. 6.1A and B).

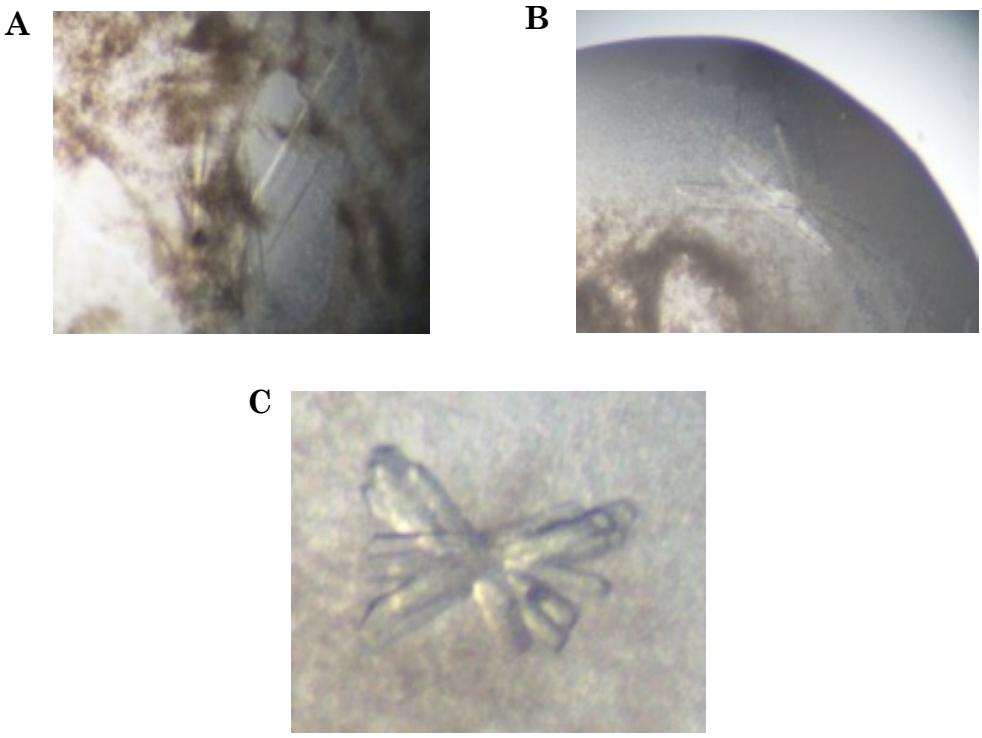

Figure 6.1 Examples of crystals grown from solutions containing Gel-NL and GsnVHH11. Crystals obtained using **(A)** 22.5%(v/v) PEG6000, 200 mM imidazole malate buffer, pH 8.5, **(B)** 30%(w/v) PEG8000, 100 mM HEPES, pH 8.2, and **(C)** 40%(w/v) PEG8000, 100 mM HEPES, pH 8.2.

6.2.2 Data collection and refinement

Four unique data sets were collected at different resolutions from four crystals that were cryoprotected using 15% (v/v) glycerol (Table A.4, Appendix A). The two data sets with the best resolution (2 Å) were collected using a crystal of the form depicted in Figures 6.1B and 6.1C, with the same space group $P2_12_12_1$, but the unit cell parameters were determined to be slightly different (Table 6.1; a = 33.9 Å, b = 55.4 Å, c = 150.5 Å, and α = β = γ = 90.0° *versus* a = 33.6 Å, b = 56.3 Å, c = 57.5 Å, and α = β = γ = 90.0°, respectively). The latter is shorter along the c-axis.

The structure of a crystal of the form depicted in Figure 6.1B was solved using molecular replacement methods with the structure of GsnVHH9 (PDB code 2X1P; Van den Abbeele *et al.*, 2010) as a search model in Phenix AutoMR in the PHENIX suite (Adams *et al.*, 2010). The structure reveals two molecules of GsnVHH11 (called GsnVHH11-F1) in the crystallographic asymmetric unit, but Gel-NL is not evident.

The structure of GsnVHH11-F1 then was used as the search model to solve the structure of a crystal of the form depicted in Figure 6.1C using Phenix_AutoMR. The structure reveals only one molecule of GsnVHH11 (called GsnVHH11-F2) in the crystallographic asymmetric unit. Both the structures GsnVHH11-F1 and GsnVHH11-F2 were refined using Phenix_Refine. Data collection and refinement statistics are presented in Table 6.1.

Table 6.1 Data collection and refinement statistics for GsnVHH11-F1 and GsnVHH11-F2

Statistic	**GsnVHH11-F1**	**GsnVHH11-F2**
Beamline	CMCF-08ID-1, CLS	BL13B1, NSRRC
Wavelength, Å	0.9795	1.000
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions	a = 33.9 Å, b = 55.4 Å, c = 130.5 Å α = β = γ = 90°	a = 33.6 Å, b = 56.3 Å, c = 57.5 Å α = β = γ = 90°
Resolution range, Å	34.2 - 2.0 (2.06 - 2.00)[d]	28.9 - 2.0 (2.26 - 2.00)[d]
Unique reflections	16,714 (2,346)	7865 (2,371)
Multiplicity	6.9 (6.7)	5.3 (5.1)
Completeness, %	97.2 (86.0)	97.0 (94.0)
Average I/σ	15.7 (2.9)	12.1 (3.0)
R_{merge}[a], %	8.0(29.1)	8.5(32.3)
R_{factor}[b], %	22.7 (34.8)	16.8 (15.7)
R_{free}[c], %	29.8 (37.1)	22.3 (23.7)
Molecules in asymmetric unit	2	1
Residue range	Mol 1: 2-128 Mol 2: 2-24, 31-128	2-129
Non-hydrogen atoms (waters)	1,881 (109)	980 (64)
Mean derived B- factor, Å^2	38	35
RMSD bonds, Å	0.008	0.007
RMSD angles, °	1.144	1.030

[a] $R_{merge} = (\Sigma | I - \langle I \rangle | / \Sigma \langle I \rangle)$

[b] $R_{factor} = (\Sigma | F_o | - | F_c | / \Sigma | F_o |)$

[c] based on 5% of the data

[d] highest resolution data shell

6.2.3 Structural analysis

The structures of GsnVHH11-F1 (Fig. 6.2A) and GsnVHH11-F2 (Fig. 6.2B) adopt the general nanobody fold, consisting of a pair of β-sheets, one four-stranded and one five-stranded, that serve as a framework to cluster the CDR1, CDR2 and CDR3 loops (conserved sequence regions), at one end of the molecule (Muyldermans *et al*., 2001; Padlan, 1994). A disulfide bond connects residues Cys22 and Cys96, which lie in β-strands near CDR1 and CDR3, respectively (Fig. 6.2).

Inspection of the GsnVHH11-F1 structure shows the CDR2 region to be traceable in Mol1 but not in Mol2. Superimposition of the two GsnVHH11 structures in the asymmetric unit yields an RMSD = 0.293 Å for α-carbon positions (Fig. 6.2A). The structure of GsnVHH11 closely resembles that of the published structure of GsnVHH9 (PDB code 2X1P; Van den Abbeele *et al*., 2010) and superimposition of the two structures results in an RMSD = 0.506 Å for α-carbon positions.

Superimposition of the structures of GsnVHH11-F1 and GsnVHH11-F2 shows the major differences to be in the CDR1 loop, which is missing in Mol2 in GsnVHH11-F1 and that in GsnVHH-F2 is more extended when compare with Mol 1 in GsnVHH11-F1 (Fig. 6.2C). RMSD = 0.401 and 0.301 Å, respectively, for α-carbon positions when GsnVHH11-F2 is superimposed onto GsnVHH11-F1 Mol1 and Mol2.

