2010

Structural characterization of cytoskeleton regulating protein villin and its C-terminal modular domains

Danielle A. Pfaff
Western Washington University

Structural Characterization of Cytoskeleton Regulating Protein Villin and its C-terminal Modular Domains

By

Danielle A. Pfaff

Accepted in Partial Completion

Of the Requirements for the Degree

Master of Science

Moheb A. Ghali, Dean of the Graduate School

ADVISORY COMMITTEE

Chair, Dr. Serge L. Smirnov

Dr. P. Clint Spiegel

Dr. Spencer Anthony-Cahill

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permission from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Danielle A. Pfaff

August 3, 2010

STRUCTURAL CHARACTERIZATION OF CYTOSKELETON REGULATING PROTEIN VILLIN AND ITS C-TERMINAL MODULAR DOMAINS

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Danielle A. Pfaff
July 2010

Abstract

Villin is a modular protein that regulates F-actin bundles in the microvilli of absorptive epithelial cells in the intestine. At low (10-100 nM) calcium levels, Villin is an F-actin bundling agent supporting the specialized brush border membrane of the absorptive epithelium. At intermediate micromolar calcium levels, Villin nucleates and caps the barbed ends of F-actin and in high (> 100 μM) calcium Villin is an F-actin severing agent (Bretsher & Weber, 1980; Glenney et al., 1980, 1981; Mooseker et al. 1980). The amino acid sequence of Villin has seven modular domains. The first six Villin domains (D1-D6) form a "core" of ~50% sequence identity with Gelsolin; and contain a Ca^{2+}-dependent actin-binding site associated with the D1-D3 fragment. The last domain, Villin's unique C-terminal headpiece (HP), contains the other F-actin binding site, which is Ca^{2+}-independent (Bretsher & Weber, 1980; Glenney et al., 1980, 1981; Mooseker et al. 1980). Recent investigation by Nuclear Magnetic Resonance (NMR) Spectroscopy and Negative-Stain Electron Microscopy (EM) of the backbone dynamics and actin-binding of Villin's D6-HP, 208-residue, C-terminal modular fragment, revealed that: a) folded domains D6 and HP are interacting only via a largely unfolded 40-residue linker, and b) at millimolar calcium levels, the monomeric D6-HP fragment bundles F-actin and has two actin binding sites; one, which is previously known on HP, and the other is novel, cryptic and Ca^{2+}-dependent, associated with domain D6 or the linker (Smirnov et al., 2007).

We have investigated how the domain structure, domain-domain and linker-domain interactions in D6-HP fragment of Villin define its actin regulation properties.

Toward this goal, we are: a) making the D6 and D6-HP NMR samples; b) determining the NMR resonance assignment of isolated D6; and c) elucidating the solution structure of D6 domain in isolation and within the D6-HP fragment. Our NMR data indicate that the D6 protein fragment in isolation likely adopts a Gelsolin-like fold and that HP and D6 structures in isolation resemble those in the context of the larger modular fragment D6-HP. The potential effect of the linker on the D6 and HP domains structure is exemplified by the noticeable chemical shift differences for residue 84 of D6 and residue 166 of HP (^{15}N-HSQC spectrum of D6-HP vs. D6 and HP in isolation). These two positions are ~23 residues away from either end of the linker and located on the surface of these domains.

In the absence of calcium, Gelsolin adopts a compact, inactive conformation stabilized by the 12-residue C-terminal helix. This helix was suggested to keep together Gelsolin domains D2 and D6 as a "latch" closed in low calcium and released at higher calcium levels (Robinson et al., 1999). Our ensuing structural study of D6-HP will clarify whether the linker sequence in D6-HP corresponding to this C-terminal helix of Gelsolin forms a helix as well and thus may or not undergo a gelsolin-like, calcium-induced rearrangement.

The solution structure of D6 will be determined by NMR and analyzed in combination with the complete solution structure of HP and known structural properties of D6-HP. Together with the calcium and F-actin binding properties of D6 in isolation (currently under study), these data will clarify the role of the C-terminal domains of Villin in its activity as a physiologically principal actin regulator of microvilli.

Acknowledgements

I give much gratitude to my Master's thesis advisor, Dr. Serge Smirnov for sharing his wealth of knowledge, encouraging and guiding me throughout this structural investigation and being a great role model in the field of biochemistry.

I would like to thank Dr. Spencer Anthony-Cahill for his patience, being a great teacher, and generously providing me the guided access to his FPLC protein purification setup at a time-sensitive stage of my work. I would also like to thank Dr. P. Clint Spiegel for allowing me to use his lab space and materials, his availability and help with my project and reassuring me every step of the way. I greatly appreciate both of my thesis committee members for advising me throughout my time here at Western.

A special thanks to Dr. Fengli Zhang at the National High Magnetic Field Laboratory (Florida State University, Tallahassee, FL) for his assistance in utilizing their 720 MHz and 600 MHz NMR spectrometers. His expertise of these instruments has been crucial to the progress of my work. I would also like to thank Charles Wandler and Dr. Tom Pratum for their assistance with NMR experiments on the local 500 MHz NMR instrument.

Much appreciation is also given to Dan VanPelt in IT for expedient setup of my Linux workstation, which was crucial for the analysis of the NMR data. Also to past and present group members, especially, Lucian Burns who performed initial purification of D6 and Alex Nelson, my coworker on this project.

Table of Contents

List of Figures

List of Tables

Introduction

Microvilli of Epithelial Absorptive Cells in the Gastrointestinal Tract

In vertebrates, most absorption of nutrients occurs via specialized epithelial cells (enterocytes) of the small intestine. This organ is structured with vast circular folds which line the intestinal wall with finger-like projections. These projections are called villi and are comprised of epithelial cells that have many microscopic appendages exposed to the intestinal lumen (Campbell & Reece, Biology; 2002). These structured protrusions are called, microvilli, and are polarized at the apical membrane of the epithelial cells allowing a significant increase in the total surface area of the cell and are responsible for the level of absorption of nutrients (Campbell & Reece, Biology; 2002).

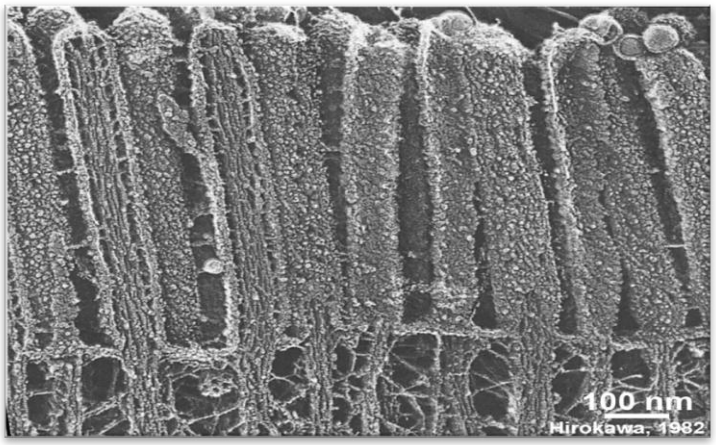

Figure 1: Infrastructure of Microvilli. Transmission Electron Micrograph (TEM) of microvilli of a nutrient-absorbing intestinal cell. The microvilli are sustained by bundles of actin filaments (F-actin), the principal protein of the cytoskeleton.
(Hirokawa et al. 1982. J. Cell Biol. 94, pp. 425-443, Fig. 1.)

The cytoskeleton of an enterocyte is comprised of a network of fibers called microfilaments and intermediate filaments. These components of the cytoskeleton provide structural support and also play a role in cell motility (Campbell & Reece, Biology; 2002). The ordered array of microvilli shown, (Figure 1) form a brush border and are sustained by parallel bundled actin filaments (F-actin), the major protein comprising microfilaments of the cytoskeleton. Filaments form a solid rod that is comprised of a twisted double chain helix built from molecules of globular actin (Campbell & Reece, Biology; 2002). F-actin is mechanically rigid and acts as a structural pillar to maintain the necessary cell shape. Multiple actin-binding proteins work together and are responsible for the formation and maintenance of the actin cytoskeleton in microvilli (Revenu et al., 2004). One of these proteins, Villin (Bretshcer & Weber, 1980), is the focus of our study. Together with the protein Fimbrin (Bretscher & Weber, 1980), Villin forms the actin bundling system of microvilli (Mooseker et al., 1980). Additionally, Villin regulates F-actin rearrangement under conditions of stress in a calcium sensitive manner (Glenney and Weber, 1981).

Gelsolin/Villin Family of Actin Regulating Proteins

There are multiple actin-binding proteins which exist; however, the Gelsolin/Villin family includes members that specifically play a role in actin filament remodeling (Silacci et al., 2004). Gelsolin is viewed as a common predecessor of these proteins (Yin & Stossel, 1979; Silacci et al., 2004), and gives rise to the modular, six-domain scaffold around which the family is organized. Gelsolin domains are 100-125

amino acids, share considerable sequence homology and much more obvious structural similarity (Kwiatkowski et al., 1986). Based on their sequence homology, the "core" scaffold domains [D1-D2-D3-D4-D5-D6] form a three-fold double repeat, also denoted as [D1-D2-D3]-[D1'-D2'-D3'] (Kwiatkowski et al., 1986; Bazari et al., 1988; Jannmey and Matsudaira, 1988). Gelsolin is capable of binding, severing, capping, and nucleating F-actin in a Ca^{2+}-dependent manner (Yin et al., 1981; Janmey et al., 1987).

Villin, which is the second member of this family, has seven modular domains (Figure 2), with the first six homologous to those of the Gelsolin core and containing at least one F-actin-binding site (Arpin et al., 1988; Bazari et al., 1988, Smirnov et al., 2007). The last domain, the C-terminal headpiece (HP), is unrelated in sequence to the other "core" domains and only homologous to other C-terminal headpiece domains (Arpin et al., 1988). The ~70 residue headpiece and the ~40 residue core-to-headpiece linker constitute the principal difference between Gelsolin and Villin. In addition to the functions of Gelsolin, Villin is also capable of F-actin bundling in a Ca^{2+}-dependent manner (Glenney and Weber., 1981; Matsudaira, 1982). Despite very high sequence similarity between Villin and Gelsolin, their respective activities are apparently not regulated via identical mechanisms (Kwiatkowski et al., 1986; Bazari et al., 1988; Janmey et al., 1988). The members of this superfamily contain at least three, and up to six homologous Gelsolin-like domains and perform at least one of these functions. Members include: Gelsolin (Yin & Stossel, 1979), Villin (Bretshcer & Weber, 1980), Severin (Andre et al., 1988), Adseverin (Bader et al., 1986), CapG (Prendergast et al.,

1991), Flightless I (Straub et al., 1996), Supervillin (Pestonjamasp et al., 1997) and Advillin (Marks et al., 1998).

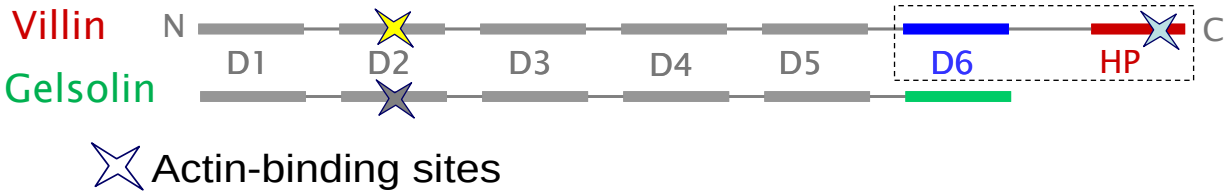

Figure 2: Modular Proteins Gelsolin and Villin. Domains [D1-D2-D3-D4-D5-D6] form a three-fold double repeat. These domains display ~50% sequence identity between Gelsolin and Villin and contain Ca^{2+}- dependent F-actin binding site. The headpiece [HP] domain is unique to Villin and houses a Ca^{2+}-independent F-actin binding site.

The Gelsolin/Villin superfamily includes other proteins where the structural and functional modularity is even more pronounced and intriguing. For example, another protein under study in our group is Supervillin (Pestonjamasp, et al., 1997). In Figure 3, it is shown that Supervillin has a C-terminal sequence that is homologous to Villin; however, these segments have been proven to not function in the same manner as Villin (Wulfkuhle, et al., 1999). Surprising results have actually designated the Supervillin N-terminus fragments to be responsible for binding and bundling F-actin. Although, completely unique in sequence, fragments A1, A2, and A3 bind F-actin in isolation and domain M binds non-muscle myosin II (Chen, et al., 2003).

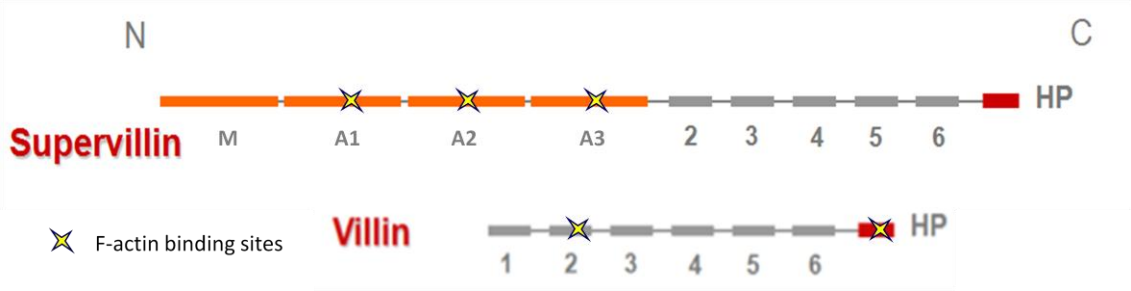

Figure 3: Supervillin: Modular protein in the Gelsolin/Villin Superfamily. F-actin binding and bundling activities are generally designated to the Gelsolin/Villin structural motif. However, these functions are provided by Supervillin's unique N-terminal domains A1, A2, and A3 instead.

Roles of Villin in the Microvillus

Villin is restricted and localized to the microvilli of absorptive epithelial cells (Ezzel et al., 1992) specifically expressed in the intestine and kidney proximal tubules (Matsudaira and Burgess, 1979; Bretscher and Weber, 1980; Robine et al., 1985) and functions as a Ca^{2+}-controlled, F-actin regulatory protein. *In vitro* studies have shown that in low calcium (10-100 nM) Villin is an F-actin bundling agent supporting the specialized brush border membrane of the absorptive epithelium. At intermediate micromolar calcium levels, Villin nucleates and caps the barbed ends of F-actin, and when intracellular calcium concentrations are high (Ca^{2+} >100 μM), Villin becomes an F-actin severing agent (Bretsher & Weber, 1980; Glenney et al., 1980, 1981; Mooseker et al. 1980). Villin's bundling function is fulfilled by the existence of an additional F-actin-binding site in HP (Glenney and Weber, 1981). Villin domains D1 and D2 are known to form an F-actin binding site and are given the responsibility of capping and

severing F-actin (Matsudaira et al., 1985; Janmey and Matsudaira, 1988) and the Villin domains associated with nucleation are thought to be D4-D6 (Friederich et al., 1999).

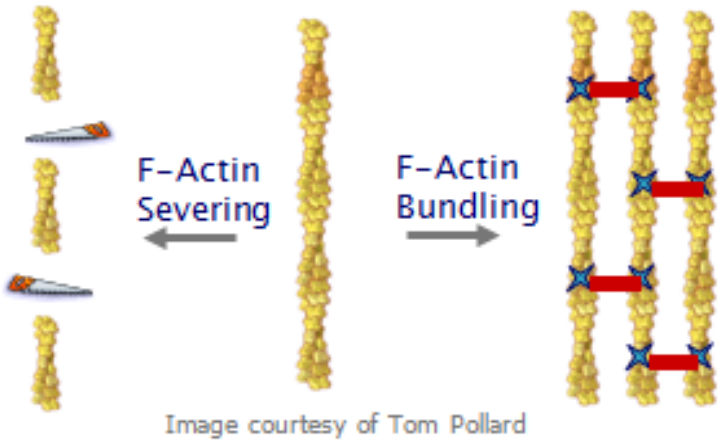

Figure 4: Two Major Functional Roles of Villin. Villin physically bundles F-actin in sub-μM concentrations of Ca^{2+} and severs F-actin into short filaments at higher (>100 μM) Ca^{2+} concentrations.

Due to Villin's ability to have multiple functionalities such as bundling and severing of actin filaments, it has been hypothesized to play a role in the induction of the formation of microvilli. Supporting evidence of this hypothesis were given when Villin was overexpressed in fibroblastic-like, non-polarized CV1 cells, which do not normally express this protein nor have microvilli. With Villin expressed, these cells were capable of the formation of an ordered brush border (Friederich et al., 1989). Villin suppression was also tested, in regards to Villin's direct role in structural support of microvilli, by the use of transcriptional silencing techniques in CACO-2 cells (Costa de Beauregard et al., 1995). In the TEM of the apical domains of the transfected

CACO-2 subclones, an absence of well-developed brush border was observed (data not shown). This experiment demonstrated that Villin is necessary for the formation of functional microvilli (Costa de Beauregard et al., 1995). The data from both of these experiments support a role for Villin in the formation and function of microvilli.

