Structural and Mechanistic Studies of

two regulatory factors in actin cytoskeletal signaling:

Vav And VopL

APPROVED BY SUPERVISORY COMMITTEE

Michael K. Rosen, Ph.D.

Youxing Jiang, Ph.D.

Jose Rizo-Rey, Ph.D.

Diana R. Tomchick, Ph.D.

This thesis is dedicated to my family

Structural and Mechanistic Studies of two regulatory factors in actin cytoskeletal signaling: Vav And VopL

by

Bingke Yu

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

June, 2011

Acknowledgement

I would like to thank all the people who have helped and inspired me during my doctoral studies. First, I want to give my sincerest gratitude to my advisor, Dr. Michael Rosen, for his guidance throughout my research and study. His enthusiasm for science, wisdom and patience have motivated all his advisees, including me. I thank him for his many insightful comments and suggestions. I also thank him for the support and encouragement he has given to me over the course of my graduate studies.

I want to thank my thesis committee members Dr. Youxing Jiang, Dr. Jose Rizo-Rey, and Dr. Diana Tomchick for providing numerous advices and inspiring suggestions over the past years.

It is a great pleasure for me to thank all past and current members of the Rosen Lab. The lab has been a convivial place to work and I have truly benefited from interacting with everyone on a daily basis. In particular, I would like to thank Dr. Pilong Li, Dr. Ilidio Martins and Dr. Hui-chun Cheng whose work in the lab has provided much of the foundation for my dissertation. Pilong and Ilidio gave me highly valuable instructions and guidance in NMR spectroscopy, and never hesitated to help me whenever I encountered difficulties in work. You guys are great collaborators and deserve my sincere gratitude. And Hui-chun, thank you for always being there when I want to find a person to consult or talk to. You are a wonderful friend for me all this time. I want to thank Dr. Shae Padrick, Dr. Zhucheng Chen, Dr. Ali Yunus and Dr. Xiaolan Yao who have provided me many stimulating scientific discussions, technical help and support for my thesis work. Also my appreciation goes to our five other graduate students, Soyeon Kim, Jacob

Zahm, Sudeep Banjade, Yuan Lin and Salman Banani for the great time we had together. Also thank Dr. Da Jia, Dr. Stone Chen, Dr. Chi Pak, Junko Umetani, and Dr. Takanori Otomo for your friendship and help.

Dr. Diana Tomchick, Dr. Mischa Machius, Dr. Chad Brautigam, Dr. Dominika Borek, thank you all for your support and help in crystallographic studies. Without you guys, I would still be struggling with model building right now. I really enjoyed the time we worked together.

Finally, I must thank my family and friends for all of their love, support, understanding and patience over the last six years. Without their presence in my life, this dissertation would simply have been impossible. In particular, I have to especially thank my parents Jungong Yu and Xiaochun Feng and my sister Ying Yu for their unflagging care and love. My parents have provided me the best possible environment to grow up and attend school. They always believe in me and are supportive of me whenever I make a decision. I am so lucky to have them be my parents. My sister, you are my best friend in my life. Thank you for taking care of me in many aspects since my childhood. I also would like to thank my dearest boyfriend Tingwan Sun for his constant support and encouragement. It is your patient love that enables me to finish this work.

Publications

Yu B., Cheng H.C., Brautigam C., Tomchick D.R. and Rosen M.K. **Mechanism of Actin Filament Nucleation by the Bacterial Effector VopL.** ***Nature Structure and Molecular Biology*** In press 2011

Yu B., Martins I.R., Li P., Amarasinghe G.K., Umetani J., Fernandez-Zapico M.E., Billadeau D.D., Machius M., Tomchick D.R., Rosen M.K.. **Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1.** *Cell.* 2010 Jan 22;140(2):246-56.

Yu W., Hu J., **Yu B.**, Xia W., Jin C., Xia B. **Solution structure of GSP13 from Bacillus subtilis exhibits an S1 domain related to cold shock proteins.** *J Biomol NMR*. 2009 Apr;43(4):255-9.

Yu W., **Yu B.**, Hu J., Xia W., Jin C., Xia B. **1H, 13C, and 15N resonance assignments of a general stress protein GSP13 from Bacillus subtilis.** *Biomol NMR Assign.* 2008 Dec;2(2):163-5.