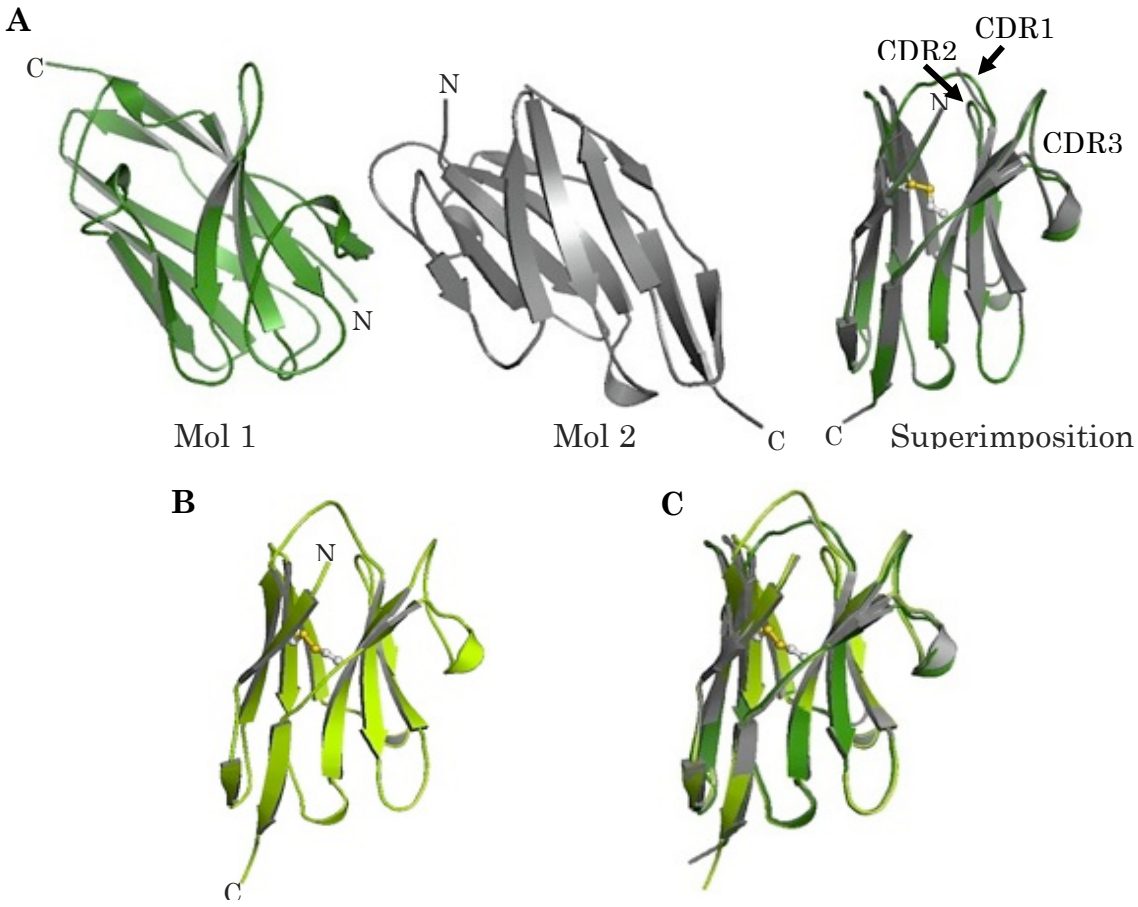

Figure 6.2 The structure of GsnVHH11. **(A)** The structure of GsnVHH11-F1. Both molecules of GsnVHH11 in the crystallographic asymmetric unit and their superimposition are shown. The structures are highly similar, with RMSD = 0.293 Å for α-carbon positions. The loop regions containing CDR1, CDR2 and CDR3 are labelled. **(B)** The structure of GsnVHH11-F2. **(C)** Superimposition of the structures in (A) and (B). RMSD = 0.401 and 0.301 Å for α-carbon positions when GsnVHH11-F2 is superimposed onto GsnVHH11-F1 Mol 1 and Mol 2, respectively. The major differences are found in the CDR1 loop, which is missing in Mol 2 in GsnVHH11-F1, and which in GsnVHH-F2 is more extended when compare with Mol 1 in GsnVHH11-F1. A disulfide bond between Cys22 and Cys96 is shown in ball and stick format.

6.3 Positioning GsnVHH11-F1 and GsnVHH11-F2 on G2-G3

GsnVHH11 binds to activated forms of gelsolin fragments G1-G2 and G2-G3 (Van den Abbeele *et al*., 2010), inferring that GsnVHH11 binds domain G2. Also, from the structure of G1-G3:actin, G1 is situated away from the complex formed by G2 and G3 on the surface of actin. For these reasons, G2-G3 was selected as a candidate for protein-protein docking studies. Atomic coordinates for G2-G3 were taken from the 2.8-Å resolution structure of human G1-G3:actin reported in Chapter 3 of this thesis.

Structures of lysozyme with its nanobody (PDB code 1MEL; Desmyter *et al*., 1996) and of β2-adrenoceptor with its nanobody (PDB code 3P0G; Rasmussen *et al*, 2011), show the nanobodies to plug their extended CDR3 regions into clefts on their target proteins, with stabilization provided by CDR1 and CDR2. It is likely, therefore, that binding specificity in nanobodies is conferred by these regions. This was used as a criterion in sorting the results of *in silico* docking runs.

The best ten clusters of solutions, ranked on the basis of minimum binding energy calculated using shape and electrostatic surface interactions that emerged from evaluation of the docking of GsnVHH11-F1 and G2-G3 were analyzed (Table B.1, Appendix B). Four out of 22 solutions, including solutions 4 (cluster 2), 7 (cluster 6), 9 (cluster 8), and 140 (cluster 10) were found to bind G2-G3 through the CDRs (Fig. 6.3). The binding of GsnVHH11 from solutions 4, 7 and 8 to G2-G3 is in a similar manner and the interaction

involves the binding of CDR1 and CDR3 in the groove near the G2-G3 linker. Similarly, GsnVHH11 from solution 140 uses CDR1 and CDR3 to bind exclusively on G2.

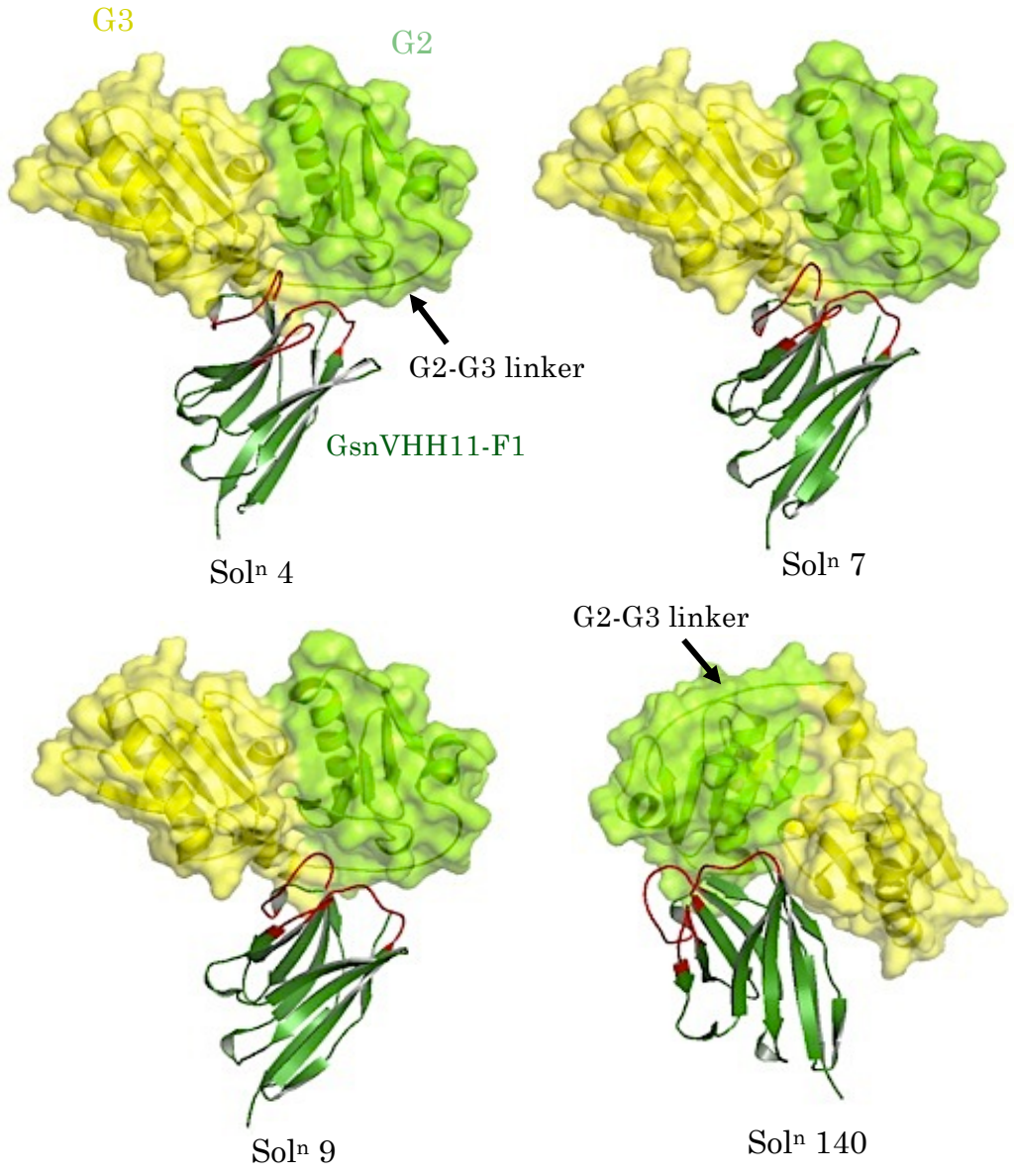

Figure 6.3 Four *in silico* protein-protein docking solutions for human G2-G3 and GsnVHH11-F1. The human G2-G3 surface is represented. The CDR1, CDR2 and CDR3 regions of the GsnVHH11-F1 are colored red.

Docking studies using G2-G3 and GsnVHH11-F2 yielded only two out of 28 solutions employed the CDRs to bind G2-G3 when selected from the top 10 clusters ranked on the basis of minimum binding energy that emerged from evaluation of the docking runs (Table B.2, Appendix B). These include solutions 1 (cluster 1) and 8 (cluster 5) (Fig. 6.4), which both involve binding the CDR1 and CDR3 regions in the groove near the G2-G3 linker. In addition, both of these solutions for interaction of GsnVHH11-F2 with G2-G3 are highly similar to solutions 4, 7, and 9 for the interaction of GsnVHH11-F1 with G2-G3 (Figs. 6.3 and 6.4).