Additional experiments have shown that Villin knock-out mice (V-null), leads to no apparent ultrastructural changes in the microvilli of enterocytes (Ferrary et al., 1999). This assessment would suggest that the bundling function of Villin could be compensated for by other actin-binding proteins present in microvilli. However this study also showed interesting results for viable V-null mice epithelial cells when induced with lesions created by dextran sodium sulfate (DSS), which then gives rise to intracellular calcium (Okayasu, et al., 1990). For the reconstitution of these injured cells into functional enterocytes to occur, the intestinal epithelium undergoes an epithelial-mesenchymal transition (EMT) to reconstruct the epithelial layer (Boyer et al., 1996). This investigation of the V-null mice demonstrated that the reorganization of F-actin structural support in EMT was impaired in the V-null mice epithelial cells. This calcium-controlled functionality suggests that Villin may be involved in the repair process of the epithelial lining after lesion or infection (Ferrary et al., 1999).

Previous studies found progressive cholestasis and hepatic failure to be linked to a dysfunction in Villin gene expression (Philips, et al., 2003). In addition, reports have suggested a decrease in the levels of Villin expression in enterocytes from ulcerative colitis, Crohn's disease, and other inflammatory diseases, such as, chronic pancreatitis (Elsasser, et al., 1991; Kersting, et al., 2004). Others have shown evidence of an

increase in apoptosis and oxidative stress in the colon of patients with Crohn's disease and ulcerative colitis (Hagiwara, et al., 2002). Furthermore, patients with a range of gastrointestinal diseases were found to have higher amounts of Villin and Villin-specific autoantibodies in their serum (Rimm, et al., 1995). In response to these studies, a novel role for Villin was proposed recently. Wang and coworkers found Villin to be an epithelial cell-specific anti-apoptotic protein (Wang, et al., 2008). In this study, V-null mice were again induced with DSS, which demonstrated a direct correlation between the absence of Villin and the increase in apoptosis leading to severe ulcerative lesions. Together, these studies suggest a role for Villin in the regulation of cellular plasticity associated with epithelial cell injury.

Other experiments probing Villin function have added to its repertoire of activities in microvilli. It has been verified that Villin's severing activity enhances actin-based motility and morphogenesis in microvilli (Athman et al., 2003). These cellular processes of microvilli happen in response to extracellular signals during physiological (i.e. hormonal stimulation or fasting/feeding) or pathological situations (i.e. gastrointestinal diseases or bacterial infections). Villin plays an essential role in actin cytoskeleton dynamics by severing F-actin and contributing to actin cycling, which would increase the number of free barbed and pointed ends and allow active polymerization at the edge of the motile cells (Athman et al., 2003). Villin is also responsible for an increase in the velocity of *Shigella flexneri* and is implicated in the invasion of this harmful bacterium into humans via the gut (Heymann, et al., 2004; Athman et al., 2005). This hypothesis was also confirmed by the design of a severing

mutant of Villin. The mutant retained bundling, capping and nucleating activities, but showed a major reduction in severing activity. The mutant was tested *in vivo*, by monitoring the velocity of *S. flexneri* and in an *in vitro* motility assay that analyzed Villin's ability to regulate actin-based movement of beads. The results of both studies are consistent, by reporting a significant decrease in velocity of the bacteria and the beads in the presence of the Villin mutant (Revenu., 2007). Therefore, severing by Villin is an important functional role in the upkeep of microvilli.

Structural Comparison of Gelsolin and Villin Modular Domains

Multiple studies of Gelsolin and Villin's three-dimensional structures have been undertaken to shed new light on their functional properties. Generally, more is known about Gelsolin's three-dimensional structure than that of Villin's. Gelsolin structure has been rigorously studied by X-ray crystallography whereas Villin has not been crystallized. Only the solution structure of isolated Villin domain one (D1) (Markus et al., 1997) has been determined by NMR, and the 67-residue headpiece domain (HP) in isolation has been determined by NMR spectroscopy (Vardar et al., 1999) and X-ray crystallography (Meng et al., 2005). Additionally, a recent study of Villin fragment [D4-D5-D6] in a 1:1 mixture with Actin, reports a crystal structure of domain six (D6), the only domain able to crystallize in this three-domain fragment combined with Actin (Wang et al., 2009). The structure of D6 in Villin (*Homo sapiens*) resembles a regular Gelsolin-like fold and has no Ca^{2+} ion bound despite the calcium present in the mixture. Interestingly, when this D6 of Villin is compared to the corresponding domain 6 of

intact Gelsolin under the same conditions, there are conformational differences seen in the long helix of this domain (Wang et al., 2009). Villin D6 displays a straight helix but, the Gelsolin helix is kinked (Figure 5). Robinson and coworkers proposed a mechanism for the activation of Gelsolin (Robinson, et al., 1999). The kink in the long helix of calcium-free domain 6 is required to prevent clashes with the long helix of domain 4. Upon calcium-activation, the domain 6 helix is straightened to maintain a regular α-helical hydrogen bonding pattern (Robinson, et al., 1999).

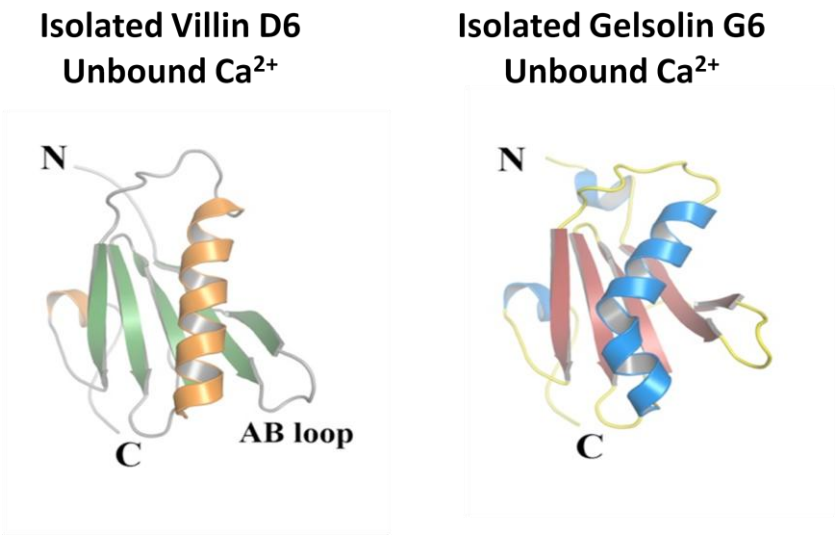

Figure 5: X-ray Crystal Structures of Domain 6 (D6) in Villin and Gelsolin. Structural comparison of D6, calcium free, in Villin vs. Gelsolin displays an evident conformational difference with respect to the long helix. This helix is straight in Villin and bent in Gelsolin (Wang et al., 2009).

The backbone dynamics of D6-HP, a modular fragment of Villin where domain six (D6) is linked to the headpiece (HP), has also been investigated by solution NMR

(Smirnov et al., 2007). This study indicates that: a) folded domains D6 and HP are connected with a flexible, unstructured, ~40-residue linker and b) the linker is the only tether between these two domains. The order parameters (S^2) based on the Lipari-Szabo model free analysis (Lipari and Szabo, 1980) designates two distinct structured moieties within D6-HP (residues 1-105 and 145-206) and a disordered region connecting them (residues 105-145) (Figure 6). In compliance with this data is the backbone ^{1}HN chemical shift profile of these three structural entities. This analysis indicates the same two structured regions having a wide dispersion of values, which also gives evidence of a folded polypeptide. Lastly, the linker sequence between the other two regions displays ^{1}HN chemical shift values in the narrow region of 8.0 and 8.5 ppm indicative of a random coil, unfolded polypeptide (Figure 6). The structural and dynamic characterization of D6-HP is the first study of a modular fragment within Villin that contains both a region of the Gelsolin-like core and the unique C-terminal headpiece domain (Smirnov et al., 2007).

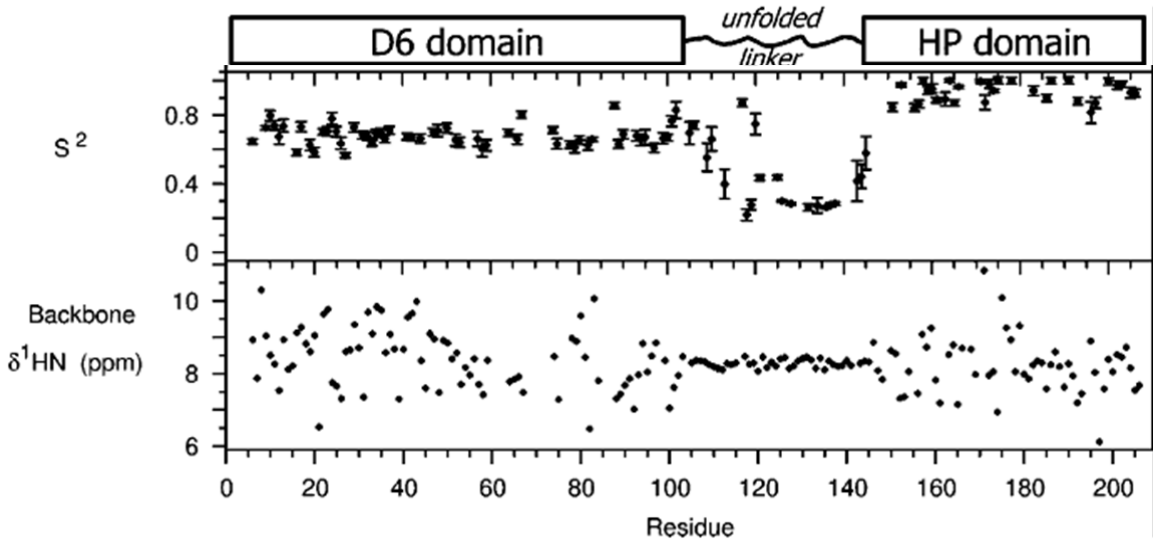

Figure 6: D6-HP: Two Domains (D6 & HP) and an Unfolded Linker. D6-HP: D6, HP and the linker. The order parameter (S^2) of Lipari-Szabo model-free analysis (R1, R2, ^{15}N-NOE) (Lipari and Szabo, 1980) and the ^{1}HN chemical shift profile (Smirnov et al., 2007).

Additionally, this NMR investigation revealed that Ca^{2+} is needed to maintain the structural integrity of D6-HP in the monomeric form thus suggesting a calcium-binding site within this fragment (Figure 7). From structural and functional investigations of the HP domain, it has been determined that this domain has no Ca^{2+}-binding sites (Vardar et al., 1999). Therefore, D6-HP calcium binding associates with D6, the linker, or their compilation. Furthermore, the crystal structure of D6 showed no Ca^{2+} ions bound, although present in the crystal, indicating that the Ca^{2+}-binding site in D6-HP may be associated with the linker region instead, an observation in agreement with previous reports (Kumar & Khurana, 2004). Surprisingly, this evaluation of D6-HP revealed that this modular fragment is capable of bundling F-actin in high, millimolar Ca^{2+}. With that, it is hypothesized that D6 or the linker has a cryptic F-actin binding site

(Smirnov et al., 2007). Our solution structure of isolated D6 will also help to further address these issues by clarifying whether D6 alone can bind calcium and/or bind F-actin.

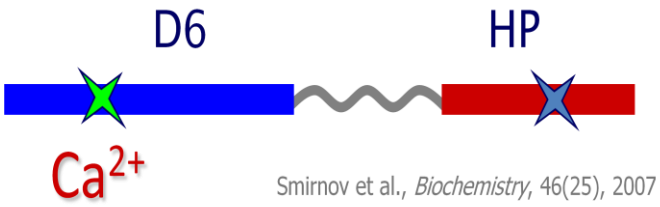

Figure 7: Villin's C-terminal Modular Fragment D6-HP. HP retains the same fold in D6-HP and in isolation. Villin and Gelsolin D6 have similar secondary structure. The inter-domain linker is mostly unfolded and ~40 residues long. D6 and HP only interact through the flexible linker. D6-HP bundles F-actin in high Ca^{2+} concentration. D6 or the linker has a cryptic F-actin binding site.

Supporting evidence has shown that the C-terminal half of Gelsolin is required for calcium regulation (Bryan et al., 1986). This hypothesis is also maintained by the observation of significant shape changes in the C-terminal half of Gelsolin, yet not seen in the N-terminal half when the protein is induced with calcium (Hellweg et al., 1993). A calcium binding site has been determined in Gelsolin's domain six, and pertaining to D6-HP of Villin's position numbering this site involves residues Asn29, Asp30, and Glu52 (Smirnov, et al., 2007). Due to sequence homology this could also be a calcium binding site on D6 of Villin. Residues Asp30 and Glu52 are conserved in Villin and Gelsolin and seen in the calcium-bound structure of Gelsolin, these residues contribute their side chains to interact with a calcium ion (Kolappan, et al., 2003). Residue Asn29 is not conserved between Villin and Gelsolin, however it binds calcium via its backbone

carbonyl oxygen. Another possible calcium binding site could be associated with residues located at positions 742 to 755 the C-terminal helix in domain 6 of Gelsolin, however this site, due to poor electron density, is unidentified in the electron-density maps and is not visible in the calcium-bound crystal structure (Robinson, et al., 1999; Kolappan, et al., 2003). In regard to the Villin D6-HP fragment, residues 742-755 in Gelsolin refer to residues 103-115 in the D6-linker inter-phase. In the absence of calcium, Gelsolin adopts a compact, inactive conformation stabilized by the 12-residue C-terminus helix. This helix was suggested to keep together Gelsolin domains D2 and D6 as a latch closed in low calcium and released at higher calcium levels (Burtnick et al., 1997; Lueck et al., 2000; Kolappan et al., 2003; Nag et al., 2009). However, this helix is not visible in the calcium-bound crystal structure (Kolappan, et al., 2003). We hope to investigate this C-terminal helix of Gelsolin in our structure of isolated D6 to probe the possible Gelsolin-type, calcium-induced structural remodeling and activation of Villin.

Nuclear Magnetic Resonance Spectroscopy (NMR)

NMR spectroscopy is a powerful technique used to elucidate the three-dimensional (3D) structures of biological molecules (up to 50 kDa) at atomic resolution (Teng, 2005). The application of NMR to structural biology is permitted by relatively recent developments within different fields, such as: Fourier transformation (FT) NMR, the development of stable magnets at high fields (up to 1 GHz), multidimensional,

heteronuclear NMR methods, as well as, overexpression and isotopic-labeling of proteins. Together, these scientifically engineered advancements have provided greater sensitivity and resolution in solution NMR, which is necessary in the investigation of large and complex biomolecular systems (Teng, 2005).

The first principal structure determination technique, X-ray Crystallography, emerged a quarter-of-a-century earlier than biological NMR and thus has been used to solve more structures. Nevertheless, NMR is unique in the sense that structural studies of biomolecules are performed in physiological relevant solution environments and can measure the molecular dynamics seen in biological systems. The NMR technique also allows the study of structurally dynamic protein domains or weakly bound protein complexes that are therefore, difficult to crystallize (Teng, 2005). This improvement over X-ray crystallography provides information about dynamics, flexibility, folding/unfolding transitions etc. for biopolymers, specifically for proteins.

Specific Aims

My thesis research utilizes the C-terminal modular domains D6 and D6-HP of *Gallus gallus* (chicken) Villin to further investigate the structural and functional relationship of this cytoskeleton regulating protein. The Specific Aims of my thesis work are as follows:

1) Expression of the C-terminal fragment D6-HP, as well as D6 in isolation, with high yield and purity, in the presence of calcium, with proper protein concentration for the structural characterization by solution NMR.

2) Complete assignment of NMR resonances (^{1}H, ^{13}C, ^{15}N) of D6 as the chemical shift reference for solution structure calculations.

3) The solution structure (2° and 3°) of D6 will be determined based on the NMR dihedral and distance restraints. From the structure of D6 we can further understand the actin-binding properties of D6 (if any), assist in determination of the fold of D6-HP, the positioning of a Ca^{2+}-binding site, and interactions within the Villin core domain D6, headpiece, and linker.

NMR Experimental Design

A 1D-NMR spectrum displays the NMR signal intensity versus the nuclei magnetic resonance frequency referred to as the chemical shift measured in units of parts-per-million (ppm). Each type of nuclei (e.g. ^{1}H, ^{13}C, ^{15}N) in proteins has a typical range of chemical shift values which may depend on the chemical environment in which it resonates. The task of NMR resonance assignment for isolated D6 requires NMR experiments beyond the 1D proton spectrum to assign each peak to a particular nucleus. For example, the 1D spectrum of D6-HP, a 208 amino acid polypeptide chain, is greatly complicated by signal overlap (Figure 8). When resonance peaks are not resolved, they cannot be assigned to their respective nuclei. However, when evaluating the resonance peaks of this 1D spectrum, we are able to determine whether or not the protein is folded. The broad region between 5-10 ppm refers to the chemical shift range for amide protons that belong to the backbone of the polypeptide chain. This spectrum displays peaks that are spread out within this range, which means many protons have a

unique location in the protein fold (Figure 8). If all of these proton peak intensities were shown in a narrow range (e.g. 7.5-8.5 ppm) and looked very similar, this would imply that all amide protons in D6-HP were exposed to the same chemical environment, meaning exposed to solvent (water), and the protein would be identified as unfolded.