Structural and Mechanistic Studies of

two regulatory factors in actin cytoskeletal signaling:

Vav And VopL

Bingke Yu

UT Southwestern Medical Center at Dallas, 2011

Supervising Professor: Michael Rosen, Ph.D.

Proper control of actin cytoskeletal dynamics is essential for cell survival. The goals of my thesis work have been to characterize the structural and biophysical properties of two regulatory proteins in actin cytoskeletal rearrangement pathway: Vav and Vibrio outer protein L (VopL).

Vav proteins are guanine nucleotide exchange factors for Rho family GTPases. They play key roles in actin regulatory pathways and control diverse cellular processes like T cell maturation and activation, cell migration and phagocytosis. They belong to a group of multi-domain signaling proteins which display complex behaviors because of the collective regulation from multiple domains. Previous work has shown that Vav is autoinhibited in the resting state through the cooperative suppression of N-terminal Calponin domain and Acidic region, with the physical mechanism yet to be determined. Here through structural, energetic and biochemical studies, I demonstrate that the Calponin homology domain of Vav binds to the Pleckstrin homology domain, restrains the inhibitory helix in the Acidic region, and shifts the Dbl homology domain - inhibitory helix equilibrium to a more closed state. This construction enables strong suppression and an efficient activation process. The energetic basis of Vav autoinhibition may turn out to be widespread in multi-domain systems.

VopL, a pathogenic effector from *Vibrio parahaemalyticus*, is an actin nucleation factor that induces stress fibers during bacterial infection. It contains three N-terminal Wiskott-Aldrich Homology 2 (WH2) motifs and a unique VopL C-terminal domain (VCD). It potently promotes actin filament nucleation *in vitro*. However, the physical basis of VopL mediated nucleation has not been understood. Here I performed structural and biochemical studies to investigate the mechanism of actin filament nucleation by VopL. I found that both the WH2

element and VCD are required for VopL activity. The crystal structure of VCD revealed a U-shaped dimer that is stabilized by a terminal coiled-coil. Dimerization of the WH2 motifs as well as contacts between VCD and actin contribute to the nucleation activity of VopL. My studies suggest the formation of a structurally organized actin cluster involving lateral contacts during nucleation. Stabilization of these lateral contacts may be a common feature of actin filament nucleation by WH2-based factors.

Table of Contents

Title fly i

Dedication ii

Title page iii

Copyright iv

Acknowledgement v

Publications vii

Abstract vii

Table of contents xi

List of figures (41 figures) xiv

List of tables (4 tables) xvii

List of abbreviations (54 abbreviations) xviii

Chapter 1 introduction I 1

Vav biology 1

The structural elements of Vav 10

Mechanism of Vav autoinhibition by Ac region 16

Internal dynamics control activation and activity of the autoinhibited Vav DH domain 19

Cooperative suppression of Vav1 transforming activity through the CH

Domain 20

Cooperative autoinhibition in other multi-domain signaling proteins 26

Activation of Vav through tyrosine phosphorylation 29

Chapter 2 Structural and Energetic Mechanisms of Cooperative Autoinhibition and Activation of Vav1 32

Results 33

Structure of Autoinhibited Vav1 33

Validation of the Structure by Mutagenesis 36

Energetic and Structural Mechanisms of Autoinhibition 51

Validation of the Energetic Mechanism by Mutagenesis 57

Stepwise phosphorylation and activation of Vav1 61

Discussion 64

Materials and Methods 72

Chapter 3 Concluding Remarks I 79

Chapter 4 Introduction II 84

Actin molecule 86

Structure of the actin filament 92

Actin nucleating factors 94

Arp2/3 complex 94

Formin proteins 98

WH2 Based nucleators: Spire, Cobl, Lmod 101

Spire 104

Cobl 108

Lmod 109

Pathogenic Actin nucleators: VopL, VopF, TARP 111

VopL 113

VopF 114

TARP 115

Chapter 5 Mechanism of Actin Filament Nucleation by the Bacterial Effector VopL 117