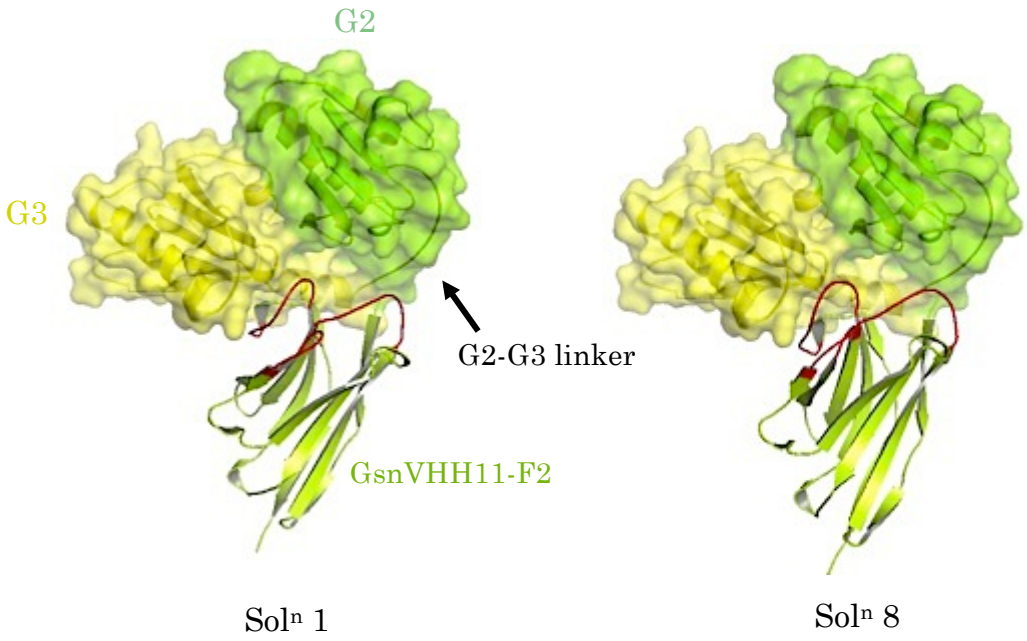

Figure 6.4 Two *in silico* protein-protein docking solutions for human G2-G3 and GsnVHH11-F2. The human G2-G3 surface is represented. The CDR1, CDR2 and CDR3 regions of the GsnVHH11-F2 are colored red.

6.4 Positioning GsnVHH13 on activated G4-G6

Activated G4-G6 binds actin independently of GsnVHH13 and GsnVHH13 binds gelsolin fragment G4-G5 but not G5-G6 (Van den Abbeele *et al.*, 2010). GsnVHH13 does not affect the nucleation or severing activities of gelsolin, suggesting that the binding site for GsnVHH13 on G4-G6 would not be the same as the actin-binding site. *In silico* docking runs were conducted holding activated gelsolin G4-G6 stationary while GsnVHH13 (PDB code 2X1O) was set to be the search ligand to find its possible binding site on G4-G6.

From the ten clusters that were most highly ranked for the docking of GsnVHH13 and activated G4-G6 (Table B.3, Appendix B), only one out of 27 solutions, solution 24 (cluster 10), was found to bind domains G4-G5 through interaction with the CDRs, specifically having CDR1 and CDR3 bind to the linker region between G4 and G5 (Fig. 6.5).

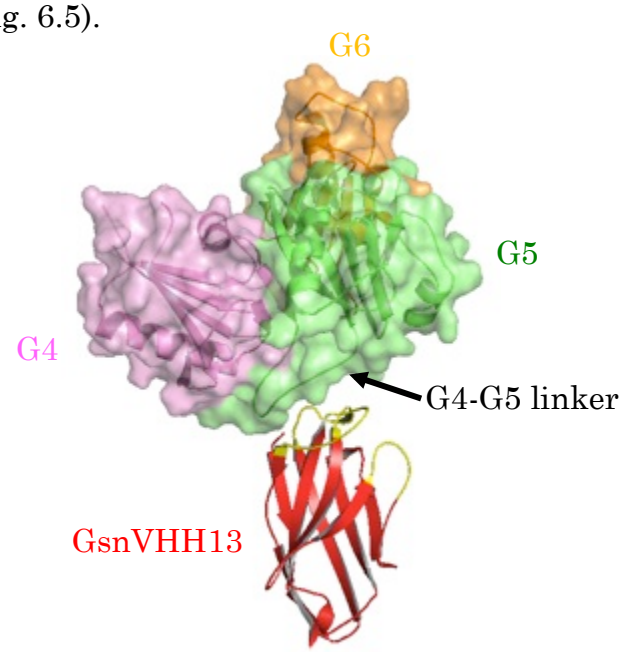

Figure 6.5 Docking GsnVHH13 (PDB code 2X1O) onto activated G4-G6. Docking solution 24 is shown. The G4-G6 surface is represented. The CDR1, CDR2 and CDR3 regions of GsnVHH13 are colored yellow.

6.5 Structure of GsnVHH9

6.5.1 Crystallization

Crystals of GsnVHH9 grew in solutions that were prepared using purified FL-gelsolin and GsnVHH9 at a mole ratio of 1:2, respectively (Fig. 6.6). The precipitant solution consisted of 33.0%(v/v) PEG600, 200 mM imidazole malate buffer, pH 5.5, at 4 °C. Crystallization conditions reported here differ from those previously reported (Van den Abbeele *et al*., 2010), which were: 30%(w/v) PEG1000, 100 mM $LiSO_4$, 100 mM phosphate/citrate buffer, pH 4.4, using the sitting drop vapor diffusion method at 20 °C.

Figure 6.6 Examples of crystals grown from solutions containing FL-gelsolin and GsnVHH9. The precipitant used was 33.0%(v/v) PEG600, 200 mM imidazole malate buffer, pH 5.5.

6.5.2 Data collection and refinement

The cryoprotectant used for these crystals was the well buffer supplemented with 20%(w/v) trehalose. One data set was collected from a single crystal that diffracted to a resolution of ~2.0 Å, with space group $P22_12_1$ and unit cell parameters a = 55.3 Å, b = 58.7 Å, c = 81.5 Å, and α = β = γ = 90.0°. The structure was solved by molecular replacement methods, using the previously reported structure of GsnVHH9 (PDB code 2X1P; Van den Abbeele *et al.*, 2010) as a search model in Phenix_AutoMR. The structure reveals two molecules of GsnVHH9 in the crystallographic asymmetric unit, but does not reveal any gelsolin. Furthermore, analysis of the unit cell volume revealed it to be too small to contain gelsolin. Refinement was carried out using Phenix_Refine. Data collection and refinement statistics are presented in Table 6.2.

The previously reported structure of GsnVHH9 was solved in space group P2 at a resolution of 1.1 Å with unit cell parameters a = 55.9 Å, b = 59.3 Å, c = 81.7 Å, and α = β = γ = 90.0°. Four molecules of GsnVHH9 were found in each crystallographic asymmetric unit (Van den Abbeele *et al.*, 2010).

Table 6.2 Data collection and refinement statistics for GsnVHH9

Statistic	**GsnVHH9**
Beamline	BL13B1, NSRRC
Wavelength, Å	1.000
Space group	$P22_12_1$
Unit cell dimensions	a = 55.3 Å, b = 58.7 Å, c = 81.5 Å α = β = γ = 90°
Resolution range, Å	29.4 - 1.97 (2.09 - 1.97)[d]
Unique reflections	18,741 (2,405)
Redundancy	4.5 (4.2)
Completeness, %	96.3 (69.0)
Average I/σ	12.1 (2.5)
R_{merge}[a], %	10.5 (32.8)
R_{factor}[b], %	27.3 (32.4)
R_{free}[c], %	29.5 (39.8)
Molecules in asymmetric unit	2
Residue range	Mol 1: 2-24, 31-126 Mol 2: 3-24, 31-126
Non-hydrogen atoms (waters)	1,798(75)
Mean derived B- factor, Å^2	35.6
RMSD bonds, Å	0.009
RMSD angles, °	1.261

[a] $R_{merge} = (\Sigma | I - \langle I \rangle | / \Sigma \langle I \rangle)$

[b] $R_{factor} = (\Sigma | F_o | - | F_c | / \Sigma | F_o |)$

[c] based on 5% of the data

[d] highest resolution data shell

6.5.3 Structural analysis

The overall structure of GsnVHH9 contains two β-sheets, one four-stranded and one five-stranded, that are formed by four conserved sequence regions. The CDR regions in the connecting loops between these β-sheets are clustered at one end of the structure, in which the overall folding resembles that for nanobodies in general (Muyldermans *et al.*, 2001; Padlan, 1994). However, part of the CDR1 region, containing residues Ser25 to Thr30, is not traceable in either molecule in the asymmetric unit. This region also was not traceable in the previously reported structure of this nanobody. The disulfide bond between residues Cys22 and Cys96 is evident. Superposition of the two molecules in the asymmetric unit shows them to be very similar, as judged by an RMSD = 0.309 Å for superposition of α-carbon positions (Fig. 6.7).