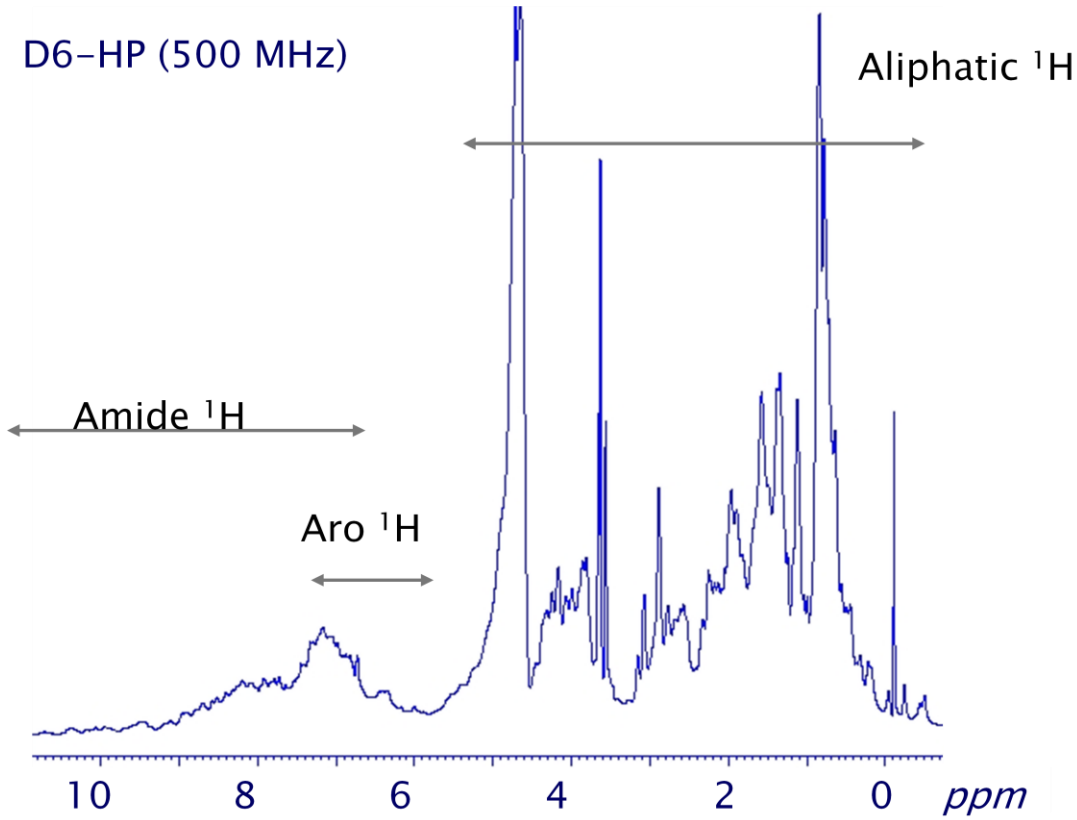

Figure 8: 1D proton NMR spectrum of D6-HP. Unresolved spectrum of D6-HP makes it difficult to assign proton resonances.

Certain types of NMR experiments can produce two-dimensional (2D) spectra where each signal or cross-peak reveals a physical interaction of two nuclei measured in the experiment. Introducing heteronuclear isotopes, such as, ^{13}C and ^{15}N, will allow "through bond" sequential assignments to be performed with the help of 2D and 3D

heteronuclear experiments. The additional heteronuclear dimensions (^{13}C, ^{15}N) will increase the spectral resolution by reducing the degeneracy of ^{1}H resonances (Teng, 2005). A principal type of a two-dimensional NMR experiment fundamental for protein NMR is called ^{15}N-HSQC (Heteronuclear Single-Quantum Coherence) where each cross-peak corresponds to the interaction of the proton and nitrogen atoms in an amide group (Palmer et al., 1991; Kay et al., 1992). The ^{15}N-HSQC spectrum represents a "fingerprint" of the protein being investigated. Each residue backbone amide group has a cross-peak shown on the spectrum, with the position of the peak being very sensitive to the local fold of the polypeptide chain. Inspection of the ^{15}N-HSQC spectrum for D6-HP (Figure 9), you can see better resolved peaks that correspond to the same 1D spectrum, which can be assigned. However, any proline residues in the primary sequence will not be manifested in this spectrum due to their lack of an amide group.

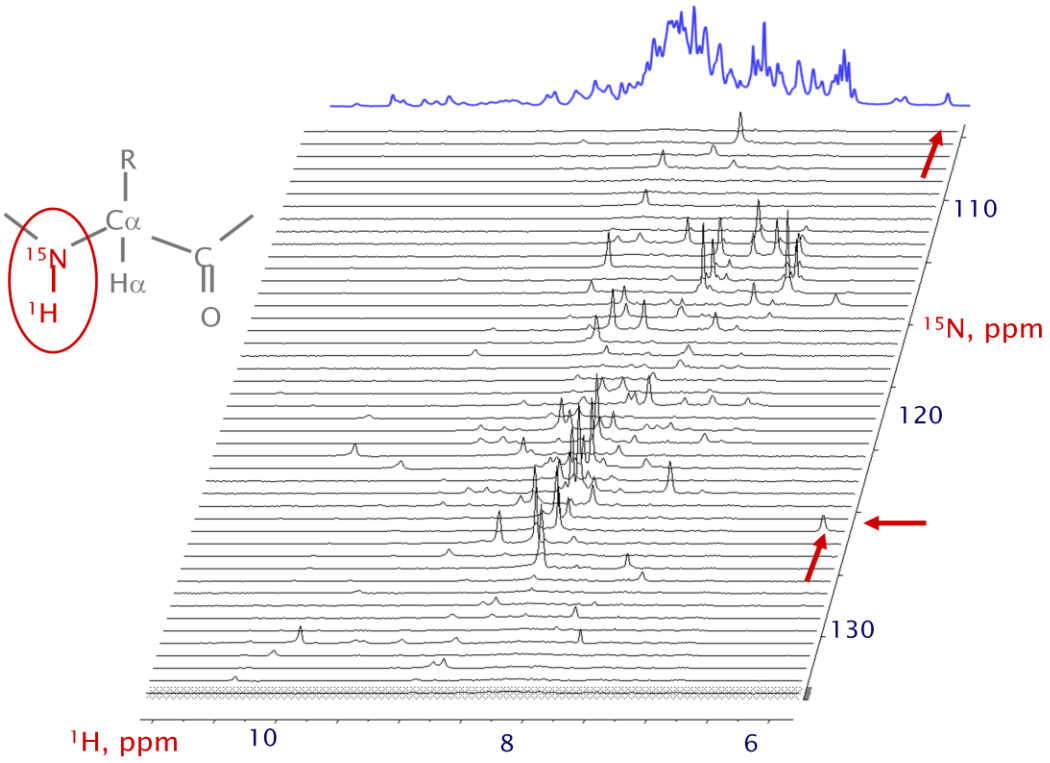

Figure 9: 2D ^{15}N-HSQC Spectrum of D6-HP. Fingerprint region; displaying chemical shifts for D6-HP backbone amide groups. One cross-peak gives rise to one amide group.

A combination of diverse complimentary heteronuclear NMR experiments is necessary to perform a sufficiently complete NMR resonance assignment, with the ^{15}N-HSQC being the common chemical shift reference for ^{1}H and ^{15}N of all these datasets. Multidimensional NMR experiments performed on ^{13}C/^{15}N-labeled proteins rely on the through-bond scalar-coupling interactions for resonance assignment and through-space dipole-dipole interactions for the solution structure of the protein. By introducing a ^{13}C and ^{15}N frequency dimension in combination with the ^{1}H dimension, the resonance assignments are significantly simplified.

The modern procedure for sequence-specific resonance assignment is based on the combined use of the heteronuclear "through bond" NMR data and the sequence of the protein. Each amino acid type reports its NMR resonances within a specific range of values (e.g. $^{13}C_{\beta}$ on a Thr resonates near 70 ppm, whereas, $^{13}C_{\beta}$ of Ala resonates around 20 ppm). Yet, the chemical shift ranges for the resonances on different residue types often overlap, thus it is necessary to incorporate supplementary information when mapping the resonance location onto the sequence of the polypeptide. This comes in the form of sequential "through bond" NMR cross peaks reporting interactions between nuclei from the adjacent residues, thus identifying sequential di-peptides unambiguously. For example, two specific experiments titled, CBCA(CO)NH (Grzesiek and Bax, 1992) and HNCACB (Wittekind et al., 1993; Muhandiram et al., 1994), are essential in linking a residue in the specific sequence to its neighboring residue. 3D NMR experiments are often named after the first half of the coherence transfer formed during the experiment (Teng, 2005). The parentheses designate the nucleus in which the chemical shifts do not evolve. The CBCA(CO)NH experiment, for example, is so-named because in it the β Carbon (C_{β}) atom is excited first, then the magnetization is transferred to the α Carbon (C_{α}) , through the Carbonyl Carbon (CO) and onto the amide Nitrogen and Hydrogen (NH) of the next residue. In Figure 10, the two spectra produced, have on display their resonance cross-peaks along an identical ^{15}N dimension. The spectrum to the left, CBCA(CO)NH, is showing the relationship of the current residue's amide group to the previous residue C_{α} and C_{β}. The experiment HNCACB on the right shows the relationship of the current residue's amide group to

their own C_α and C_β and the C_α and C_β of the previous residue. These two 3D experiments used in combination unambiguously relate the amide group of the residue with the C_α and C_β atoms of the previous one. The chemical shift values of the C_α and C_β atoms of the current and previous residues can indicate the possible chemical types of the adjacent amino acids which are often sufficient to find their exact position on the polypeptide sequence. A high resolution structure can be obtained when the assignment is complete for a sufficiently high number of atoms in the sequence (Teng, 2005).

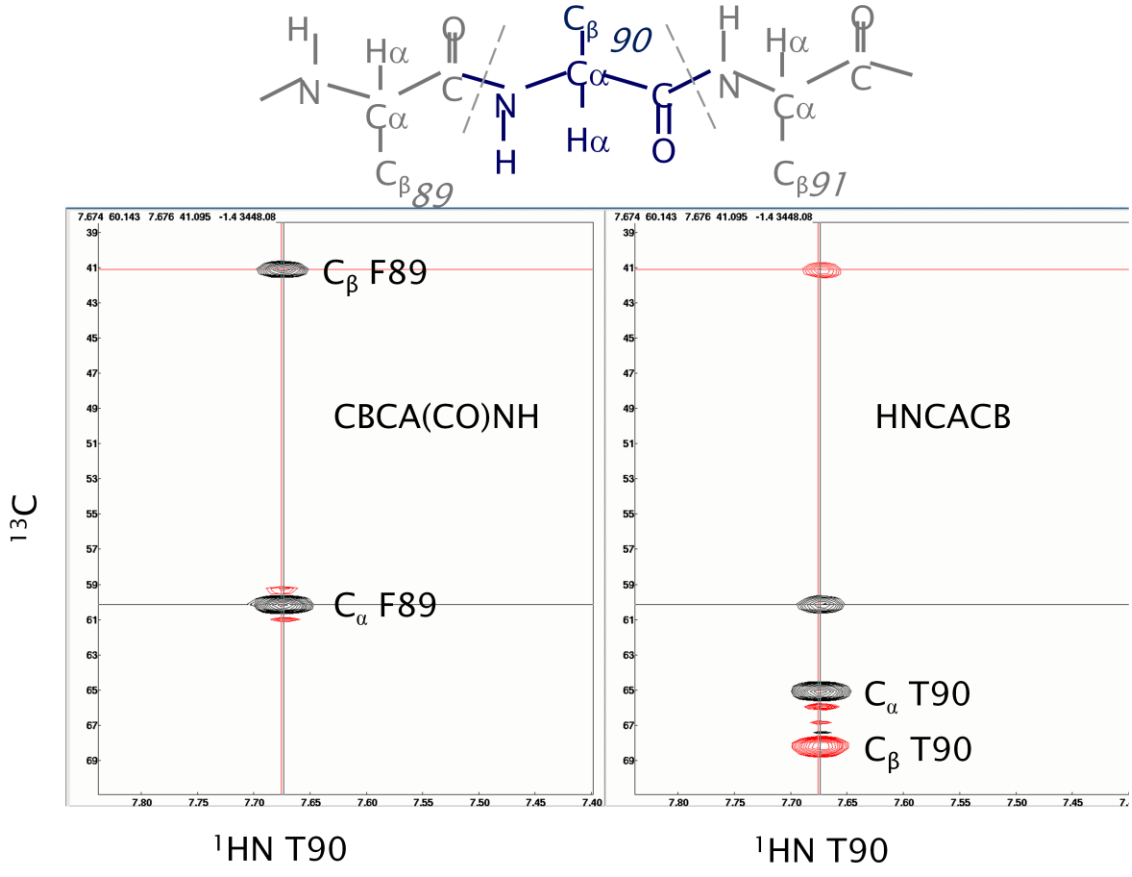

Figure 10: Triple Resonance Experiments CBCA(CO)NH & HNCACB of D6-HP. Two essential NMR experiments, which display the "walk" along the backbone of D6-HP (Smirnov et al., 2007).

One of the most important experiments to be acquired is the through-space Nuclear Overhauser Effect Spectroscopy (NOESY) data set. This 2D NOESY data set contains NOE cross-peaks, which are indicative of through-space proton-proton interactions between hydrogen atoms that are in short-range and long-range positions, but less than 5Å apart from each other in the protein structure (Teng, 2005). The intensity (I) of a NOE cross-peak is created by the cross-relaxation of these nuclei. The intensity of the cross-relaxation is inversely proportional to r^6 ($I = \frac{A}{r^6}$), with r being the distance between two nuclei (Teng, 2005). With the intensity being governed by this formula, cross-relaxation will fall rapidly with greater distance between two nuclei, which are why NOE cross-peaks only arise within the 5Å limit. Thus, this experiment provides the information that defines the unique protein fold, which is the most essential for the elucidation of a NMR-based solution structure.

With the sequence specific NMR resonance assignment nearly complete, the chemical shifts of the NMR-active nuclei will be used to predict the secondary structure with a program named TALOS (**t**orsion **a**ngle **l**ikelihood **o**btained from the **s**hift and sequence similarity) (Cornilescu et al., 1999). This algorithm utilizes the chemical shift values of the nuclei (^{15}N, $^{13}C_\alpha$, ^{13}CO, $^{13}C_\beta$, $^{1}H_\alpha$) pertaining to the current, previous, and the following residue to predict secondary structure (φ and ψ angles) of the current (central) residue by analogy to other previously determined protein structures. The predicted torsion angle ranges for the residues are then used as backbone dihedral restraints in the 3D structure calculation along with the NOE-based distance restraints.

CYANA (**C**ombined assignment and d**y**namics **a**lgorithm for **N**MR **a**pplications) (Guntert, et al., 2003; 2004) is the software that performs the automated determination of the protein solution structure. CYANA assigns the NOE peaks and determines the solution structure of the protein in an integrated, iterative process utilizing the combination of protein NMR resonance assignment, dihedral angles, and raw NOE peak data.

Results and Discussion

Protein NMR Sample Preparation

D6-HP and D6 are readily expressed in *E.coli*, BL21(DE3) strain, with a yield of purified $^{13}C/^{15}N$ labeled proteins of 40 mg/L of labeled medium (Marley et al., 2001; Smirnov et al., 2007). The optimal expression conditions that produced sufficient yields for NMR were with an induction time of 3-5 hours, in shake flasks at 37°C. Both protein fragments are soluble, (Figure 11) to at least 1 mM, and stable for several weeks at room temperature in the presence of DTT (10 mM) and Ca^{2+} (5 mM). The addition of Ca^{2+} is essential to prevent slow, irreversible aggregation of D6-HP which suggests a key interaction with Ca^{2+} (Smirnov et al., 2007). The necessity of calcium ions for the fold and structural stability of D6 in isolation is being currently investigated. The detailed expression and purification processes for these proteins are described in "Materials and Methods".

Isotopically labeling ($^{13}C/^{15}N$) recombinant proteins is an essential technique developed to study large biomolecules by NMR. The current production protocol for $^{13}C/^{15}N$ labeled protein, utilizes the "M9 minimal medium" protocol, which achieves more than 90% incorporation of the desired isotopic labels (Marley et al., 2001). The 90% incorporation of isotopic labels would be otherwise unattainable in rich medium because a competition between labeled $^{13}C/^{15}N$ sources and the natural occurring

nitrogen and carbon sources would occur in the Luria-Bertani broth and a low percentage of isotopically labeled protein would be the result.

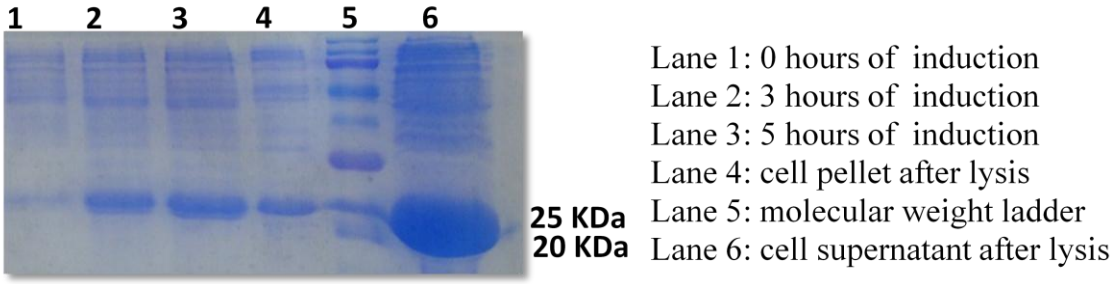

Figure 11: SDS-PAGE results for induction and lysis of D6-HP. 3-5 hours of induction was optimal for D6-HP and D6 (data not shown). Most of D6-HP and D6 remained in the soluble fraction after lysis.