Results 118

WH2 motifs and VCD are both required for actin nucleation 118

The VopL VCD is a U-shaped dimer stabilized by a coiled-coil 122

Dimerization and actin contacts of the VCD are important for activity 135

VopL activity is sensitive to the length of linker between WH2c and VCD 142

Discussion 146

Methods and Materials 151

Chapter 6 Concluding remarks II 157

List of figures

Figure 1-1 Domain structure of Vav protein 11

Figure 1-2 Structure of DPZ-Rac complex 18

Figure 1-3 Solution structure of Vav1 AD module (pdb ID: 1F5X).....21

Figure 1-4 Model for the regulation and activation mechanism of Vav

AD module 23

Figure 1-5 Structure of Vav3 after EM reconstruction and docking 25

Figure 1-6 Overlaid 1H/13C methyl TROSY spectra of murine AsD and CADPZ proteins (left and right panels, respectively), showing Leu180δ1 and Leu325 δ 1 signals(top and bottom panels, respectively) 28

Figure 2-1 Crystal of Vav1 CADPZ 37

Figure 2-2 Ribbon diagram of Vav1 CADPZ 38

Figure 2-3 Overlay of DPZ elements from the CADPZ and DPZ/Rac

Complex 41

Figure 2-4 Contacts of Y142 (a) andY160 (b) to other residues in the CADPZ crystal structure 42

Figure 2-5 Modulatory contacts contribute to autoinhibition of Vav activity 45

Figure 2-6 Interdomain contacts in Vav1 49

Figure 2-7 Disrupting modulatory contacts increases Vav activty in vivo 50

Figure 2-8 Inhibitory Effects of the CH Domain Require an Intact Ac Element 54

Figure 2-9 Titrations of CH/CA into ADPZ/AsDPZ 55

Figure 2-10 CH-PH and helix-DH equilibria are thermodynamically coupled 60

Figure 2-11 Phosphorylation or mutation of Y142 and Y160 destabilize CH-PH and helix-DH equilibria, increasing kinase access to Y174 62

Figure 2-12 Fluctuations in the Ac Element Enable Stepwise Activation of Vav1 65

Figure 2-13 Structural and energetic models of autoinhibition in Vav1 67

Figure 4-1 Structure of monomeric actin and actin filament 88

Figure 4-2 Domain structures of actin nucleators 91

Figure 4-3 Proposed mechanisms of actin assembly factors 97

Figure 4-4 Structure of WH2 bound actin 106

Figure 4-5 Sequence alignment of WH2 domains 107

Figure 5-1 Domain structure of VopL constructs used and activities of VopL proteins represented by half times 120

Figure 5-2 Isothermal titration calorimetry analysis of the binding of the individual VopL WH2 domains to actin in G-buffer 121

Figure 5-3 Both the WH2 motifs and C-terminal domain contribute to actin nucleation by VopL 123

Figure 5-4 Critical concentration of actin measured to be 0.02 μM, 0.21 μM, 0.40 μM, 0.77 μM in the presence of 0 μM, 1 μM, 2 μM or 4 μM VopL W_3 peptide 124

Figure 5-5 Actin assembly assays contained 4 μM actin and 50 nM W_3-C, W_2-C, W_1-C or VCD 125

Figure 5-6 Size exclusion chromatography—multi-angle laser light scattering (MALLS) analysis of VopL constructs W_3-C (red) and VCD (blue) 127

Figure 5-7 Structure of the VopL VCD dimer 131

Figure 5-8 Alternative VopL VCD dimer from that shown in Figure 5-7 132

Figure 5-9 MALLS analysis of VopL VCD Δ h. Data plotted as in Fig. 5-6 133

Figure 5-10 Structure alignment of the two VopL VCD molecules in the dimer 134

Figure 5-11 Actin polymerization assays performed with 4 μM actin and 50 nM W_3-C, W_2-C , W_1-C, C, W_3-GST, W_2-GST or W_1-GST 137

Figure 5-12 VCD dimerization is important for VopL nucleating activity 138

Figure 5-13 The basic arm of VCD subdomain 2 contributes to VopL nucleating activity 140

Figure 5-14 VCD subdomain 1 may also contact acin and contribute to nucleating activity 141