In addition, both structures closely resemble that previously reported for GsnVHH9, with RMSD = 0.440 Å and 0.223 Å for α-carbon positions when Mol 1 and Mol 2 from the current structure are superimposed, respectively, onto the earlier structure. While CDR1 loops were not present in any of the structures, the CDR2 and CDR3 loops in the present and previous structures, respectively, of GsnVHH9 are identical. Thus, the crystal structure of GsnVHH9 from this study confirms the previously reported structure of GsnVHH9 obtained using different starting materials and crystallization conditions (Van den Abbeele *et al.*, 2010).

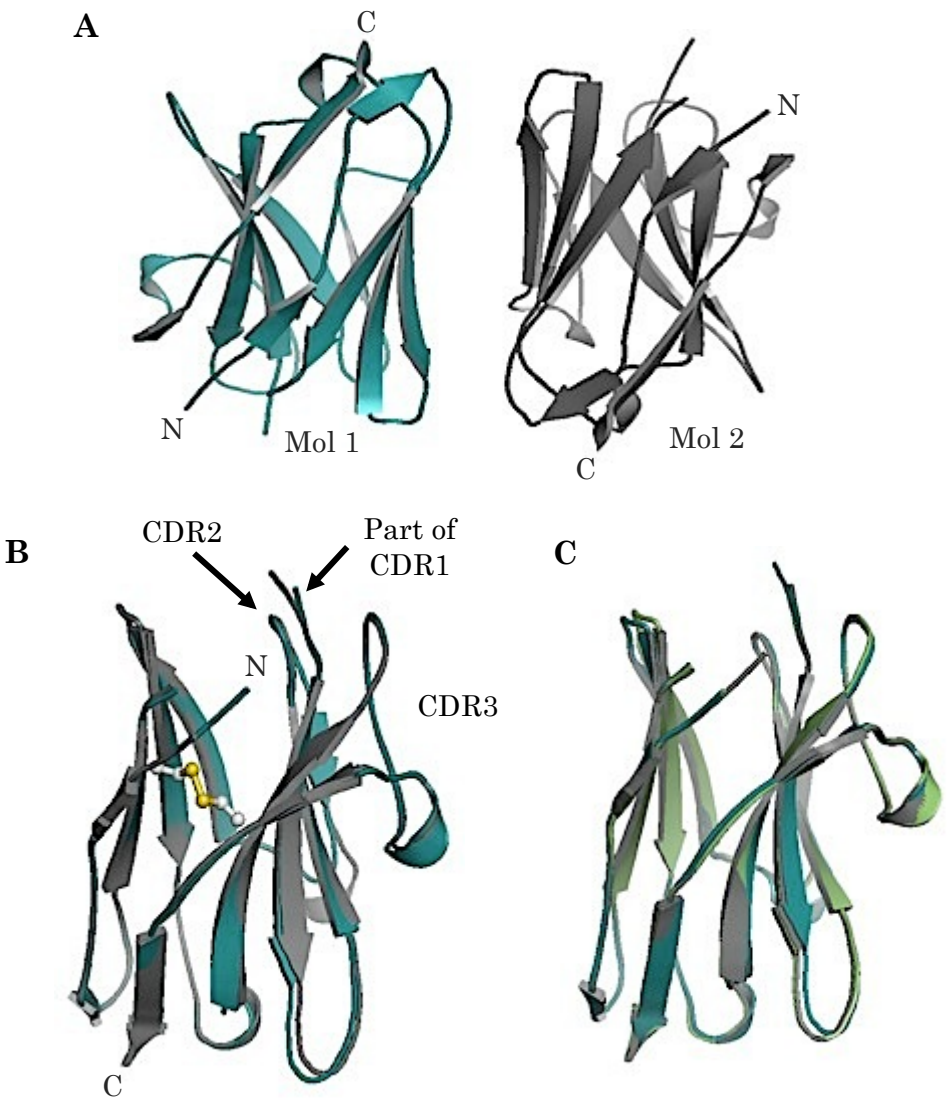

Figure 6.7 The structure of GsnVHH9. **(A)** The two molecules of GsnVHH9 in the crystallographic asymmetric unit. **(B)** Superimposition of the two molecules in (A). The locations of CDR1, CDR2 and CDR3 are labelled. CDR2 is not fully traceable in either molecule. The disulfide bond between Cys22 and Cys96 is shown in ball and stick format. **(C)** Superimposition of the two molecules of GsnVHH9 from (A) onto the previously reported GsnVHH9 structure (coloured lime green).

6.6 Summary and discussion

Surprisingly, the two structures of GsnVHH11 (GsnVHH-F1 and GcnVHH-F2) that were obtained from crystals of complex solutions of Gel-NL and GsnVHH11 grew in precipitants that differ in only 10%(w/v) of PEG8000 [30%(w/v) PEG8000, 100 mM HEPES, pH 8.2 versus 40%(w/v) PEG8000, 100 mM HEPES, pH 8.2] results in two structures of GsnVHH11. The structures of GsnVHH11-F1 and GsnVHH-F2 differ in the CDR1 loop in that one adopts a more extended form than the other. The two different conformations of the CDR1 loop might play a role in gelsolin binding.

Positioning GsnVHH11 on gelsolin domains G2-G3 by using *in silico* protein-protein docking suggested GsnVHH11 used CDR1 and CDR3 loops in binding to two possible binding regions of G2-G3: First, the region in the groove near the G2-G3 linker was obtained from docking runs with either GsnVHH11-F1 or GsnVHH11-F2 (Fig. 5.3 and 5.4). The other region is exclusively on G2 (Fig. 5.3, sol^{n} 140). The docking of GsnVHH13 (PDB code 2X1O) on active G4-G6 suggested the binding of CDR1 and CDR3 to the G4-G5 linker region (Fig. 5.5).

The crystal structure of GsnVHH9 from different starting materials and crystallization conditions presented in this study confirms the previously reported structure of GsnVHH9 (Van den Abbeele *et al.*, 2010).

Chapter 7

Conclusions

7.1 Gelsolin constructs

Gelsolin constructs, including FL-Gelsolin, Gel-NL and Gel-CL, were successfully cloned and purified. Their purities and identities were confirmed by SDS-PAGE and by matching their activities in actin polymerization and depolymerization assays to those previously published for gelsolin and a variety of its fragments.

The higher resolution (2.8 Å) structure of human G1-G3:actin was obtained from crystals that grew from concentrated solutions of Gel-NL:actin. In this human G1-G3:actin complex, the G1-G3 used to make the complex included a C-terminal extension (the G3-G4 linker polypeptide). However, no part of this extension was evident in the refined structure. It seems that the linker does not bind to actin in a sufficiently defined way to contribute to diffraction patterns, or it has been proteolytically removed during or prior to the crystallization process. The new structure does confirm the corresponding 3.0-Å resolution structure in two aspects related to the structural stability of gelsolin. First, the disulfide bridge between residues Cys188 and Cys201 that stabilizes G2 in the native gelsolin structure is present. Second, there is

a Ca^{2+} ion bound in the type-II Ca^{2+}-binding site of G2, the site that is disrupted in mutant gelsolins characteristic of the heritable disease FAF.

The structure of isolated activated G3 has not been published previously. This thesis reports that its structure, obtained from crystals grown from solution of activated G2-G4 in isolation, is not different in any significant way from that reported previously for activated G3 within complexes of G1-G3 and actin. A feature that differentiates activated G3 from the inactive form, aside from the occupied type-II calcium-binding site, is that the kinked long helix evident within the inactive form of intact gelsolin is straight in the activated form. The straightening of this helix weakens the hydrophobic interactions and salt-bridges between it and the long helix of domain G1 in the inactive structure, while simultaneously stabilizing G3 in the activated form.

While needle crystal clusters were obtained from solutions containing a Gel-CL:actin complex, they were of insufficient quality for X-ray diffraction analysis. However, crystals that grew from concentrated solutions of FL-Gelsolin:2actin yielded an activated structure of isolated G4-G6. The crystal structure of activated isolated G4-G6 reported in this thesis had very different origins from the previously reported structures, yet confirms that Ca^{2+} ions alone can fully activate G4-G6 in the absence of actin. Activated G4-G6 and actin can then dock essentially as rigid bodies, without further conformational adjustments.

7.2 Gelsolin nanobodies

Attempts to grow crystals of G1-G3 bound to a gelsolin-specific nanobody (GsnVHH11) yielded a crystal structure for GsnVHH11. Two conformations that differ in the CDR1 loop region were obtained, one being more extended than the other. *In silico* protein-protein docking runs suggest that GsnVHH11 binds to the grove region near the G2-G3 linker. In addition, docking of GsnVHH13 with activated G4-G6 suggests possible contact through G4. Attempts to grow crystals of gelsolin with two actins and gelsolin-specific nanobody 9 (GsnVHH9) yielded a crystal structure for GsnVHH9 that is in good agreement with the previous structure of GsnVHH9.