In order to reach the sample purity optimal for NMR, two steps of purification were necessary. The first step used a gravity-flow size exclusion chromatography (SEC) column as the initial step to fractionate all the soluble proteins found in the cell supernatant after lysis. Separation by apparent hydrodynamic radius is achieved as the large proteins elute first and the smaller proteins next. Optimal separation was a challenge; however, SEC clearly enriched our target proteins in the later eluting fractions (Figure 12). Confirmed by SDS-PAGE, the samples collected displayed a distinct increase in concentration of the desired protein versus the other contaminants in the fractions which were known to contain our target protein (Figure 12). Protein Spot Tests (Figure 13) were another confirmation method used to identify fractions with a noticeable presence of protein. We would use these results to then perform a more in depth SDS-PAGE analysis of only those fractions containing purified D6 or D6-HP.

The fractions containing the majority of our target protein (e.g., fractions 50-60 of D6; see Figure 13) were combined and used for the final purification step. SDS-PAGE expression gel shows large protein contaminants present with our target protein (D6/D6-HP) in the supernatant after lysis. As proven by SDS-PAGE (Figure 11), fractions (25-35) are mostly contaminants and fractions (50-60) contain most of D6 (Results for D6-HP are not shown).

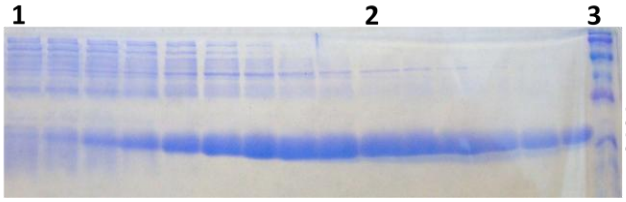

Figure 12: Gravity-Flow SEC Analysis by SDS-PAGE for D6-HP. Displayed by SDS-PAGE are protein containing fractions collected after Gravity Flow SEC. Lane 1 is an example of a fraction containing other contaminant proteins. Lane 2 is an example of a fraction containing the majority of our target protein, D6-HP (23.5 kDa), but still impure. Lane 3 is the molecular weight reference (D6 SDS-PAGE results not shown).

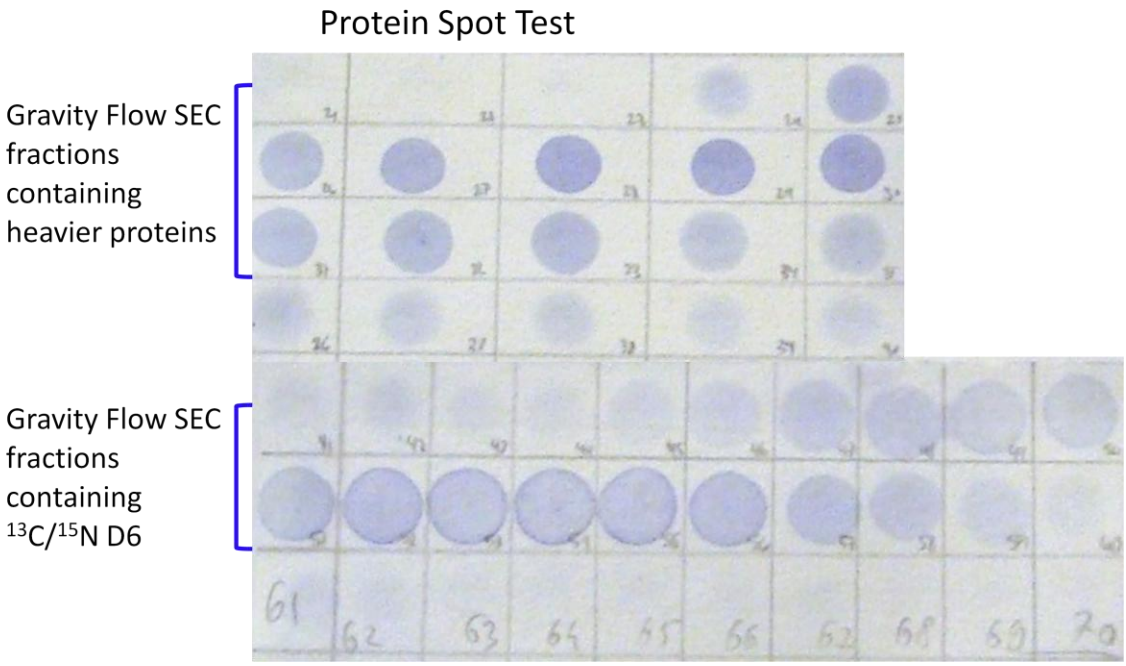

Figure 13: Protein Spot Test to Identify D6 Fractions after Gravity Flow SEC. Lower numbered fractions (25-35) were the first protein containing fractions eluted from the column. Higher numbered fractions (50-60) were the second protein containing fractions to be eluted from the column.

The combined fractions containing most of D6 or D6-HP were purified by High performance size exclusion chromatography (HP-SEC). This resulted in the final purified product necessary for NMR. Figure 14, displays the HP-SEC chromatogram obtained by monitoring the column effluent at 280 nm. The SDS-PAGE gel demonstrates the purity of the desired protein (D6 and D6-HP). UV-Vis spectroscopy was also used to quantify the concentration of purified D6 and D6-HP NMR samples, as described in Materials and Methods.

Figure 14: Pure $^{13}C/^{15}N$-D6 NMR sample. HP-SEC results of D6 (D6-HP data not shown). Peak 1 shows the heavier protein collected sample labeled "B3". Peak 2 shows the collected fraction "C3" containing pure D6. Lane 3 verifies the purity of C3 fraction containing D6 by SDS-PAGE. Lane 4 is the molecular weight reference.

^{15}N-^{1}H-HSQC spectral Analysis

Figure 15 shows the ^{15}N Heteronuclear Single Quantum Coherence (^{15}N-HSQC) spectrum for $^{13}C/^{15}N$-D6 protein sample at ~1 mM recorded at 25°C and pH 7.0 in 5 mM Ca^{2+}. The dispersion pattern of the ^{15}N-HSQC cross-peaks are very sensitive to the fold of the protein and are often compared to the fold of related protein samples, or of the same sample, at different conditions (i.e. physiological, chemical, etc.). In Figure 15, the majority of ^{15}N-HSQC cross-peaks displayed for D6-HP (red) vs. isolated D6 and HP (Vardar et al., 1999) (black) spectra overlay correspond to those of isolated D6 and HP. This suggests that both domains in D6-HP are folded and retain the same fold as in isolation. Any cross-peaks observed with a difference in amide backbone environment most likely occur in residues near the linker. Currently, ~95% of the

backbone and side chain resonances in D6-HP are assigned (Smirnov et al., 2007, BMRB). The D6 assignment will be greatly facilitated by the strong similarity of the ^{15}N-HSQC D6 spectra in the context of D6-HP and in isolation. The corresponding ^{15}N-HSQC data for the HP domain was obtained from C.J. McKnight (Boston University School of Medicine, Boston, MA).

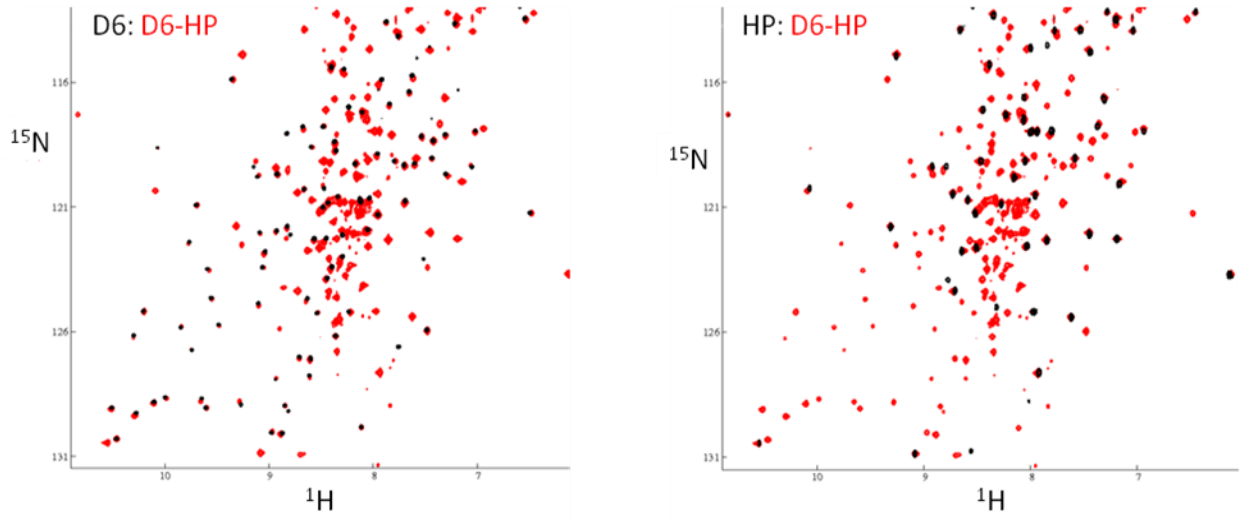

Figure 15: ^{15}N-HSQC Spectra Overlay. These spectra show the amide backbone chemical shifts of D6 (black) vs. D6-HP (red) and HP (black) vs. D6-HP (red). The D6 spectrum was collected on a Varian 720 MHz at 25°C in H_2O (see Materials & Methods). Spectra were processed using NMRPipe (Delaglio, et al. 1995) and visualized with NMRView (B.A. Johnson, 1994).

The chemical shift differences seen in Figure 16, for isolated D6 and isolated HP with D6-HP, were calculated based on the amide backbone resonance assignment. Most of the ^{1}HN chemical shift differences are less than 0.04 ppm, which indicates very similar structural environments and are therefore negligible. The potential effect of the

inter-domain linker on the domain structures is exemplified by the observation of two proton chemical shifts of F84 and D166, which differ from the corresponding values in isolated D6 and HP. The quantified chemical shift differences are 0.63 ppm and 0.18 ppm, respectively (Figure 16). Noticeable chemical shift differences such as these are likely to occur in residues which reside at the C-terminus of isolated D6 and the N-terminus of isolated HP, since they will have a change in chemical environment when the linker is not present. However, these two positions are well away (23 residues) from the termini of D6 and HP, and are located on the surface of the isolated domains (Vardar et al., 1999; Wang et al., 2009). These noticeable chemical shift differences indicate potential alterations of the local conformation of domain 6 and the headpiece possibly due to the presence of the linker residues in the vicinity of F84 and D166. In the previous study of D6-HP, the ^{15}N-HSQC was taken of the fragment in the presence and absence of calcium where the D166 amide resonance remained essentially unaltered (Smirnov et al., 2007). This displayed evidence that again no residues in the headpiece domain bind calcium and that this extreme chemical shift difference is most likely caused by the presence of the linker in D6-HP. Current efforts are underway to investigate the effects of the absence of calcium on D6 amide backbone resonances (A. Nelson and S. Smirnov).

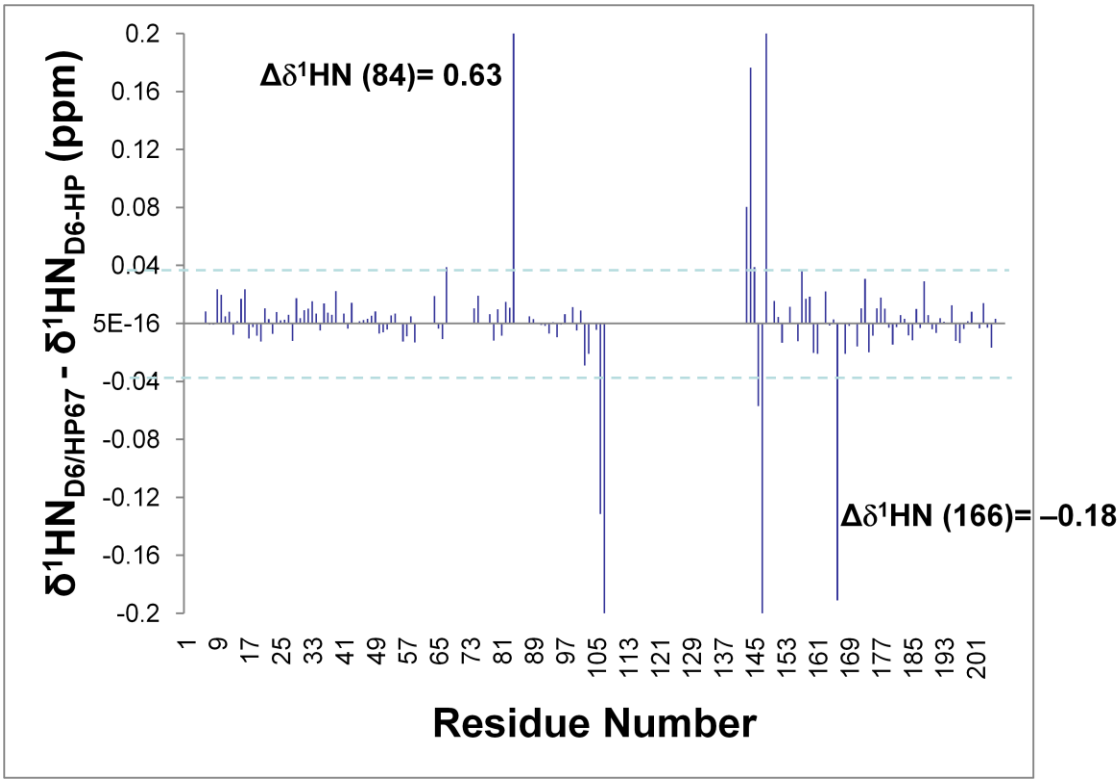

Figure 16: Backbone δ ^{1}H Differences: D6 and HP67 vs. D6-HP. Qualitative analysis of the backbone ^{1}HN chemical shift differences of isolated D6 and isolated HP with D6-HP samples at 25°C and pH 7.0.

D6 Resonance Assignments by Multidimensional Heteronuclear NMR

Sequence-specific NMR resonance assignments of isolated D6 were generated by Nuclear Magnetic Resonance Spectroscopy in solution. To obtain a high resolution structure, it is necessary to complete the assignment for a sufficient number of atoms in the sequence. The sequential resonance assignment of the amide backbone for D6 in isolation is shown in Figure 17 and listed in Table 1. Only 101 of the 107 residues in D6 are assigned here because six residues in the sequence are Proline (1, 68,

76, 86, 87, and 98). Any resonance assignments for these residues come from additional heteronuclear triple resonance experiments such as HBHA(CO)NH, CBCA(CO)NH, HNCO, HCC(CO)NH, and CC(CO)NH. Other unassigned peaks that appear in this spectrum belong to other amide groups that reside on the aromatic heterocyclic side chain of tryptophan (5 Trp) or side chains of histidine (1 His) and arginine (5 Arg). Also the amino groups found on the side chain of asparagine (3 Asn) and glutamine (4 Gln) resonates in this spectrum. The majority of the amide backbone cross-peaks resonate at a similar intensity, however, the 15N-HSQC cross-peak of residue Gly69 (124.2, 8.17), which resides on a turn loop next to Pro68 and Ser70 is very weak. This observation can be interpreted as evidence for intermediate structural exchange on the NMR time scale, in the region of Gly69, and therefore results in a weak cross-peak. We have utilized the heteronuclear 3D NMR data recorded with the $^{13}C/^{15}N$-labeled D6 sample to unambiguously assign the backbone resonances (^{1}HN, ^{15}N, $^{13}C_{\alpha}$, and ^{13}CO) for 106 residues of the 107 in D6 using established procedures (Cavanagh et al; 1996).

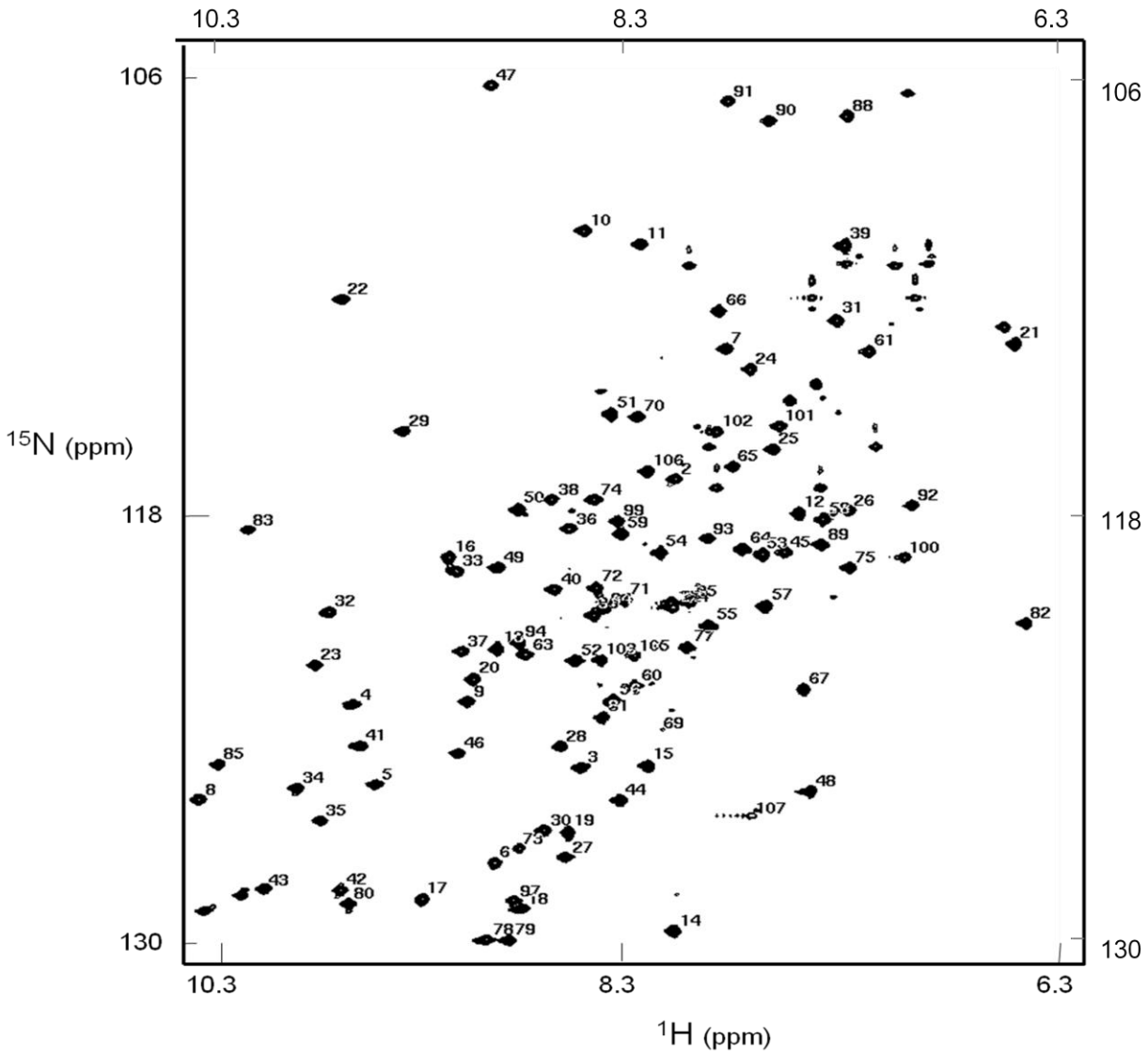

Figure 17: Isolated D6 Backbone Resonance Assignments. ^{15}N-HSQC "fingerprint" region for isolated D6 with each backbone amide cross-peak labeled with its corresponding residue number.

Table 1: ^{15}N & ^{1}H Resonance Assignment for Backbone Amide groups in D6. Chemical shifts are designated from the ^{15}N-HSQC spectrum for isolated D6. The peak assignment was determined by PINE online server (Bahrami, et al., 2009).

Res. #	Res. Name	^{15}N-^{1}H	δ (ppm)	^{1}H-^{15}N	δ (ppm)	Res. #	Res. Name	^{15}N-^{1}H	δ (ppm)	^{1}H-^{15}N	δ (ppm)
2	ARG	N	117.203	H	8.108	**52**	GLU	N	122.271	H	8.566
3	LEU	N	125.261	H	8.542	**53**	LYS	N	119.319	H	7.704
4	PHE	N	123.49	H	9.591	**54**	GLU	N	119.268	H	8.173
5	GLU	N	125.718	H	9.484	**55**	ALA	N	121.311	H	7.954
6	CYS	N	127.915	H	8.937	**56**	ALA	N	123.395	H	8.393
7	SER	N	113.574	H	7.873	**57**	ALA	N	120.754	H	7.69
8	ASN	N	126.153	H	10.299	**58**	GLU	N	118.337	H	7.423
9	LYS	N	123.42	H	9.064	**59**	THR	N	118.744	H	8.354
10	THR	N	110.292	H	8.525	**60**	ALA	N	122.991	H	8.294
11	GLY	N	110.668	H	8.266	**61**	GLN	N	113.649	H	7.215
12	ARG	N	118.17	H	7.538	**62**	GLU	N	120.684	H	8.126
13	PHE	N	121.956	H	8.926	**63**	TYR	N	122.118	H	8.796
14	LEU	N	129.824	H	8.114	**64**	LEU	N	119.163	H	7.799
15	ALA	N	125.201	H	8.234	**65**	ARG	N	116.868	H	7.842
16	THR	N	119.393	H	9.149	**66**	SER	N	112.527	H	7.907
17	GLU	N	128.928	H	9.269	**67**	HIS	N	123.09	H	7.515
18	ILE	N	129.185	H	8.814	**69**	GLY	N	124.202	H	8.166
19	VAL	N	127.068	H	8.6	**70**	SER	N	115.467	H	8.28
20	ASP	N	122.793	H	9.037	**71**	ARG	N	120.585	H	8.341
21	PHE	N	113.443	H	6.542	**72**	ASP	N	120.258	H	8.471
22	THR	N	112.2	H	9.64	**73**	LEU	N	127.518	H	8.826
23	GLN	N	122.406	H	9.766	**74**	ASP	N	117.787	H	8.48
24	ASP	N	114.153	H	7.76	**75**	THR	N	119.682	H	7.303
25	ASP	N	116.379	H	7.655	**77**	ILE	N	121.906	H	8.052
26	LEU	N	118.095	H	7.309	**78**	ILE	N	130.054	H	8.977
27	ASP	N	127.757	H	8.612	**79**	VAL	N	130.072	H	8.872
28	GLU	N	124.674	H	8.632	**80**	VAL	N	129.055	H	9.608
29	ASN	N	115.882	H	9.36	**81**	LYS	N	123.854	H	8.441
30	ASP	N	127.013	H	8.71	**82**	GLN	N	121.237	H	6.493
31	VAL	N	112.784	H	7.364	**83**	GLY	N	118.629	H	10.072
32	TYR	N	120.917	H	9.698	**84**	PHE	N	120.834	H	8.431
33	LEU	N	119.79	H	9.111	**85**	GLU	N	125.174	H	10.207
34	LEU	N	125.829	H	9.847	**88**	THR	N	107.088	H	7.316
35	ASP	N	126.739	H	9.74	**89**	PHE	N	119.031	H	7.437
36	THR	N	118.603	H	8.596	**90**	THR	N	107.216	H	7.675
37	TRP	N	122.021	H	9.09	**91**	GLY	N	106.674	H	7.865
38	ASP	N	117.786	H	8.676	**92**	TRP	N	117.951	H	7.018
39	GLN	N	110.693	H	7.325	**93**	PHE	N	118.863	H	7.956
40	ILE	N	120.287	H	8.663	**94**	MET	N	121.781	H	8.825
41	PHE	N	124.653	H	9.555	**95**	ALA	N	120.648	H	8.037
42	PHE	N	128.681	H	9.647	**96**	TRP	N	121.009	H	8.481
43	TRP	N	128.634	H	9.997	**97**	ASP	N	128.963	H	8.85
44	ILE	N	126.166	H	8.359	**99**	LEU	N	118.381	H	8.372
45	GLY	N	119.26	H	7.603	**100**	CYS	N	119.396	H	7.052
46	LYS	N	124.868	H	9.104	**101**	TRP	N	115.737	H	7.625
47	GLY	N	106.23	H	8.954	**102**	SER	N	115.896	H	7.916
48	ALA	N	125.925	H	7.484	**103**	ASP	N	122.262	H	8.45
49	ASN	N	119.683	H	8.921	**104**	ARG	N	120.795	H	8.117
50	GLU	N	118.057	H	8.828	**105**	LYS	N	122.131	H	8.294
51	SER	N	115.397	H	8.401	**106**	SER	N	116.993	H	8.235
						107	TYR	N	126.603	H	7.755

Resonance assignments are also supported by NOESY and TOCSY spectra acquired in H_2O and D_2O. The NOESY spectrum displays typical sequential NOE connectivity as well as long range NOE interactions which were used in the 3D structure calculations by CYANA software (Guntert, et al., 2003; 2004). Side chain resonance assignments are also of great importance for the elucidation of a high resolution solution NMR structure. CYANA software requires a minimum of 50% resonances assignments of a protein for structure calculations. Currently we have a total of 91% resonances assigned for D6, which were determined by heteronuclear 3D NMR data.

Secondary Structure Analysis of D6

The isolated D6 secondary structure corresponds to that of Gelsolin domain 6, as well as, that of the D6 domain in the context of D6-HP (Figure 18). The secondary structure elements such as, α-helices and β-sheets, in D6 were determined by TALOS software (Cornilescu et al., 1999). The protein nuclei $^{1}H_{\alpha}$, ^{15}NH, $^{13}C_{\alpha}$, ^{13}CO, and $^{13}C_{\beta}$ are adjacent to the φ and ψ backbone angles. Therefore the chemical shift values of these nuclei are sensitive to the secondary structure elements and are utilized by TALOS to predict the secondary structure pattern. The D6 secondary structure resembles that of domain 6 of Gelsolin (Burtnick, et al., 1997; Kolappan, et al., 2003). The only major structural difference between D6 and D6-HP with domain 6 of Gelsolin is the C-terminal helix (residues 103-113 in D6-HP). Due to the fact that D6 is truncated

after 107 residues, this structural entity is challenging to study in the isolated-form of D6. Together with the high degree of sequence identity and secondary structure similarity, we expect that the D6 folds as a Gelsolin-like "core" domain, that is five-stranded β-sheet placed in between a long α-helix and another short α-helical element of the domain (Smirnov, et al., 2007).

Figure 18: Secondary Structure: D6-HP and D6 vs. Gelsolin Domain 6 and HP67. Comparison of the secondary structure of D6 (blue), D6-HP (black), Gelsolin (Burtnick et al., 1997; Kolappan et al., 2003) (green), and HP (Vardar et al., 1999; Meng et al., 2005) (red). The D6 & D6-HP sequence is 50% identical to the Gelsolin domain 6 (dashed underline). The linker sequence (no underline) and Villin headpiece (dotted underline) are also designated. The elements of secondary structure (arrows represent β-sheets and cylinders helices) for D6 and D6-HP were predicted by TALOS (Cornilescu, et al., 1999). Secondary structure elements found in the calcium-free Gelsolin domain 6 and the isolated Villin headpiece, HP, are labeled Gelsolin and HP, respectively. Only one residue (proline 86) in D6 is completely unassigned and is denoted by a lowercase letter.

Three-Dimensional Fold of D6 in Isolation

The general 3D fold of isolated D6 polypeptide has been determined. Currently, we are in the process of refining the solution structure of D6.

In the beginning of the calculation process our α-helical features in the 3D structure were continuously present. According to our secondary structure calculation based on dihedral angles found by TALOS (Cornilescu et al., 1999), our structure should have at least 3 α-helices and 5 β-strands. At most we were able to obtain 3 of our 5 β-strands in the structure calculations. According to Cornilescu and coworkers, CYANA software had great difficulty with structure convergence of their beta-rich protein and needed the assistance of manual NOE assignments to generate a structure (Cornilescu, et al., 2006). The helical elements possess high structural stiffness in all three dimensions due to their geometry and therefore the intra-helical proton contacts are easier for CYANA to decipher. The flexible β-strand elements are characterized by longer-range hydrogen-bonding interactions than helices, and in our experience CYANA software is limited in deciphering this type of NOE data.

The β-strands of D6 were a challenge to elucidate in our 3D structures due to their intrinsic lack of the 3D bulk and rigidity. To guide CYANA software to the proper fold of the β-strands into a sheet, we manually assigned some of the unambiguous inter-strand NOE cross-peaks. The most efficient way of doing this was to look for NOE cross-peaks between different amide protons of residues known to reside in a β-strand (based on the secondary structure determined for D6; Figure 18). The 2D ^{1}H-NOESY was used to locate any amide proton cross-peaks, for example, in this spectrum a NOE

cross-peak was identified with the chemical shift values of 8.11 ppm and 9.74 ppm. The 9.74 ppm value correlates only with Asp35; however the 8.11 ppm value correlates with Arg2, Leu14, and Arg104. Based on sequence homology and secondary structure Asp35 is most likely interacting with Arg2, but the ^{15}N-edited NOESY spectrum was utilized to verify this cross-peak. With the use of each residue's amide ^{15}N chemical shift, we were able to sift through different ^{15}N dimensions in the ^{15}N-edited NOESY. Seen in Figure 19, the spectrum on the left shows the amide proton cross-peak along the ^{15}N plane of residue Asp35 (126.739 ppm) and on the right shows the same amide proton cross-peak along the ^{15}N plane of residue Arg2 (117.203 ppm). The other candidate residues, Leu14 and Arg104, demonstrated no ^{15}N-NOESY cross-peaks with Asp35. Therefore, these inter-strand residues were validated to be within 5Å of each other and we were able to manually assign this cross-peak to the CYANA cross-peak file. This procedure was performed for 17 more inter-strand residues.

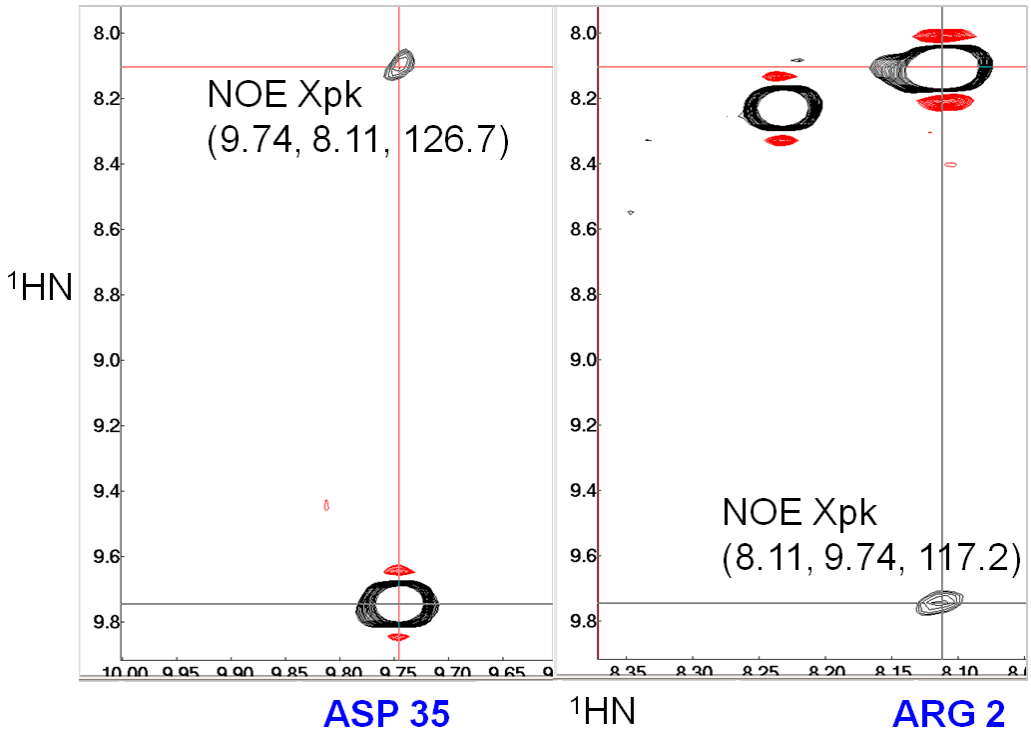

Figure 19: ^{15}N-edited NOESY spectra of amide protons from inter-strand residues Arg2 and Asp35. NOE cross-peaks in both ^{15}N dimensions for residues Arg2 and Asp35 verify a backbone amide proton interaction between these two residues, which reside on two different β-strands.

The addition of manual assignments between multiple inter-strand residues resulted in a solution structure displaying the isolated D6 three-dimensional fold (Figure 20). Table 2 summarizes the structural features of our current NMR structure of D6 (*Gallus gallus*) and the X-ray Crystal structure of D6 (*Homo sapiens*) discussed in this study (Wang et al., 2009). In terms of the structures fold, both display a core sheet of 5 β-strands sandwiched between a long α-helix and a shorter C-terminal α-helical element. The NMR structure of our isolated D6, displays a 3D backbone fold that resembles a typical Gelsolin-like fold.

Table 2: Comparative Summary of Structural Features from the Current NMR and Crystal Structures of D6.