7.3 Future work

The crystal structure of activated full-length gelsolin has eluded all who have sought it. As a means to access this structure, the Burtnick and Robinson groups have made many attempts to crystallize complexes formed between full-length gelsolin and one, two or three actins. To date, these have resulted in structures for the fragments G1-G3:actin (Burtnick *et al*., 2004; Nag *et al*. 2009; Wang, 2009) and activated G4-G6 (the present work), but not for intact gelsolin alone or in a complex with actin. Trials continue with new crystallization conditions.

Nanobodies GsnVHH3 and GsnVHH9 are found to slow down proteolysis of gelsolin (Van den Abbeele *et al*., 2010). Perhaps crystallization of complexes made of GsnVHH3 and/or GsnVHH9 with full-length gelsolin should be pursued further to obtain a structure for activated full-length gelsolin. Results from similar crystallographic experiments using GsnVHH11 and GsnVHH13 to form complexes with gelsolin are required to decide which of the *in silico* protein-protein docking models discussed in Chapter 6 for the nanobodies with G1-G3 and G4-G6, are correct. In addition, use of other methods such as site-directed mutagenesis, crosslinking, small-angle X-ray scattering (SAXS), or protein footprinting (either by chemical or synchrotron radiolytic modification of amino acid side-chains) could help validate the *in silico* docking models.

The structure of the 50-residue long G3-G4 polypeptide linker has not been visualized in any crystal structure of gelsolin or its fragments. Binding of the linker to actin would determine the relative positioning of the two halves of gelsolin across an actin filament during the severing process. Efforts reported here to append the linker to the C-terminus of G1-G3 or the N-terminus of G4-G6 failed to identify how the linker extends across actin. Perhaps studies with synthetic sections of the linker in the presence of a crystallizable form of actin should be pursued in order to clarify this issue.

Bibliography

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

Adelman, M.R. and Taylor, E.W. (1969). Isolation of an actomyosin-like protein complex from slime mold plasmodium and the separation of the complex into actin- and myosin-like fractions. *Biochemistry* 8: 4964-4975.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002). Molecular Biology of the Cell, 4th edition, "The Cytoskeleton". New York: Garland Science.

Allen, P. G. (1997). Functional consequences of disulfide bond formation in gelsolin. *FEBS Lett* 401: 89–94.

Arbabi Ghahroudi, M., Desmyter, A., Wyns, L., Hamers, R. and Muyldermans, S. (1997). Selection and identification of single-domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett* 414: 521–526.

Ashish, Paine, M.S., Perryman, P.B., Yang, L., Yin, H.L. and Krueger, J.K. (2007). Global structure changes associated with Ca^{2+} activation of full-length human plasma gelsolin. *J Biol Chem* 282: 25884-25892.

Bergfors, T.M. (1999). Protein crystallization: techniques, strategies, and tips. California: International University Line.

Bryan, J. (1988). Gelsolin has three actin-binding sites. *J Cell Biol* 106: 1553-1562.

Bryan, J. and Hwo, S. (1986). Definition of an N-terminal actin-binding domain and a C-terminal Ca^{2+} regulatory domain in human brevin. *J Cell Biol* 102:1439-46.

Bubb, M.R., Govindasamy, L., Yarmola, E.G., Vorobiev, S.M., Almo, S.C., Somasundaram, T., Chapman, M.S., Agbandje-McKenna, M. and McKenna, R. (2002). Polylysine induces an antiparallel actin dimer that nucleates filament assembly - crystal structure at 3.5-angstrom resolution. *J Biol Chem* 277: 20999-21006.

Bugyi, B. and Carlier, M.F. (2010). Control of actin filament treadmilling in cell motility. *Annu Rev Biophys* 39: 449-470.

Burtnick, L.D. and Chan, K.W. (1980). Protection of actin against proteolysis by complex formation with deoxyribonuclease I. *Can J Biochem* 58: 1348-1354.

Burtnick, L.D., Koepf, E.K., Grimes, J., Jones, E.Y., Stuart, D.I., McLaughlin, P.J. and Robinson, R.C. (1997). The crystal structure of plasma

gelsolin: implications for actin severing, capping, and nucleation. *Cell* 90: 661-670.

Burtnick, L.D., Urosev, D., Irobi, E., Narayan, K. and Robinson, R.C. (2004). Structure of the N-terminal half of gelsolin bound to actin: roles in severing, apoptosis and FAF. *EMBO J* 23: 2713-2722.

Carlier, M.F., Valentin-Ranc, C., Combeau, C., Fievez, S. and Pantoloni, D. (1994). Actin polymerization: regulation by divalent metal ion and nucleotide binding, ATP hydrolysis and binding of myosin. *Adv Exp Med Biol* 358: 71-81.

Chaponnier, C., Janmey, P.A. and Yin, H. (1986). The actin filament-severing domain of plasma gelsolin. *J Cell Biol* 103: 1473-1481.

Chayen, N.E. (1997). The role of oil in macromolecular crystallization. *Structure* 5: 1269-1274.

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

Choe, H., Burtnick, L.D., Mejillano, M., Yin, H.L., Robinson, R.C. and Choe, S. (2002). The Calcium Activation of Gelsolin: Insights from the 3 Å Structure of the G4–G6/Actin Complex. *J Mol Biol* 324: 691-702.

Chumnarnsilpa, S., Loonchanta, A., Xue, B., Choe, H., Urosev, D., Wang, H., Lindberg, U., Burtnick, L.D. and Robinson, R.C. (2006). Calcium ion exchange in crystalline gelsolin. *J Mol Biol* 357: 773-782.

Collaborative Computational Project, Number 4. (1994). The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst D* 50: 760-763.

Conrath, K.E., Lauwereys, M., Galleni, M., Matagne, A., Frere, J.M., Kinne, J., Wyns, L. and Muyldermans, S. (2001). Beta-lactamase inhibitors derived from single-domain antibody fragments elicited in the camelidae. *Antimicrob Agents Chemother* 45: 2807–2812.

Conrath, K.E., Wernery, U., Muyldermans, S. and Nguyen, V.K. (2003). Emergence and evolution of functional heavy-chain antibodies in Camelidae. *Develop Comp Immunol* 27: 87–103.

Costantini, S., Colonna, G. and Facchiano, A.M. (2008). ESBRI: a web server for evaluating salt bridges in proteins. *Bioinformation* 3: 137-138.

Coue, M. and Korn, E.D. (1985). Interaction of plasma gelsolin with G-actin and F-actin in the presence and absence of calcium ions. *J Biol Chem* 260: 15033-15041.

Davies, J. and Riechmann, L. (1996). Single antibody domains as small recognition units: design and in vitro antigen selection of camelid, human VH domains with improved protein stability. *Protein Eng* 9: 531–537.

Decanniere, K., Desmyter, A., Lauwereys, M., Ghahroudi, M.A., Muyldermans, S. and Wyns, L. (1999). A single-domain antibody fragment in complex with RNase A: non-canonical loop structures and nanomolar affinity using two CDR loops. *Structure* 7: 361–370.

De Genst, E., Saerens, D., Muyldermans, S. and Conrath, K. (2006a). Antibody repertoire development in camelids. *Develop Comp Immunol* 30: 187–198.

De Genst, E., Silence, K., Decanniere, K., Conrath, K., Loris, R., Kinn, J., Muyldermans, S. and Wyns, L. (2006b). Molecular basis for the preferential cleft recognition by dromedary heavy-chain antibodies. *Proc Natl Acad Sci USA* 103: 4586–4591.

DeLano, W.L. "The PyMOL Molecular Graphics System." DeLano Scientific LLC, San Carlos, CA, USA [online]. Available at: http://www.pymol.org

Desmyter, A., Transue, T. R., Ghahroudi, M.A., Thi, M.H., Poortmans, F., Hamers, R., Muyldermans, S. and Wyns, L. (1996). Crystal structure of a camel single-domain VH antibody fragment in complex with lysozyme. *Nat Struct Biol* 3: 803–811.

Doi, Y. and Frieden, C. (1984). Actin polymerization. The effect of brevin on filament size and rate of polymerization. *J Biol Chem* 259: 11868-11875.

Dominguez, R. (2004). Actin-binding proteins-a unifying hypothesis. *Trends Biochem Sci* 29: 572-578.

Dominguez, R. and Holmes, K. C. (2011). Actin structure and function. *Annu Rev Biophys* 40: 169-186.

Emsley, P and Cowtan, K. (2004). COOT: model-building tools for molecular graphics. *Acta Cryst D* 60: 2126-2132.

Fujii, T., Iwane, A.H., Yanagida, T. and Namba, K. (2010). Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* 467: 724-728.

Galkin, V.E., Van Loock, M.S., Orlova, A. and Egelman, E.H. (2002). A new internal mode in F-actin helps explain the remarkable evolutionary conservation of actin's sequence and structure. *Curr Biol* 12: 570-575.

Garman, E. (2003). 'Cool' crystals: macromolecular cryocrystallography and radiation damage. *Curr Opin Struct Biol* 13: 545-551.