	# of H-Bonds	# of Helices	# of Strands	# of Turns
Current NMR Structure of D6	44	2	5	11
X-Ray Crystal Structure of D6	71	3	5	10

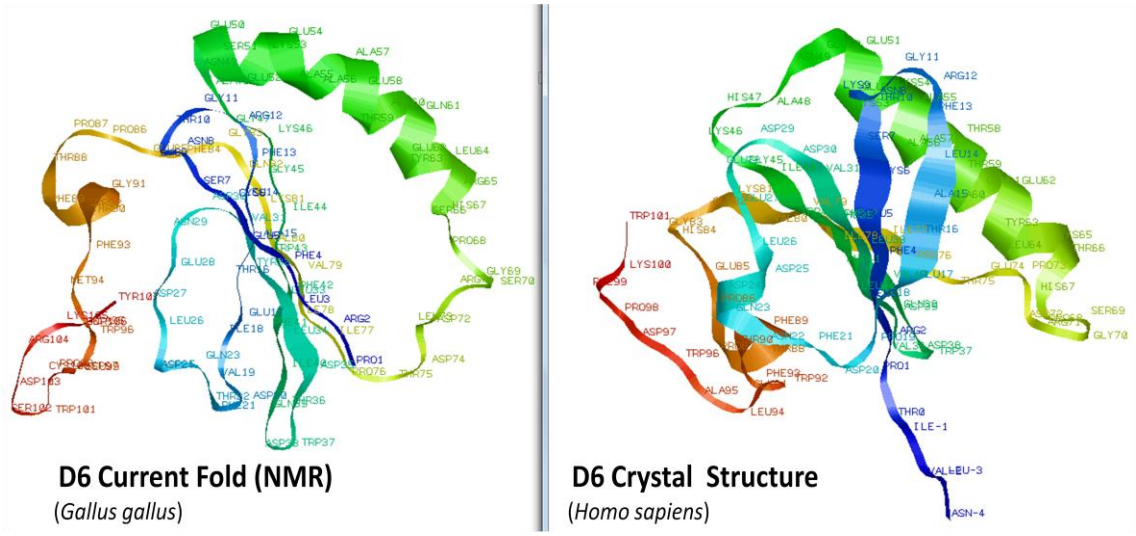

Figure 20: Isolated D6 3D structures. The RasMol (Sayle and Milner-White, 1995) molecular depictions show the current NMR structure of D6 from chicken Villin (left) which displays its 3D-fold and is in the final refinement steps for a final solution structure. The crystal structure of D6 from human Villin (right) shows its final calcium-free conformation (Wang et al., 2009). Both folds depict 5 β-strands spread into a core sheet sandwiched between a long α-helix and a shorter α-helix.

The current NMR structure simulation of isolated D6 was initiated with 500 starting structures, which were then minimized to conformations based on the experimentally derived dihedral angles and NOE distance restraints. The top 10 structures reported gave the lowest average RMSD and NOE violations out of all the converging structures. Currently 65% of the ^{15}N-edited NOESY cross-peaks were

assigned by CYANA and gave no apparent violations. This results give us confidence in our current backbone assignment, however doesn't mean that violations don't exist. CYANA leaves an NOE cross-peak unassigned if it cannot find a pair of protons with matching chemical shifts. Some ambiguous cross-peaks may also be unassigned by the program if that allows for a lower value of the target function (i.e., conformational energy). Figure 21 is a D6 structure overlay of the top 10 structures. The current average backbone RMSD is 3.0 Å, which is not sufficient for a final structure, but allows prediction of the 3D fold for D6.

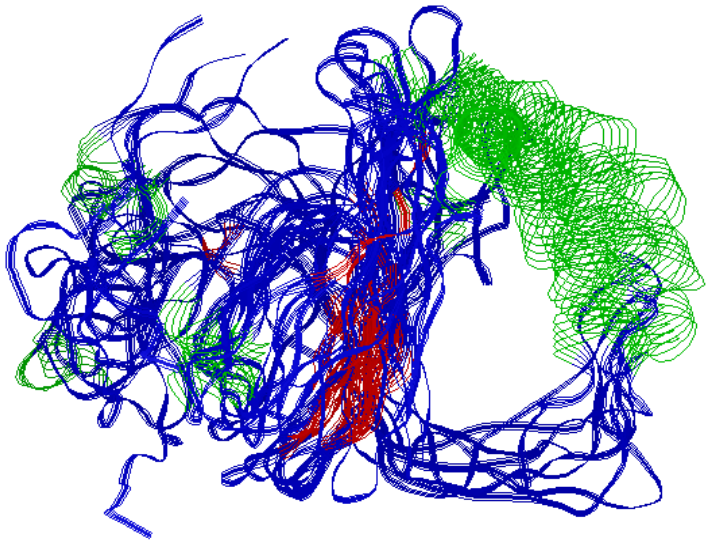

Figure 21: Overlay of Top 10 Isolated D6 Solution Structures. The α-helices are labeled in green, the β-strands in red, and the turns/coils in blue. The average backbone RMSD is 3.0 Å as calculated by CYANA. The NOE cross-peaks used in this simulation are only from the ^{15}N-edited NOESY data set and are ~2/3 assigned.

CYANA structure calculations have assigned 605 total NOE cross-peaks in addition to folding isolated D6. Out of the 605 total NOEs assigned, 457 of these cross-peaks are dedicated to inter-residue NOE interactions. On average, each residue has a total of 6 NOEs with 4 of these being inter-residue NOEs. Figure 22 displays the assigned NOEs per residue results which reveal that the most defined region of the structure belongs to residues 7-59. The least defined residues in this current fold are 60-88, which reside in the C-terminal half of the structure.

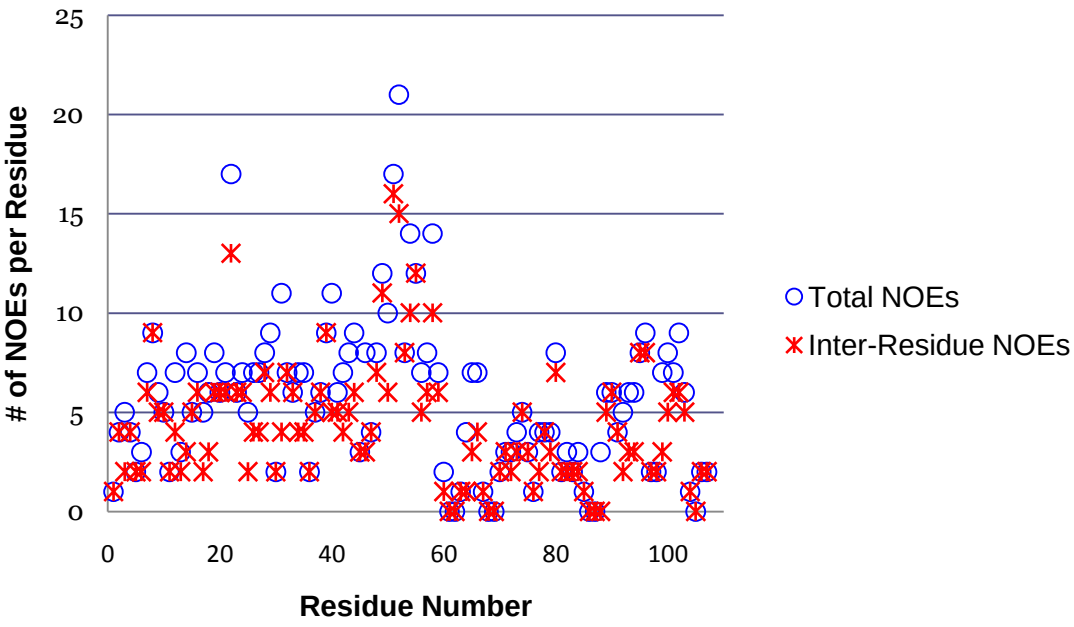

Figure 22: Analyses of Assigned NOEs per Residue. Currently there are 605 total NOEs assigned and 457 of these are inter-residue NOEs. On average, each residue has a total of 6 NOEs and out of the total 4 are inter-residual NOEs. The most defined region of the structure is residues 7-59. The least defined region of the structure is residues 60-88 and the termini.

To evaluate inter-residue NOE cross-peaks further, a plot with both axes designated to residue number identifies the specific residues involved in the inter-residue NOE interactions (Figure 23). In this analysis it is seen that a majority of inter-residue NOEs are short-range interactions between adjacent residues. This is displayed in the step-wise diagonal pattern plotted. According to strong inter-residue NOE peaks of adjacent residues seen in the NOESY spectrum and unambiguously assigned by CYANA we are confident in our backbone assignment of D6 and related NOE cross-peaks. The residues that are distant in sequence, but rather close in space and evolve a NOE cross-peak are extremely important to the prediction of the 3D fold of a protein. There are only a few long-range NOEs, which reside away from the diagonal plot and are defined more in the N-terminal half of D6. This data directly relates to the N-terminus being better defined because these residues have more NOE interactions occurring between them and with a larger proportion of them being long-range.

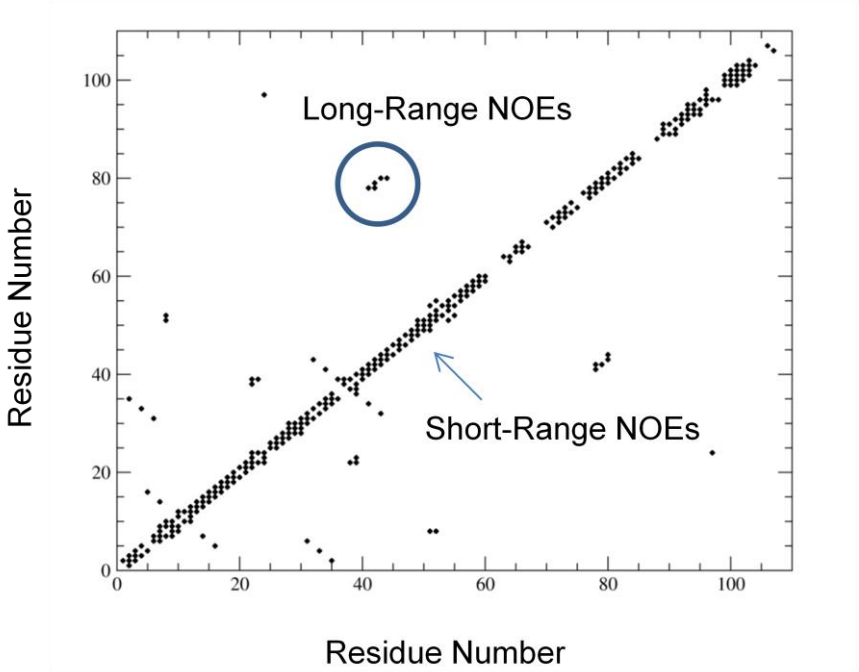

Figure 23: Plot of NOE Interactions between Residues. Short-range NOEs form a diagonal pattern due to most residues NOE interactions with adjacent residues. Long-range NOEs are seen away from the diagonal. Most long-range NOEs appear with N-terminal residues.

A final solution structure of a protein plays a pivotal role in understanding the functional properties of a protein. The refinement of the isolated D6 solution structure is the first task at hand in the future direction of this work. Utilizing two extra multidimensional heteronuclear experiments, recently collected HCCH-TOCSY and the ^{13}C-edited NOESY, the side chains of each residue can be further defined increasing the resonance assignment for D6, as well as lowering the RMSD value.

Once the structures of isolated D6 converge to nearly atomic resolution, the long helix conformation can be analyzed. The question remains of whether or not our isolated D6 in the presence of 5 mM calcium displays a kinked long helix or a straight

long helix. The Gelsolin domain six, which is free of calcium displays a kinked long α-helix at residues 53-56. Our Villin D6 shows the long α-helix to be slightly bent, but as of yet, there is not a distinct kink at residues 53-56. Unique NOE cross-peaks which could possibly designate this similar kinked-helix feature need to be located (if any). Residues Lys46 and Asp30 are conserved in both Villin and Gelsolin and could form a major interaction between their two side chains, and could cause the kink of the long helix in Gelsolin. In addition, residue Glu52 is also conserved between the two proteins, and this residue could interact with residue 7 which could also stabilize the kink. Residue 7 is not conserved; it is a serine in Villin and an asparagine in Gelsolin.

In terms of calcium being bound or not we are still uncertain and an investigation of calcium-free isolated D6 via ^{15}N-HSQC is necessary. This additional information could locate a calcium binding site (if any), by the observation of chemical shift differences seen between ^{15}N-HSQC spectra of "calcium present" and "calcium absent" D6 environments. Chemical shift differences can be global or localized to a calcium binding site. The calcium binding site of Gelsolin's domain six involves residues Asn29, Asp30, and Glu52 so these residues should be observed closely in our Villin D6 structure (Smirnov, et al., 2007).

Other future directions of this project will be performing F-actin binding assays on isolated D6 in order to identify the cryptic F-actin binding site within this structured domain. In the end, our acquired NOESY data for D6-HP along with the final solution structure of isolated D6 will aid in the elucidation of D6-HP solution structure and Ca^{2+}-binding sites residing in the fragment.

Conclusions

The research project presented, utilized the C-terminal modular domains D6 and D6-HP of Villin from *Gallus gallus*, to further investigate the structural and functional properties of this principal cytoskeleton regulating protein. Our specific aims were 1) to express and purify the isotopically labeled C-terminal fragment D6-HP, and, D6 in isolation, with high yield, in the presence of calcium, and with proper protein concentration for the structural characterization by solution NMR; 2) to perform the heteronuclear NMR resonance assignment (^{1}H, ^{13}C, ^{15}N) of D6 and 3) to determine the solution structure of isolated D6.

The expression and purification of isotopically labeled D6 and D6-HP was fulfilled (Aim 1). The ^{13}C/^{15}N-D6 sample was used in multi-dimensional heteronuclear NMR experiments to assign the backbone and side chain NMR resonances (^{1}H, ^{15}N, ^{13}C) for this isolated fragment. With ~91% of NMR resonances assigned, sufficient for the structural characterization, Aim 2 has been achieved. Analysis of the ^{15}N-HSQC spectral overlay of isolated D6 and D6-HP and the calculated chemical shift differences of the backbone resonances from both, revealed that D6 adopts a similar fold in isolation and in the modular fragment D6-HP. Lastly, all the major elements of the secondary structure have been identified and the D6 “backbone fold” determined based on the NMR dihedral and distance restraints available, is concurrent with a typical Gelsolin-like fold.

In conclusion, the refinement of the D6 NMR solution structure needs to be finalized in order to locate the possible calcium and F-actin binding sites and address any protein interactions within the Villin core domain D6, headpiece, and linker.

Materials and Methods

Rich Medium Protein Expression

D6/D6-HP Expression: The D6-HP expression plasmid based on a pET-24a vector (Novagen), was obtained previously; the construction of the D6-HP plasmid has been described previously (Smirnov et al., 2007). The expression plasmid for the D6 fragment was obtained from the D6-HP coding vector by inserting a TGA stop codon after the 107th codon with a Quick-Change site-directed mutagenesis kit (Stratagene, La Jolla, CA). Both D6-HP and D6 protein fragments, sequenced from Villin (*Gallus gallus*), were overexpressed in *Escherichia coli* (*E. coli*) cell line BL21-(DE3) (Novagen) in a similar way. First, 50 mL of overnight cultures (incubated 16-18 hours at 37 °C) were used to inoculate 1 L of Luria-Bertani broth (LB) containing 10 μg/ml of Kanamycin in 2.8 L Fernbach flasks. All cultures were incubated at 37 °C in a Lab Line incubator shaking the flasks at 150 rpm. The cells proliferated until mid-log phase (OD_{600} = 0.4-0.6), at which time they were induced with 0.8 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and harvested after 5 hours by centrifugation at 4,000 g for 20 minutes. The cell pellet was then resuspended in ~30 mL of lysis buffer (50 μg/mL lysozyme, 20 mM sodium phosphate buffer pH 7.0, 200 mM NaCl, 50 mM DTT, 5 mM $CaCl_2$, and 2 mM PMSF) and stored in the -80 °C freezer.

Minimal Medium Protein Expression

$^{13}C/^{15}N$- D6 Expression: The cell growth first occurred in rich medium to obtain a high cell density. These cells were then transferred to minimal medium where ^{15}N and ^{13}C isotopic labeling occurred (Marley et al., 2001).

An overnight culture consisted of 100 mL of LB broth in a 1 L flask, with appropriate Kanamycin concentration and inoculated with a single transformed *E. coli* colony. This was left to shake for 16-18 hours at 37 °C. This overnight culture was separated evenly into four, 2.8 L Fernback flasks, each containing 1 L of LB broth and 10 μg/mL Kanamycin and incubated at 37 °C with shaking at 150 rpm. At an OD_{600} between 0.6-0.8, the cells were harvested by centrifugation at 4,000 g for 20 minutes. The pellet was resuspended and washed in 1 L of sterilized M9T minimal media, (6.0 g Na_2HPO_4, 3.0 g KH_2PO_4, 0.5 g NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 10 μM $FeCl_3$ per 1 L) and 10 μg/mL Kanamycin, but lacked the isotope labeling chemicals. The cells were reharvested and then transferred to ^{13}C and ^{15}N-labeled M9T medium complemented with 0.8 g/L $^{15}NH_4Cl$ and 3.0 g/L ^{13}C6-d-glucose for the expression of the ^{15}N-and ^{13}C-labeled recombinant proteins. The cells were equilibrated in the labeled M9T medium for 1 hour, then induced with 0.8 mM IPTG, and harvested after 5 hours by centrifugation at 4,000 g for 20 minutes. All cultures were incubated at 37 °C, shaking the flasks at 150 rpm. The cell pellet was then resuspended in ~30 mL of lysis buffer (50 μg/mL Lysozyme, 20 mM Sodium Phosphate buffer pH 7, 200 mM NaCl, 50 mM DTT, 5 mM $CaCl_2$, and 2 mM PMSF) and stored in the -80 °C freezer.

Protein Purification

Lysis: The resuspended cell pellets were thawed and then put on ice for lysis via sonication with a Branson Instruments, Inc. Sonifier (model 450) equipped with a 1.8 cm diameter horn. Two rounds of sonication were performed for 30 seconds each, with 2 minute resting intervals in between rounds using a 50% duty cycle and a power output level of 6. The lysate was incubated for 30 minutes at room-temperature on the bench with 10 mM $MgCl_2$ and 1 μg/mL DNAse I added. The lysate was then centrifuged at 40,000 g for 25 minutes. The supernatant was then filtered with a 20 mL BD Luer-Lok Tip syringe through 0.45 μm and then 0.2 μm cellulose acetate VWR sterile syringe filters. The final volume of supernatant was ~25-30 mL.