Geng, Y.J., Azuma, T., Tang, J. X., Hartwig, J.H., Muszynski, M., Wu, Q., Libby, P. and Kwiatkowski, D.J. (1998). Caspase-3-induced gelsolin fragmentation contributes to actin cytoskeletal collapse, nucleolysis, and apoptosis of vascular smooth muscle cells exposed to proinflammatory cytokines. *Eur J Cell Biol* 77: 294-302.

Gill, S.C. and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182: 319-326.

Gonzalez, A and Tsai, Y. (2010). AutoXDS. Stanford Synchrotron Radiation Lightsource (SSRL). Available: http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script

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

Grochulski, P., Fodje, M.N., Gorin, J., Labiuk, S.L. and Berg, R. (2011). Beamline 08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility. *J Synchrotron Rad* 18: 681-684.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Bajyana Songa, E., Bendahman, N. and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains. *Nature* 363: 446–448.

Hanson, J. and Lowy, J. (1963). The structure of F-actin and the actin filaments isolated from muscle. *J Mol Biol* 6: 46-60.

Harris, H. E. and Weeds, A. G. (1984). Plasma gelsolin caps and severs actin filaments. *FEBS Lett* 177: 184-188.

Hartwig, J.H., Bokoch, G.M., Carpenter, C.L., Janmey, P.A., Taylor, L.A., Toker, A. and Stossel, T.P. (1995). Thrombin receptor ligation and activated Rac uncap actin filament barbed ends through phosphoinositide synthesis in permeabilized human platelets. *Cell* 82: 643-653.

Hatano, S. and Oosawa, F. (1966). Isolation and characterization of plasmodium actin. *Biochem Biophys Acta* 127: 488-498.

Hertzog, M., van Heijenoort, C., Didry, D., Gaudier, M., Coutant, J., Gigant, B., Didelot, G., Preat, T., Knossow, M., Guittet, E. and Carlier, M. F. (2004). The beta-thymosin/WH2 domain: Structural basis for the

switch from inhibition to promotion of actin assembly. *Cell* 117: 611-623.

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

Irobi, E., Burtnick, L.D., Urosev,, D., Narayan, K., and Robinson, R.C. (2003) From the first to the second domain of gelsolin: a common path on the surface of actin? *FEBS Lett* 552: 86-90.

Irobi, E., Aguda, A.H., Larsson, M., Guerin, C., Yin, H.L., Burtnick, L.D., Blanchoin, L. and Robinson, R.C. (2004). Structural basis of actin sequestration by thymosin-beta 4: implications for WH2 proteins. *EMBO J* 23: 3599-3608.

Iwasa, M., Maeda, K., Narita, A., Maeda, Y. and Oda, T. (2008). Dual roles of Gln137 of actin revealed by recombinant human cardiac muscle alpha-actin mutants. *J Biol Chem* 283: 21045-21053.

Jacobson, G.R. and Rosenbusch, J.P. (1976). ATP binding to a protease-resistant core of actin. *Proc Natl Acad Sci USA* 73: 2742-2746.

Janmey, P.A., Iida, K., Yin, H.L. and Stossel, T.P. (1987). Polyphosphoinositide micelles and polyphosphoinositide-containing vesicles dissociate endogenous gelsolin-actin complexes and promote actin assembly from the fast-growing end of actin filaments blocked by gelsolin. *J Biol Chem* 262: 12228-12236.

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

Kamada, S., Kusano, H., Fujita, H., Ohtsu, M., Koya, R. C., Kuzumaki, N. and Tsujimoto, Y. (1998). A cloning method for caspase substrates that uses the yeast two-hybrid system: cloning of the antiapoptotic gene gelsolin. *Proc Natl Acad Sci USA* 95: 8532–8537.

Kazmirski, S.L., Isaacson, R.L., An, C., Buckle, A., Johnson, C.M., Daggett, V. and Fersht, A.R. (2002). Loss of a metal-binding site in gelsolin leads to familial amyloidosis-Finnish type. *Nat Struct Biol* 9: 112-116.

Kiselar, J.G., Janmey, P.A., Almo, S.C. and Chance, M.R. (2003). Visualizing the Ca^{2+}-dependent activation of gelsolin by using synchrotron footprinting. *Proc Natl Acad Sci USA* 100: 3942-3947.

Kolappan, S., Gooch, J.T., Weeds, A.G. and McLaughlin, P.J. (2003). Gelsolin domains 4-6 in active, actin-free conformation identifies sites of regulatory calcium ions. *J Mol Biol* 329: 85-92.

Kothakota, S., Azuma, T., Reinhard, C., Klippel, A., Tang, J., Chu, K., McGarry, T.J., Kirschner, M.W., Koths, K., Kwiatkowski, D.J. and Williams, L.T. (1997). Caspase-3-generated fragment of gelsolin: effector of morphological change in apoptosis. *Science* 278: 294-298.

Kouyama, T. and Mihashi, K. (1981). Fluorimetry study of N-(1-pyrenyl)iodoacetamide-labelled F-actin. Local structural change of

actin protomer both on polymerization and on binding of heavy meromyosin. *Eur J Biochem* 114: 33-38.

Korn, E.D., Carlier, M.F. and Pantaloni, D. (1987). Actin polymerization and ATP hydrolysis. *Science* 238: 638-644.

Kumar, S. and Nussinov, R. (1999) Salt bridge stability in monomeric proteins. *J Mol Biol* 293: 1241-1255.

Kumar, S. and Nussinov, R. (2002). Relationship between ion pair geometries and electrostatic strengths in proteins. *Biophys J* 83: 1595-1612.

Kwiatkowski, D.J., Janmey, P.A. and Yin, H. (1989). Identification of critical functional and regulatory domains in gelsolin. *J Cell Biol* 108: 1717-1726.

Kwiatkowski, D.J., Stossel, T.P., Orkin, S.H., Mole, J.E., Colten, H.R. and Yin, H.L. (1986). Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin-binding domain. *Nature* 323: 455-458.

Lauwereys, M., Ghahroudi, M.A., Desmyter, A., Kinne, J., Hölzer, W., De Genst, E., Wyns, L. and Muyldermans, S. (1998). Potent enzyme inhibitors derived from dromedary heavy-chain antibodies. *EMBO J* 17: 3512– 3520.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D. and Darnell, J. (2000). "The Actin Cytoskeleton", *in* Molecular Cell Biology. 4th edition. New York: W. H. Freeman.

Lohkamp, B. and Dobritzsch, D. (2008). A mixture of fortunes: the curious determination of the structure of *Escherichia coli* BL21 Gab protein. *Acta Cryst D* 64: 407-415.

Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M.D. and Ritchie, D. W. (2010). HexServer: an FFT-based protein-docking server powered by graphics processors. *Nucleic Acids Research* 38: 445-449.

Maury, C.P. (1991). Gelsolin-related amyloidosis. Identification of the amyloid protein in Finnish hereditary amyloidosis as a fragment of variant gelsolin. *J Clin Invest* 87: 1195 - 1199.

Margarit, S.M., Davidson, W., Frego, L. and Stebbins, C. E. (2006). A steric antagonism of actin polymerization by a salmonella virulence protein. *Structure* 14: 1219-1229.

McGough, A.M., Staiger, C.J., Min, J.K. and Simonetti, K.D. (2003). The gelsolin family of actin regulatory proteins: modular structures, versatile functions. *FEBS Lett* 552: 75-81.

McLaughlin, P.J., Gooch, J.T., Mannherz, H.G. and Weeds, A.G. (1993). Structure of gelsolin segment-1-actin complex and the mechanism of filament severing. *Nature* 364: 685-692.

Mustard, D. and Ritchie, D.W. (2005). Docking essential dynamics eigenstructures. *Proteins: Struct Funct Bioinf* 60: 269-274.

Murshudov, G., Vagin, A. and Dodson, E. (1996). "Application of Maximum Likelihood Refinement" *in* The refinement of protein structures.

Proceedings of Daresbury Study Weekend.

Murshudov, G.N., Vagin, A. and Dodson, E. J. (1997). Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst D* 53: 240-255.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.., Long, F. and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst D* 67: 355-367.

Muyldermans, S., Baral, T.N., Retamozzo, V.C., De Baetselier, P., De Genst, E., Kinne, J., Leonhardt, H., Magez, S., Nguyen, V.K., Revets, H., Rothbauer, U., Stijlemans, B., Tillib, S., Wernery, U., Wyns, L., Hassanzadeh-Ghassabeh, G. and Saerens, D. (2009). Camelid immunoglobulins and nanobody technology. *Vet Immunol Immunopathol* 128: 178-183.

Muyldermans, S., Cambillau, C. and Wyns, L. (2001). Recognition of antigens by single-domain antibody fragments: the superfluous luxury of paired domains. *TIBS* 26: 230-235.

Muyldermans, S. and Lauwereys, M. (1999). Unique single-domain antigen binding fragments derived from naturally occurring camel heavychain antibodies. *J. Mol. Recognit* 12: 1-10.