Gravity Flow Liquid Chromatography; Size Exclusion Column: The supernatant (unconcentrated; ~25-30 mL) was loaded onto the Sephadex G-50 beaded gel stationary phase in our Kontes Chromaflex Chromatography column (2.5 cm x 100 cm). The column was prepared for loading by washing with 2 column volumes of buffer (20 mM PIPES pH 7.0, 5 mM dithiothretol (DTT), 5 mM $CaCl_2$, 150 mM NaCl, and 0.02 % NaN_3). The buffer was filtered before use with GelmanSciences Metricel membrane 0.2 μm filter. The flow rate was set to 1 mL/min for protein fraction collection. D6 or D6-HP fractions were collected using an Advantec SF-2120 Super Fraction Collector, which was set to 70 tubes with 100 drops/tube (~2-3 mL). Out of the 70 tubes collected with this setting, fractions 22-34 had significant concentrations of D6-HP and fractions 46-59 had significant concentrations of D6. The D6/D6-HP containing fractions were

determined by spot tests and SDS-PAGE. Fractions 22 & 23 were combined and concentrated to ~1 mL. This process was repeated for fractions 24 & 25, 26 & 27…etc., together yielding 6 pooled fractions. Samples of D6/D6-HP were concentrated using 15 mL Amicon Ultra Centrifugal units with an Ultracel membrane (3,000 MWCO) made by Millipore and centrifugation at 3350 g for 30 minutes.

High Performance- Size Exclusion Chromatography (HP-SEC): Gravity flow size exclusion purified D6/D6-HP samples were purified by HP-SEC using a GE Healthcare Superdex 75 10/300 GL size exclusion column (~24 mL bed volume) connected to a Waters 650E Advanced Protein Purification System. A D6 sample was purified by HP-SEC using the same size exclusion column; however it was connected to a Biocad Sprint Perfusion Chromatography System. A HiLoad 16/60 Superdex 75 Prep Grade size exclusion column (120-124 mL column volume) was used to HP-SEC purify a D6-HP sample on the Biocad workstation. The flow rate was always set to 1 mL/min. The columns were prepared for loading of D6/D6-HP protein samples by washing with 2-3 column volumes of buffer (same buffer used as above). The buffer was filtered before use through a GelmanSciences Metricel membrane 0.2 μm filter. Maximum sample volumes of 500 μL were injected into the Superdex 75 10/300 GL size exclusion column (2 injections per 1 mL concentrated D6/D6-HP sample) and sample volumes up to 5 mL were injected into the HiLoad 16/60 Superdex 75 Prep Grade size exclusion column. The absorbance was monitored at 280 nm using a Waters 486 Tunable Absorbance Detector and recorded by a NGI Servogor 124 for the first system and

monitored by the Biocad Sprint internal Absorbance flow cell for the second system. All peaks with a significant absorbance at 280 nm were collected as fractions and then analyzed by SDS-PAGE for purity determination.

Protein SDS-PAGE Characterization

Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis (SDS-PAGE): This is a technique used to evaluate the abundance and purity of D6 (12.4 kDa) and D6-HP (23.5 kDa) protein samples throughout expression and purification steps. Gels prepared for this experiment contained 7.5 % acrylamide in the stacking gel and 15 % acrylamide in the resolving gel. The sample loading volumes ranged from 3-30 μL, depending on concentration of protein sample. Protein expression samples were mixed with the appropriate volume of cell cracking buffer (120 mM Tris pH 6.8, 4 % SDS, 0.2 % Bromophenol Blue, 6 % glycerol, and 1 % β-mercaptoethanol (BME)) and heated at 100 °C for 2 minutes, then vortexed before loading the sample onto the gel. Protein purification samples were mixed with appropriate volume of 5X loading buffer (250 mM Tris pH 6.8, 500 mM DTT, 10 % SDS, 0.2 % Bromophenol Blue, 50 % glycerol) and then applied to the gel. The SDS-PAGE running buffer contained 25 mM Tris, 200 mM glycine, 0.1 % SDS. Gels were run at 90 V through the stacking gel and once the samples reached the resolving gel the voltage was increased to 115 V. SDS-PAGE took approximately 2-3 hours per run. Gels were then stained with Coomassie Brilliant Blue staining solution (50 % methanol, 10 % acetic acid, 0.1 % Coomassie Brilliant Blue R250) and then destained with 1: 3: 4 glacial acetic acid: methanol: water solution.

Spot Tests: This is a relatively quick analysis to identify protein-containing samples with a rough estimate of total protein concentration. These samples included fractions from gravity flow size exclusion chromatography or the flow-through created from concentrating protein samples. Sample volumes range from 3-10 μL depending on concentration of the protein sample. The sample is applied, as a spot, onto Whatman chromatography paper (0.33 mm thick; medium flow rate) and allowed to completely dry. Then the paper is incubated at room-temperature in the spot test staining solution (0.25% Coomassie Brilliant Blue R250, 45% Methanol, 10% glacial acetic acid) on the shaker for 20 minutes. The spot test is then destained with SDS-PAGE destaining solution for 15 minutes. The stain only remains visible on the spot if any protein is present in the sample. The total amount of protein in the sample was estimated by the intensity of the dark-blue color of the spot with the spot intensity of another protein sample with a known concentration.

UV-Vis Spectroscopy: D6/D6-HP protein samples were diluted in dd H_2O to obtain an absorbance reading in the UV spectrum 220-360 nm between 0.1 and 1.0. The UV spectra of the samples were taken in 70μL cuvettes with a 4.5 mm path length using a Hewlett Packard 8452A diode array spectrophotometer. Protein concentrations were estimated using the Beer-Lambert equation ($A=\varepsilon lc$) and the extinction coefficients, $\varepsilon_{D6\text{-}HP}$= 37,595 $M^{-1}cm^{-1}$ or ε_{D6}= 32,000 $M^{-1}cm^{-1}$.

Nuclear Magnetic Resonance Analysis

NMR Sample Preparation: D6-HP and D6 NMR samples were prepared by exchanging the SEC buffer for NMR buffer using 15 mL Amicon Ultra Centrifugal units with an Ultracel membrane (3,000 MWCO) made by Millipore and centrifugation at 3350 g for 30 minutes. The NMR buffer contained: 10% D_2O or 100% D_2O, 5mM $CaCl_2$, 10mM d_{10}-DTT, 0.02% NaN_3, and 20mM d_{18}-PIPES (pH 7.0), made to a final volume of ~300 μL in a Shigemi NMR tube (Shigemi Inc.). The total protein concentration was ~1 mM. The pH of the buffer was adjusted with no correction for the effect of D_2O.

NMR Data Collection and Processing: Each sample's purity was first checked via ^{15}N-HSQC on the local 500 MHz Varian Inova NMR Spectrometer. The 2D and 3D NMR data for each protein sample was acquired at 25°C on Varian INOVA 720MHz and 600MHz spectrometers equipped with room-temperature probes located at National High Magnetic Field Laboratory at Florida State University in Tallahassee, FL. The processing of all the NMR data was done using the NMRPipe software (Delaglio, et al. 1995). Table 3 lists all the heteronuclear NMR experiments used to perform the resonance assignment (^{1}H, ^{13}C, ^{15}N) for the D6 fragment to > 90% completion. The experiments were performed at 600 or 720 MHz field strength with the relevant ranges of the respective dimensions. The stability of the protein samples was checked periodically with a 1D ^{1}H or 2D ^{15}N-HSQC. If degraded/deteriorated, the samples received a "d-DTT upgrade" (10 μL of 5 mM d-DTT was added) or made anew.

Chemical shifts for all nuclei are relative to the internal standard 3-(trimethylsilyl) tetradeuterosodium propionate (TMSP) (Wishart et al., 1995).

Table 3: Multidimensional Heteronuclear NMR Experiments and their respective Dimensions performed on D6.

	^{1}H (ppm)	^{13}C (ppm)	^{15}N (ppm)
Experiments	**Center/ Spec.Width**	**Center/ Spec.Width**	**Center/ Spec.Width**
HNCACB	4.85/ 15.78	48.8/ 74.7	118.5/ 30.1
CBCA(CO)NH	4.85/ 15.78	48.8/ 74.7	118.5/ 30.1
^{15}N-HSQC	4.85/ 15.78		118.5/ 30.1
HNCO	4.85/ 15.78	176.6/ 20.7	118.5/ 30.1
HN(CA)CO	4.85/ 15.78	176.6/ 20.7	118.5/ 30.1
HCC(CO)NH	4.85/ 12.99		118.5/ 30.1
HBHA(CO)NH	4.85/ 15.78		118.5/ 30.1
HNHA	4.85/ 15.78		118.5/ 30.1
HNHB	4.85/ 15.78		118.5/ 30.1
CC(CO)NH	4.85/ 15.78	48.8/ 74.7	118.5/ 30.1
HCCH-TOCSY	4.82/ 15.99	48.8/ 74.7	
NOESY-H_2O (150 ms)	4.82/ 15.99		
NOESY-D_2O (150 ms)	4.81/ 15.79		
TOCSY-H_2O (80 ms)	4.82/ 15.99		
TOCSY-D_2O (60 ms)	4.81/ 15.79		
^{15}N-NOESY (80 ms)	4.85/ 12.50		118.5/ 30.1
^{15}N-TOCSY (80 ms)	4.85/15.78		118.5/ 30.1
^{13}C-NOESY (80 ms)	4.85/ 12.50	48.8/ 74.7	

The ^{15}N-HSQC spectrum of the isolated headpiece domain (HP) was recorded at 25°C on a Bruker DMX 500MHz spectrometer at the Boston University Core Facility for Structural NMR. The assignment of the backbone ^{1}H and ^{15}N resonances of HP at 25°C was performed by matching the cross-peaks to those published for an HP sample at 20°C (BMRB entry 4428). Most of the heteronuclear (^{1}H, ^{15}N, ^{13}C) NOE data for D6-HP were recorded previously (Smirnov et al., 2007) and the D6-HP NMR resonances were reported in the same paper (BMRB entry 15097).

NMR Resonance Assignment: The backbone resonance (^{1}H, ^{15}N, and ^{13}C) assignment of the D6 sample was performed through a combined investigation of 2D ^{15}N-HSQC (Palmer, et al., 1991; Kay, et al., 1992) and the following 3D NMR data sets: HNCACB (Wittekind, et al., 1993; Muhandiram, et al., 1994), CBCA(CO)NH (Grzesiek, S. and A. Bax, 1992), HNCO (Grzesiek, S. and A. Bax, 1992, 1993), HN(CA)CO (Yamazaki, et al., 1994), HNHA (Vuister, G.W. and A. Bax, 1993), HNHB (Archer, et al., 1991), HBHA(CO)NH (Grzesiek, S. and A. Bax, 1993), ^{15}N-TOCSY (Clore, et al., 1990), HCC(CO)NH (Grzesiek, et al,1993), CC(CO)NH (Grzesiek, et al., 1993), and HCCH-TOCSY (Clore, et al., 1990). The side chain resonance (aromatic ^{1}H) assignments were obtained from 20ms and 60ms 2D TOCSY (in 100% $^{2}H_2O$), 40ms and 80ms 2D TOCSY (in 10% $^{2}H_2O$), 80ms ^{15}N-NOESY, 150ms and 250ms 2D NOESY (in 100% $^{2}H_2O$ and in 10% $^{2}H_2O$) (States, et al., 1982; Yamazaki, et al., 1994). The visualization and analysis of the NMR spectra and cross-peak picking were performed with NMRView (Johnson, 1994). The sequence specific backbone assignment was performed with the PINE online server (Bahrami, et al., 2009).

3D-Structure Characterization

Secondary Structure Element Detection: The secondary structure content via dihedral angle geometry of D6 and D6-HP was predicted by TALOS software (Cornilescu et al., 1999). This calculation is based on the chemical shift values of the backbone $^{1}H_{\alpha}$, ^{15}N, $^{13}C_{\alpha}$, and ^{13}CO and side chain $^{13}C_{\beta}$ chemical shift values. The standard TALOS reference protein database contains 3D structures of 186 proteins,

which provides more than 24,000 residue triplets for the secondary structure calculation (Cornilescu et al., 1999).

NOESY Assignment and Structure Calculation and Refinement: The homonuclear 2D ^{1}H-NOESY spectra was recorded for all the samples at 720 MHz magnetic field strength. Having the NOE data recorded at this field strength allowed the utilization of the powerful algorithm implemented in the CYANA software (L.A. Systems, Inc.) (Guntert, et al., 2003; 2004). The assignment of the off-diagonal NOE cross-peaks and protein structure determination are performed concurrently in a fully automated way. Figure 24 summarizes what a CYANA explicit input file reads to start a 3D structure calculation.

```
peaks := d6_720_noesy15nhsqc_cyana_b.xpk,d6_720_noesy_150ms_refined_Z_manual.xpk
prot := d6_Y2                          # names of chemical shift lists
constraints := d6_talos_cyana_L.aco    # dihedral constraints
tolerance := 0.025,0.025,0.30          # chemical shift tolerances
calibration :=                         # NOE calibration parameters
structures  := 500,20                  # number of initial, final structures
steps       := 10000                   # number of torsion angle dynamics steps
rmsdrange   := 8..80                   # residue range for RMSD calculation
randomseed  := 434726                  # random number generator seed

nproc=2                                # utilize 2 CPUs if available

  subroutine KEEP                      # keep the manually assigned NOEs unchanged
    peaks select "*, *"
  end

noeassign peaks=$peaks prot=$prot keep=KEEP
```

Figure 24: CYANA Explicit Input Source Code for D6 Solution Structure Calculation. File that CYANA uses as a guideline throughout structure calculations.

We are currently applying over 700 ^{15}N-NOE cross-peaks and ~200 dihedral angle restraints based upon TALOS (Cornilescu, et al., 1999). The approach for computing a structure was patterned after that of Guntert, et al, 2003, 2004. CYANA created 200 starting structures based on the D6 primary sequence. Each structure was subject to 10000 steps of torsion angle dynamic calculations during 7 cycles of restrained molecular dynamics using the 900 constraints given. In reference to the chemical shift assignments for D6, CYANA assigns inter-and intra-residue distances provided by the NOE cross-peak file. In the final cycle, the descent energy minimization of the structure calculations occurs and the 20 best producing RMSD structures are given as the 20 final structures. The structures were analyzed by a combined use of the statistics reported by CYANA and visual inspection with RasMol (Sayle and Milner-White, 1995)**.**

References

Andre, E., F. Lottspeich, M. Schleicher, and A. Noegel, "Severin, gelsolin, and villin share a homologous sequence in regions presumed to contain f-actin severing domains." *J. Biol. Chem.*, **1988**. 263: p. 722-27.

Archer, S.J., M. Ikura, D.A. Torchia, and A. Bax, "An alternative 3D NMR technique for correlating backbone ^{15}N with side chain H_β resonances in larger proteins." *J. Magn. Reson.*, **1991**. 95: p. 636-41.

Arpin, M., E. Pringault, J. Finidori, A. Garcia, J.-M. Jeltsch, J. Vandekerckhove, and D. Louvard, "Sequence of human villin: a large duplicated domain homologous with other actin-severing proteins and a unique small carboxy-terminal domain related to villin specificity." *J. Cell Biol.*, **1988**. 107: p. 1759-66.

Athman, R., D. Louvard, and S. Robine, "Villin enhances hepatocyte growth factor-induced actin cytoskeleton remodeling in epithelial cells." *Mol. Biol. Cell*, **2003**. 14: p. 4641-53.

Athman, R., M.I. Fernandez, P. Sansonetti, D. Louvard, D. Philpott, and S. Robine, "Shigella flexneri infection is dependent on villin in the mouse intestine and in primary cultures of intestinal epithelial cells." *Cell Microbiol.*, **2005**. 7(8): p. 1109-16.

Bader, M.F., J.M. Trifaro, O.K. Langley, D. Thierse, and D. Aunis, "Secretory cell actin-binding proteins: identification of a gelsolin-like protein in chromaffin cells." *J.Cell Biol.*, **1986**. 102: p. 636-46.

Bahrami, A., A.H. Assadi, J.L. Markley, and H.R. Eghbalnia, "Probabilistic interaction network of evidence algorithm and its application to complete labeling of peak lists from protein NMR spectroscopy." *PLoS Comput. Biol.*, **2009**. 5: e1000307.

Bazari, W.L., P. Matsudaira, M. Wallek, T. Smeal, R. Jakes, and Y. Ahmed, "Villin sequence and peptide map identify six homologous domains." *Proc. Natl. Acad. Sci. U S A*, **1988**. 85(14): p. 4986-90.

Boyer, B., A.M. Valles, and J.P. Thiery, "Model systems of epithelium-mesenchyme transitions." *Acta. Anat.*, **1996**. 156: p. 227-239.

Bretscher, A. and K. Weber, "Villin is a major protein of the microvillus cytoskeleton which binds both G and F actin in a calcium-dependent manner." *Cell*, **1980**. 20(3): p. 839-47.

Bretscher, A. and K. Weber, "Fimbrin, a new microfilament-associated protein present in microvilli and other cell surface structures." *J. Cell Biol.*, **1980**. 86: p. 335-40.

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

Burtnick, L.D., E.K. Koepf, J. Grimes, E.Y. Jones, D.I. Stuart, P.J. McLaughlin, and R.C. Robinson, "The crystal structure of plasma gelsolin: implications for actin severing, capping, and nucleation." *Cell,* **1997**. 90: p. 661-70.

Cambell, N.A. and J.B. Reece, "Animal Nutrition: Absorption of Nutrients", Biology Sixth Edition. **2002.** Pearson Education Inc, San Francisco, CA.

Cavanagh, J., W.J. Fairbrother, A.G. Palmer, and N.J. Skelton, Protein NMR Spectroscopy: Principles and Practice. **1996.** Academic Press, San Diego, CA.

Chen, Y., N. Takizawa, J.L. Crowley, S.W. Oh, C.L. Gatto, T. Kambara, O. Sato, X. Li, M. Ikebe, and E.J. Luna, "F-actin and myosin ii binding domains in supervillin." *J. Biol. Chem.,* **2003**. 278(46): p. 46094-106.

Clore, G.M., A. Bax, P.C. Driscoll, P.T. Wingfield, and A.M. Gronenborn, "Assignment of the side-chain ^{1}H and ^{13}C resonances of interleukin-1β using double- and triple-resonance heteronuclear three-dimensional NMR spectroscopy." *Biochemistry,* **1990**. 29(35): p. 8172-84.

Cornilescu, G., D.A. Vinarov, E.M. Tyler, J.L. Markley, and C.C. Cornilescu, "Solution structure of a single-domain thiosulfate sulfurtransferase from arabidopsis thaliana." *Prot. Sci.,* **2006**. 15(12): p. 2836-41.

Cornilescu, G., F. Delaglio, and A. Bax, "Backbone angle restraints from searching a database for chemical shift and sequence homology." *J. Biolmol. NMR,* **1999**. 13: p. 289-302.

Costa de Beauregard, M., E. Pringault, S. Robine, and D. Louvard, "Suppression of villin expression by antisense RNA impairs brush border assembly in polarized epithelial intestinal cells." *EMBO,* **1995**. 14(3): p. 409-21.

Delaglio, F., S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, "NMRPipe: a multidimensional spectral processing system based on UNIX pipes." *J. Biomol. NMR,* **1995**. 6(3): p. 277-93.

Elsasser, H.P., G. Kloppel, H.G. Mannherz, K. Flocke, and H.F. Kern, "Immunohistochemical demonstration of villin in the normal human pancreas and in chronic pancreatitis." *Histochemistry,* **1991**. 95: p. 383-90.

Ezzel, R.M., J. Leung, K. Collins, M.M. Chafel, T.J. Cardoza, and P.T. Matsudaira, "Expression and localization of villin, fimbrin, and myosin I in differentiating mouse F9 teratocarcinoma cells." *Dev. Biol.,* **1992**. 151(2): p. 575-85.

Ferrary, E., M. Cohen-Tannoudji, G. Pehau-Arnaudet, A. Lapillonne, R. Athman, T. Ruiz, L. Boulouha, F. El Marjou, A. Doye, J. Jacques Fontaine, C. Antony, C. Babinet, D. Louvard, F. Jaisser, and S. Robine, "*In vivo,* villin is required for $Ca^{(2+)}$-dependent f-actin disruption in intestinal brush borders." *J. Cell Biol*., **1999**. 146(4): p. 819-30.

Friederich, E., C. Huet, M. Arpin, and D. Louvard, "Villin induces microvilli growth and actin redistribution in transfected fibroblasts." *Cell,* **1989**. 59: p. 461-75.

Friedrich, E., K. Vancompernolle, D. Louvard, and J. Vandekerckhove, "Villin function in the organization of the actin cytoskeleton: correlation of *in vivo* effects to its biochemical activities *in vitro*." *J. Biol. Chem.,* **1999**. 274(38): p. 26751-760.

Glenney, J.R., A. Bretscher, and K. Weber, "Calcium control of the intestinal microvillus cytoskeleton: its implication for the regulation of microfilament organizations." *Proc. Natl. Acad. Sci. USA,* **1980**. 77: p. 6458-62.

Glenney, J.R. and K. Weber, "Calcium control of microfilaments: uncoupling of the f-actin –severing and –bundling activity of villin by limited proteolysis *in vitro*." *Proc. Natl. Acad. Sci. USA,* **1981**. 78(5): p. 2810-14.

Grzesiek, S. and A. Bax, "Improved 3D triple-resonance NMR techniques applied to a 31 KDa protein." *J. Magn. Reson.,* **1992**. 96: p. 432-40.

Grzesiek, S. and A. Bax, "Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR." *J. Am. Chem. Soc.,* **1992**. 114: p. 6291-93.

Grzesiek, S. and A. Bax, "Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins." *J. Biomol. NMR,* **1993**. 3: p. 185-204.

Grzesiek, S., J. Anglister, and A. Bax, "Correlation of backbone amide and aliphatic side-chain resonances in $^{13}C/^{15}N$-enriched proteins by isotropic mixing of ^{13}C magnetization." *J. Magn. Reson., Ser. B,* **1993**. 101: p. 114-19.

Guntert, P., "Automated NMR structure calculation with CYANA." *Meth. Mol. Biol.,* **2004**. 278: p. 353-78.

Guntert, P., "Automated NMR protein structure calculation." *Prog. NMR Spectrosc.,* **2003.** 43: p. 105-25.

Hagiwara, C., M. Tanaka, and H. Kudo, "Increase in colorectal epithelial apoptotic cells in patients with ulcerative colitis ultimately requiring surgery." *Gastroenterol. Hepatol.* **2002**. 17: p. 758-64.

Hellweg, T., H. Hinssen, and W. Eimer, "The Ca^{2+}-induced conformational change of gelsolin is located in the carboxyl-terminal half of the molecule." *Biophys. J.*, **1993**. 65: p. 799-805.

Heymann, D.L., "Shigellosis." American Public Health Association. Control of communicable diseases manual. 18th ed. Washington, DC.: American Public Health Association; **2004**. p. 487–91.http://wwwnc.cdc.gov/travel/yellowbook/2010/chapter-5/shigellosis.aspx

Hirokawa, N., L.G. Tilney, K. Fujiwara, and J.E. Heuser, "Organization of actin, myosin, and intermediate filaments in the brush border of intestinal epithelial cells." *J. Cell Biol.* **1982**. 94: p. 425-43

Janmey, P.A., K. Iida, H.L. Yin, and T.P. Stossel, "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.*, **1987**. 262: p. 12228-36.

Janmey, P.A. and P.T. Matsudaira, "Functional comparison of villin and gelsolin effects of Ca^{2+}, KCl, and polyphosphoinositides." *J. Biol. Chem.*, **1988.** 263(32): p. 16738-43.

Johnson, B.A. and R.A. Blevins, "NMRView: a computer program for the visualization and analysis of NMR data." *J. Biomol. NMR*, **1994**. 4: p. 603-14.

Kay, L.E., P. Keifer, and T. Saarinen, "Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity." *J. Am. Chem. Soc.*, **1992**. 114: p. 10663-665.

Kersting, S., M. Bruewer, G. Schuermann, A. Klotz, M. Utech, M. Hansmerten, C.F. Krieglstein, N. Senninger, J.D. Schulzke, H.Y. Naim, and K.P. Zimmer, "Antigen transport and cytoskeletal characteristics of a distinct enterocyte population in inflammatory bowel diseases." *Am. J. Pathol.* **2004** 165: p. 425-37.

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

Kumar, N. and S. Khurana, "Identification of a functional switch for actin severing by cytoskeletal proteins." *J. Biol. Chem.*, **2004**. 279(24): p. 24915-8.

Kwiatkowski, D.J., T.P. Stossel, S.H. Orkin, J.E. Mole, H.R. Colten, and H.L. Yin, "Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin-binding domain." *Nature*, **1986**. 323: p. 455-58.

Lipari, G., and A. Szabo, “Effect of liberation motion on fluorescence depolarization and nuclear magnetic resonance relaxation in macromolecules and membranes.” *Biophys. J.*, **1980**. 30: p. 489-506.

Lueck, A., H.L. Yin, D.J. Kwiatkowski, and P.G.Allen, “Calcium regulation of gelsolin and adseverin: a natural test of the helix latch hypothesis.” *Biochemistry,* **2000**. 39: p. 5274-79.

Marks, P.W., M. Arai, J.L. Bandura, and D.J. Kwiatkowski, “Advillin (p92): a new member of the gelsolin/villin family of actin regulatory proteins.” *J. Cell Sci.,* **1998**. 111: p. 2129-36.

Markus, M.A., P. Matsudaira, and G. Wagner, “Refined structure of villin 14T and a detailed comparison with other actin-severing domains.” *Protein Sci.*, **1997**. 6(6): p. 1197-209.

Marley, J., M. Lu, and C. Bracken, “A method for efficient isotopic labeling of recombinant proteins.” *J. Biomol. NMR*, **2001**. 20(1): p. 71-5.

Matsudaira, P.T. and D.R. Burgess, “Identification and organization of the components in the isolated microvillus cytoskeleton.” *J. Cell Biol.*, **1979**. 83: p. 667-73.

Matsudaira, P.T. and D.R. Burgess, “Partial reconstruction of the microvillus core bundle: characterization of villin as a Ca^{2+}-dependent actin-bundling/ depolymerizing protein.” *J. Cell Biol.,* **1982**. 92: p. 648-56.

Matsudaira, P.T., R. Jakes, and J.E. Walker, “A gelsolin-like Ca^{2+}-dependent actin-binding domain in villin.” *Nature,* **1985**. 315: p. 248-50.

Meng, J., D. Vardar, Y. Wang, H.C. Guo, J.F. Head, and C.J. McKnight, “High-resolution crystal structures of villin headpiece and mutants with reduced f-actin binding activity.” *Biochemistry,* **2005**. 44: p. 11963-73.

Mooseker, M.S., T.A. Graves, A. Wharton, N. Falco, and C. L. Howe, “Regulation of microvillus structure: calcium-dependent solation and cross-linking of actin filaments in the microvilli of intestinal epithelial cells.” *J. Cell Biol.*, **1980**. 87: p. 809-22.

Muhandiram, D.R. and L.E. Kay, “Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity.” *J. Magn. Reson., Ser. B,* **1994**. 103: p. 203-16.

Nag, S., Q. Ma, H. Wang, S. Chumnarnsilpa, W.L. Lee, M. Larsson, B. Kannan, M. Hernandez-Valladares, L.D. Burtnick, and R.C. Robinson, “Ca^{2+} binding by domain 2 plays a critical role in the activation and stabilization of gelsolin.” *Proc. Natl. Acad. Sci. U S A,* **2009**. 106: p. 13713-8

Okayasu, I., S. Hatakeyama, M. Yamada, T. Ohkusa, Y. Inagaki, and R. Nakaya, "A novel method in the induction of reliable experimental acute and chronic ulerative colitis in mice." *Gastroenterology,* **1990**. 98: p. 694-702.

Palmer, A.G., J. Cavanagh, P.E. Wright, and M. Rance, "Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy." *J. Magn. Reson.,* **1991**. 93: p. 151-70.

Pestonjamasp, K.N., R.K. Pope, J.D. Wulfkuhle, and E.J. Luna, "Supervillin (p205): a novel membrane-associated, f-actin-binding protein in the villin/gelsolin superfamily." *J. Cell Biol.,* **1997**. 139: p. 1255-69.

Philips, M. J., T. Azuma, S.L. Meredith, J.A. Squire, C.A. Ackerley, F.G. Pluthero, E.A. Roberts, R.A. Superina, G.A. Levy, and P.A. Marsden, "Abnormalities in villin gene expression and canalicular microvillus structure in progressive cholestatic liver disease of childhood." *Lancet,* **2003.** 362: p. 1112-19.

Prendergast, G.C. and E.B. Ziff, "Mbh1: A novel gelsolin/severin-related protein which binds actin in vitro and exhibits nuclear localization *in vivo*." *EMBO J.,* **1991.** 10(4): p. 757-66.

Revenu, C., M. Courtois, A. Michelot, C. Sykes, D. Louvard, and S. Robine, "Villin severing activity enhances actin-based motility *in vivo*." *Mol. Biol. Cell,* **2007**. 18: p. 827-38.

Revenu, C., R. Athman, S. Robine, and D. Louvard, "The co-workers of actin filaments: from cell structures to signals." *Nat. Rev. Mol. Cell. Biol.,* **2004**, 5: p. 635-46.

Rimm, D.L., T.E. Holland, J.S. Morrow, and J.M. Anderson, "Autoantibodies specific for villin found in patients with colon cancer and other colitides." *Dig. Dis. Sci.* **1995**. 40 (2): p. 389-95.

Robine, S., C. Huet, R. Moll, C. Sahuquillo-Merino, E. Coudrier, A. Zweibaum, and D. Louvard, "Can villin be used to identify malignant and undifferentiated normal digestive epithelial cells?" *Proc. Natn. Acad. Sci. USA,* **1985**. 82: p. 8488-92.

Robinson, R.C., M. Mejillano, V.P. Le, L.D. Burtnick, H.L. Yin, and S. Choe, "Domain movement in gelsolin: a calcium-activated switch." *Science,* **1999**. 286: p. 1939-42.

Sayle, R.A., E.J. Milner-White, "RasMol: biomolecular graphics for all." *Tr. Biochem. Sci.,* **1995**. 20: p. 374-76.

Silacci, P., L. Mazzolai, C. Gauci, N. Stergiopulos, H. Yin, and D. Hayoz, "Gelsolin superfamily proteins: key regulators of cellular functions." *Cell Mol. Life Sci.,* **2004**. 61: p. 2614-23.

Smirnov, S.L., N.G. Isern, Z.G. Jiang, D.W. Hoyt, and C.J. McKnight, " The isolated sixth gelsolin repeat and headpiece domain of villin bundle f-actin in the presence of calcium and are linked by a 40-residue unstructured sequence." *Biochemistry,* **2007**. 46(25): p. 7488-96.

States, D.J., R.A. Habekorn, and D.J. Ruben, "A two-dimensional nuclear Overhauser experiment with pure absorption phase in four quadrants." *J. Magn. Res.*, **1982**. 48: p. 286-292.

Straub, K.L., M.C. Stella, and M. Leptin, "The gelsolin-related flightless I protein is required for actin distribution during cellularisation in Drosophila." *J. Cell Sci.*, **1996**. 109: p. 263-70.

Teng, Q., "Protein Structure Determination from NMR data", Structural Biology: Practical NMR Applications. **2005**, Springer Science + Business Media, Inc, New York, NY.

Vardar, D., D.A. Buckley, B.S. Frank, and C. J. McKnight, "NMR structure of an f-actin-binding "headpiece" motif from villin." *J. Mol. Biol.*, **1999**. 294(5): p. 1299-1310.

Vuister, G.W. and A. Bax, "Quantitative J correlation: a new approach for measuring homonuclear three-bond $J_{HNH\alpha}$ coupling constants in ^{15}N-enriched proteins." *J. Am. Chem. Soc.*, **1993**. 115(17): p. 7772-77.

Wang, H., S. Chumnarnsilpa, A. Loonchanta, Q. Li, Y.M. Kuan, S. Robine, M. Larsson, I. Mihalek, L.D. Burtnick, and R.C. Robinson, "Helix straightening as an activation mechanism in the gelsolin superfamily of actin regulatory proteins." *J. Biol. Chem.*, **2009**. 284(32): p. 21265-9.

Wang, Y., K. Srinivasan, M.R. Siddiqui, S.P. George, A. Tomar, and S. Khurana, "A novel role for villin in intestinal epithelial cell survival and homeostasis." *J. Biol. Chem.,* **2008**. 283(14): p. 9454-64.

Wishart, D.S., C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, and B.D. Sykes, "^{1}H, ^{13}C, and ^{15}N chemical shift referencing in biomolecular NMR." *J. Biolmol. NMR,* **1995**. 6: p. 135-40.

Wittekind, M. and L. Mueller, "HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the α- and β-carbon resonances in proteins." *J. Magn. Reson., Ser. B,* **1993**. 101: p. 201-05.

Wulfkuhle, J.D., I.E. Donina, N.H. Stark, R.K. Pope, K.N. Pestonjamasp, M.L. Niswonger, and E.J. Luna, "Domain analysis of supervillin, an f-actin bundling plasma membrane protein with functional nuclear localization signals." *J. Cell. Sci.,***1999**. 112: p. 2125-36.

Yamazaki, T., W. Lee, C.H. Arrowsmith, D.R. Muhandiram, and L.E. Kay, "A suite of triple resonance NMR experiments for the backbone assignment of ^{15}N, ^{13}C, ^{2}H labeled proteins with high sensitivity." *J. Am. Chem. Soc.*, **1994**. 116(26): p. 11655-66.

Yin, H., J.H. Hartwig, K. Maruyama, and T.P. Stossel, "Ca^{2+} control of actin filament length." *J. Biol. Chem.*, **1981**. 256: p. 9693-97.

Yin, H.L. and T.P. Stossel, "Control of cytoplasmic actin gel-sol transformation by gelsolin, a calcium-dependent regulatory protein." *Nature*, **1979**. 281: p. 583-86.