Nag, S., Ma, Q., Wang, H., Chumnarnsilpa, S., Lee, W.L., Larsson, M., Kannan, B., Hernandez-Valladares, M., Burtnick, L.D. and Robinson,

R.C. (2009). Ca^{2+} binding by domain 2 plays a critical role in the activation and stabilization of gelsolin. *Proc Natl Acad Sci USA* 106: 13713-13718.

Narayan, K., Chumnarnsilpa, S., Choe, H., Irobi, E., Urosev, D., Lindberg, U., Schutt, C. E., Burtnick, L.D. and Robinson, R.C. (2003). Activation in isolation: exposure of the actin-binding site in the C-terminal half of gelsolin does not require actin. *FEBS Lett* 552: 82-85.

Nguyen, V.K., Desmyter, A. and Muyldermans, S. (2001). Functional heavychain antibodies in Camelidae. *Adv Immunol* 79: 261–296.

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

Otterbein, L.R., Cosio, C., Graceffa, P. and Dominguez, R. (2002). Crystal structures of the vitamin D-binding protein and its complex with actin: Structural basis of the actin-scavenger system. *Proc Natl Acad Sci USA* 99: 8003-8008.

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

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol A* 276: 307-326.

Padlan, E.A. (1994). Anatomy of the antibody molecule. *Molec Immunol* 31: 169–217.

Pak, C.W., Flynn, K.C. and Bamburg, J.R. Actin-binding proteins take the reins in growth cones. *Nat Rev Neurosci* 9: 136-147.

Pardee, J.D. and Spudich, J.A. (1982). Purification of muscle actin. *Methods Enzymol* 85: 164-181.

Perrin, B.J. and Ervasti, J.M. (2010). The actin gene family: function follows isoform. *Cytoskeleton* 67: 630-634.

Pflugrath, J.W. (2004). Macromolecular cryocrystallography—methods for cooling and mounting protein crystals at cryogenic temperatures. *Methods* 34: 415-423.

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

Pollard, T.D. and Berro, J. (2009). Mathematical models and simulations of cellular processes based on actin filaments. *J Biol Chem* 284: 5433-5437.

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

Pope, B.J., Gooch, J.T. and Weeds, A.G. (1997). Probing the effects of calcium on gelsolin. *Biochemistry* 36: 15848-15855.

Popp, D., Lednev, V.V. and Jahn, W. (1987). Methods of preparing well-orientated sols of F-actin containing filaments suitable for X-ray diffraction. *J Mol Biol* 197: 679-684.

Popp, D. and Robinson, R.C. (2011). Many ways to build an actin filament. *Mol Microbiol* 80: 300-308.

Rasmussen, S.G., Choi, H.J., Fung, J.J., Pardon, E., Casarosa, P., Chae, P.S., Devree, B.T., Rosenbaum, D.M., Thian, F.S., Kobilka, T.S., Schnapp, A., Konetzki, I., Sunahara, R.K., Gellman, S.H., Pautsch, A., Steyaert, J., Weis, W.I. and Kobilka, B.K. (2011). Structure of a nanobody-stabilized active state of the β_2-adrenoceptor. *Nature* 469: 175-180.

Ritchie, D.W. (2003). Evaluation of protein docking predictions using Hex 3.1 in capri rounds 1 and 2. *Proteins: Struct Funct Genet* 52: 98-106.

Ritchie, D.W. (2005). High order analytic translation matrix elements for real space six-dimensional polar Fourier correlations. *J Appl Cryst* 38: 808-818.

Ritchie, D.W. (1998). Parametric Protein Shape Recognition. PhD Thesis, Departments of Computing Science and Molecular & Cell Biology, University of Aberdeen.

Ritchie, D.W. and Kemp, G.J.L. (1999). Fast computation, rotation, and comparison of low-resolution spherical harmonic molecular surfaces, *J Comp Chem* 20: 383-395.

Ritchie, D.W. and Kemp G.J.L. (2000). Protein docking using spherical polar Fourier correlations. *Proteins: Struct Funct Genet* 39: 178-194.

Ritchie, D.W., Kozakov, D. and Vajda, S. (2008). Accelerating and focusing protein-protein docking correlations using multi-dimensional rotational

FFT generating functions. *Bioinformatics* 24: 1865-1873.

Ritchie, D.W. and Venkatraman, V. (2010). Ultra-Fast FFT protein docking on graphics processors. *Bioinformatics* 26: 2398-2405.

Robinson, R.C., Mejillano, M., Le, V.P., Burtnick, L.D., Yin, H.L. and Choe, S. (1999). Domain movement in gelsolin: A calcium-activated switch. *Science* 286: 1939-1942.

Rodgers, D.W. (1994). Cryocrystallography. *Structure* 2: 1135-1140.

Rossmann, M.G. (1972). The molecular replacement method. New York: Gordon & Breach.

Rossmann, M.G. (1990). The molecular replacement method. *Acta Cryst A* 46: 73-82.

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

Spudich, J.A. and Watt S. (1971). The regulation of rabbit skeletal muscle contraction. I. Biochemical studies of the interaction of the tropomyosin-troponin complex with actin and the proteolytic fragments of myosin. *J Biol Chem* 246: 4866-4871.

Straub, F.B. (1942). Actin. *Stud Inst Med Chem Univ Szeged* II: 3-15.

Straub, F.B. (1943). Actin, II. *Stud Inst Med Chem Univ Szeged* III: 23-27.

Sun, H.Q., Wooten, D.C., Janmey, P.A. and Yin, H.L. (1994). The actin side-binding domain of gelsolin also caps actin filaments. Implications for actin filament severing. *J Biol Chem* 269: 9473-9479.

Tirion, M.M., ben Avraham, D., Lorenz, M. and Holmes, K.C. (1995). Normal modes as refinement parameters for the F-actin model. *Biophys J* 68: 5-12.

Van den Abbeele, A., De Clercq, S., De Ganck, A., De Corte, V., Van Loo, B., Soror, S.H., Srinivasan, V., Steyaert, J., Vandekerckhove, J. and Gettemans, J. (2010). A llama-derived gelsolin single-domain antibody blocks gelsolin-G-actin interaction. *Cell Mol Life Sci* 67: 1519-1535.

Vagin, A. and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 30: 1022-1025.

Vagin, A. and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Cryst D* 66: 22-25.

Vasconcellos, C.A. and Lind, S.E. (1993). Coordinated inhibition of actin-induced platelet aggregation by plasma gelsolin and vitamin D-binding protein. *Blood* 82: 3648-3657.

Wang, H. (2009). Structural studies of actin and actin-binding proteins. Ph.D. Thesis. University of British Columbia: Canada.

Wang, H., Robinson, R.C. and Burtnick, L.D. (2010). The structure of native G-actin. *Cytoskeleton* 67: 456-465.

Way, M., Gooch, J., Pope, B. and Weeds, A. G. (1989). Expression of human plasma gelsolin in *Escherichia coli* and dissection of actin binding sites by segmental deletion mutagenesis. *J Cell Biol* 109: 593-605.

Way, M., Pope, B., Gooch, J., Hawkins, M. and Weeds, A.G. (1990). Identification of a region in segment-1 of gelsolin critical for actin binding. *EMBO J* 9: 4103-4109.

Wegner, A. (1976). Head to tail polymerization of actin. *J Mol Biol* 108: 139-150.

Yin, H. (1999). *In* Guidebook to the cytoskeletal and Motor Proteins (Kreis, T. and Vale, R., Eds), pp. 99-102, Oxford University Press, New York.

Yin, H., Hartwig, J.H., Maruyama, K. and Stossel, T.P. (1981). Ca^{2+} control of actin filament length. Effects of macrophage gelsolin on actin polymerization. *J Biol Chem* 256: 9693-9697.

Yin, H.L., Janmey, P.A. and Schleicher, M. (1990). Severin is a gelsolin prototype. *FEBS Lett* 264: 78-80.

Yin, H.L. and Stossel, T.P. (1979). Control of cytoplasmic actin gel-sol transformation by gelsolin, a calcium-dependent regulatory protein. *Nature* 281: 583-586.

Appendix A

Datasets for Crystals of G1-G3:actin, Activated G4-G6, and GsnVHH11

This appendix lists all data sets that were collected from the crystals grown from:

1) Gel-NL:actin and Gel-NL:2actin, which were solved to contain a G1-G3:actin complex (Tables A.1 and A.2; Chapter 4, section 4.2)

2) GA_2, which were found to contain activated G4-G6 (Table A.3; Chapter 5, section 5.2)

3) Gel-NL:GsnVHH11, which were found to contain only GsnVHH11 (Table A.4; Chapter 6, section 6.2)

4) GA_2 complex that were unable to be solved by molecular replacement methods using gelsolin domains G2, G3, and G4. Searches using space group and unit cell dimension parameters against the PDB database (http://www.pdb.org) showed the data did not agree with any known structure. (Table A.5; Chapter 5, section 5.3)

The precipitant used in crystallization, beamline used in data collection, space group, resolution, and unit cell dimensions for each particular data set are shown.

Table A.1 Data sets for G1-G3:actin that were obtained from crystals grown from Gel-NL:actin solutions

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
1	1.2%(w/v) PEG8000, 1.2mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	4.0 Å	a = 100.3 Å b = 147.8 Å c = 147.2 Å α = β = γ = 90°
2	6.4%(w/v) PEG8000, 6.4mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	2.8 Å	a = 100.5 Å b = 147.9 Å c = 147.0 Å α = β = γ = 90°
3	6.4%(w/v) PEG8000, 6.4 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	3.0 Å	a = 99.8 Å b = 147.9 Å c = 148.1 Å α = β = γ = 90°
4	6.4%(w/v) PEG8000, 6.4 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	4.3 Å	a = 102.2 Å b = 148.3 Å c = 148.2 Å α = β = γ = 90°

Table A.2 Data sets for G1-G3:actin that were obtained from crystals grown from Gel-NL:2actin solutions

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
1	4.0%(w/v) PEG8000, 4.0 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P6_522$	7.5 Å	a = 145.8 Å b = 145.8 Å c = 373.5 Å α = β = 90° γ = 120°
2	1.0%(w/v) PEG8000, 1.0 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	3.5 Å	a = 100.0 Å b = 146.5 Å c = 147.1 Å α = β = γ = 90°
3	2.0%(w/v) PEG8000, 2.0 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	5.2 Å	a = 100.0 Å b = 147.7 Å c = 146.9 Å α = β = γ = 90°
4	2.1%(w/v) PEG8000, 2.1 mM $CaCl_2$, 200mM NaOAc, pH 4.7	CMCF-08B1-1, CLS	$P2_12_12_1$	2.8 Å	a = 102.3 Å b = 147.6 Å c = 148.8 Å α = β = γ = 90°
5	1.5%(w/v) PEG8000, 1.5 mM $CaCl_2$, 200mM NaOAc, pH 4.7	CMCF-08B1-1, CLS	$P2_12_12_1$	3.5 Å	a = 102.3 Å b = 147.6 Å c = 148.8 Å α= β = γ = 90°

Table A.2 (cont.) Data sets for G1-G3:actin that were obtained from crystals grown from Gel-NL:2actin solutions

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
6	1.5%(w/v) PEG8000, 1.5 mM $CaCl_2$, 200mM NaOAc, pH 4.7	BL13B1, NSRRC	$P2_12_12_1$	2.9 Å	a = 101.5 Å b = 147.0 Å c = 148.0 Å α = β = γ = 90°

Table A.3 Data sets for activated G4-G6 that were obtained from crystals grown from GA_2 solutions

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
1	25.0%(w/v) PEG1500, 100mM MIB pH 6.0	BL13B1, NSRRC	$P2_12_12_1$	2.43 Å	a = 84.9 Å b = 88.6 Å c = 156.2 Å α = β = γ = 90°
2	25.0%(w/v) PEG1500, 100mM SPG buffer, pH 7.0,	BL13B1, NSRRC	$P2_12_12_1$	2.42 Å	a = 84.9 Å b = 89.6 Å c = 155.4 Å α = β = γ = 90°
3	25.0%(w/v) PEG1500, 100mM MIB pH 6.0,	CMCF-08ID1, CLS	$P2_12_12_1$	2.55 Å	a = 85.0 Å b = 87.3 Å c = 156.8 Å α = β = γ = 90°
4	25.0%(w/v) PEG1500, 100mM MIB pH 4.0	CMCF-08ID1, CLS	$P2_12_12_1$	2.73 Å	a = 84.0 Å b = 86.0 Å c = 155.0 Å α = β = γ = 90°
5	25.0%(w/v) PEG1500, 100mM MIB pH 5.5	CMCF-08B1-1, CLS (Remote)	$P2_12_12_1$	2.25 Å	a = 84.9 Å b = 88.3 Å c = 156.2 Å α = β = γ = 90°

Table A.4 Data sets for GsnVHH11 that were obtained from the crystals grown from Gel-NL:GsnVHH11 solutions

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
1	30.0%(w/v) PEG8000, 100mM HEPES, pH 8.2	CMCF-08ID1, CLS	$P2_12_12_1$	2.0 Å	a = 33.9 Å b = 55.4 Å c = 130.5 Å α = β = γ = 90°
2	30.0%(w/v) PEG8000, 100mM HEPES, pH 8.2	CMCF-08ID1, CLS	$P2_12_12_1$	2.1 Å	a = 33.9 Å b = 55.4 Å c = 130.5 Å α = β = γ = 90°
3	22.5%(w/v) PEG6000, 200mM Imidazole Malate, pH 8.5	CMCF-08B1-1, CLS	$P2_12_12_1$	3.0 Å	a = 33.9 Å b = 55.4 Å c = 130.5 Å α = β = γ = 90°
4	40.0%(w/v) PEG8000, 100mM HEPES, pH 8.2	BL13B1, NSRRC	$P2_12_12_1$	2.0 Å	a = 33.5 Å b = 56.3 Å c = 57.3 Å α = β = γ = 90°

Table A.5 Data sets obtained from crystals grown from GA_2 complex solutions that are still unsolved

Data set	Precipitant	Beamline	Space group	Resolution	Unit cell dimensions
1	20.0%(w/v) PEG3350, 200mM NaI, 100mM Bis-Tris propane, pH 8.5	BL13B1, NSRRC	P622	4.0 Å	a = 78.0 Å b = 78.0 Å c = 115.0 Å α = β = 120° γ = 90°
2	20.0%(w/v) PEG3350, 200mM KSCN, 100mM Bis-Tris propane, pH 7.5	CMCF-08ID1, CLS	C121	3.4 Å	a = 272.5 Å b = 157.5 Å c = 115.3 Å α = β = γ = 90°
3	20.0%(w/v) PEG3350, 200mM KSCN, 100mM Bis-Tris propane, pH 7.5	CMCF-08B1-1, CLS (Remote)	P622	2.85 Å	a = 155.3 Å b = 155.3 Å c = 113.5 Å α = β = γ = 90°
4	20.0%(w/v) PEG3350, 200mM KSCN, 100mM Bis-Tris propane, pH 7.5	CMCF-08B1-1, CLS (Remote)	P1	2.88 Å	a = 115.1 Å b = 157.4 Å c = 157.6 Å α = 60°, β = γ = 90°
5	20.0%(w/v) PEG3350, 200mM KSCN, 100mM Bis-Tris propane, pH 7.5	BL13B1, NSRRC	P622	3.5 Å	a = 157.0 Å b = 157.0 Å c = 114.4 Å α = β = 90° γ = 120°

Appendix B

In Silico Protein Docking Results

This appendix presents tables of the top ten clusters from *in silico* protein-docking using Hex software for G2-G3 with GsnVHH11-F1 and GsnVHH11-F2 (Chapter 6, section 6.3) (Table B.1 and B.2, respectively), as well as for activated G4-G6 with GsnVHH13 (PDB code 2X1O) (Chapter 6, section 6.4) (Table B.3).

E-value represents the total calculated pseudo interaction energy (as defined by the Hex software) of the system with units of kJ/mol, in which a negative score is favorable.

Table B.1 The top 10 clusters from docking human G2-G3 with GsnVHH11-F1

Cluster	Solution	E-value
1	1	-710.5
1	47	-494.3
1	175	-440.2
1	179	-439.0
2	2	-616.4
2	4	-609.7
3	3	-611.8
3	16	-554.9
4	5	-598.9
5	6	-594.5
5	161	-442.5
6	7	-572.9
6	53	-484.8
6	114	-456.7
7	8	-572.2
7	43	-497.3
8	9	-569.5
9	10	-565.5
9	101	-461.6
10	11	-562.4
10	85	-466.2
10	140	-449.6

Table B.2 The top 10 clusters from docking human G2-G3 with GsnVHH11-F2

Cluster	Solution	E-value
1	1	-642.7
1	4	-563.7
1	5	-558.0
1	30	-502.0
2	3	-575.9
2	9	-536.8
2	46	-484.0
2	90	-453.3
3	6	-557.6
3	39	-489.9
3	68	-470.4
4	7	-552.5
4	26	-508.3
4	41	-488.0
4	52	-478.8
4	99	-447.5
5	8	-540.4
5	17	-518.8
6	10	-529.1
7	11	-527.9
7	14	-524.9
8	12	-527.0
8	16	-521.1
9	13	-525.7
9	70	-468.6
10	15	-522.7
10	36	-491.6
10	51	-479.6

Table B.3 The top 10 clusters from docking activated G4-G6 with GsnVHH13 (PDB code 2X1O)

Cluster	Solution	E-value
1	1	-669.8
1	2	-652.1
1	5	-624.2
2	3	-648.7
2	6	-624.2
2	15	-580.6
2	21	-551.9
2	82	-502.0
3	4	-643.1
4	7	-612.7
5	8	-611.0
5	16	-579.8
5	146	-472.3
6	9	-608.1
6	113	-483.6
6	152	-470.2
7	10	-605.8
7	11	-602.9
7	18	-564.6
7	58	-518.3
8	12	-599.8
8	39	-532.6
9	13	-591.2
9	34	-537.6
9	138	-475.3
10	14	-589.2
10	24	-547.6