Figure 5-15 Stability of the VopL dimer 143

Figure 5-16 Increasing the linker between WH2c and the VCD increases actin assembly activity 145

Figure 5-17 Model for a minimal actin nucleus assembled by VopL 149

List of Tables

Table 2-1 Data collection and structure refinement statistics of Vav1 39

Table 3-1 Molecules with multiple weak interactions 82

Table 4-1 Comparison of WH2 based nucleators 102

Table 5-1 Data collection, phasing and refinement statistics of VopL VCD 128

List of Abbreviations

λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
° C	Degrees Celsius
Ac/A	Acidic
ADF	Actin depolymerization factor
Arp2/3	Actin related protein 2/3
A.U.	Arbitrary units
B	Basic
C	Central
Cdc24	Cell division control protein 24
Cdc42	Cell division cycle 42
CH	Calponin Homology
CRIB	Cdc42 and Rac interactive binding domain
DAG	Diacylglycerol
DH	Dbl homolgy
F-actin	Filamentous actin
G-actin	Globular or monomeric actin
GBD	GTPase binding domain
GDI	Guanine nucleotide inhibitor
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GST	Glutathione S-transferase

GTP	Guanosine triphosphate
K_D	Equlibrium dissociation constant
kD	kilo-Dalton
KIND	kinase non-catalytic C-lobe domain
mg	Milligram
MIM	Missing In Metastasis protein
μM	Micromolar
min	Minute
mol	Mole
ms	Millisecond
NMR	Nuclear Magnetic Resonance
nm	Nanometer
nM	Nanomolar
NPF	Nucleation promoting factor
P	Poly-proline region
PEG	Polyethylene glycol
PH	Pleckstrin homolgy
PI3K	Phosphatidylinostol 3-kinase
PIP_2	Phosphatidylinostol 4,5-biphosphate
PIP_3	phosphatidylinostol 3,4,5-triphosphate
PRM	Proline rich motif
Rho	Ras homology
SH2	Src homology 2
SH3	Src homology 3

VASP	VAsodilator-Stimulated Phosphoprotein
VCD	VopL C-terminal domain
VopF/L	Vibrio Outer Protein F/L
WH2	Wiskott-Aldrich Syndrome Protein Homology Domain 2
WASP	Wiskott-Aldrich Syndrome Protein
WAVE	WASP-family VErprolin homolog
ZF	Zinc finger

Chapter 1

Introduction I

Cells exist in stochastically variable environments, and in order to survive, they must sense changes and adapt rapidly and accordingly (Huse and Kuriyan 2002; Pawson and Kofler 2009; Lim 2010). They achieve such adaptations through the integration of sophisticated signaling networks. Multi-domain signaling proteins are delicate molecular machineries utilized to integrate the diverse and rapidly changing inputs into coherent and rational outputs. They are likely to have evolved from the selective pressure to generate new functions. Their prevalence in nature underscores their importance: They constitute a surprising majority of genomic proteins: ~66% in unicellular organisms and more than 80% in metazoans. In some multi-domain signaling proteins, the domains are linked together, but spatially isolated from one other, analogous to beads on a string. In most cases, however, the individual domains interact with one another, creating a circumstance wherein a conformational change in one domain is "sensed" by other domains in the molecule. It is through such allosteric, inter-

domain communication that multi-domain signaling proteins are able to respond to diverse input signals in complex fashion.

Many multi-domain signaling proteins consist of a core function-bearing domain and other interacting regulatory domains. Examples are Src family kinases, Dbl family guanine nucleotide exchange factors (GEFs), and actin nucleation regulatory factors formin proteins and neuronal Wiskott - Aldrich syndrome protein (N-WASP), transcription factor NF-κB etc (Pufall and Graves 2002). These proteins are autoinhibited in the resting state and activated through chemical modification or binding to allosteric activators. Very frequently, the functional domain is inhibited through intramolecular interactions from the regulatory domains, which creates steric or conformational occlusions. Steric occlusions occur when the regulatory element is located in a manner that it physically blocks the ligand from accessing the active site. Conformational autoinhibition, on the other hand, describes the situation when intramolecular interactions between regulatory and functional domains distort the conformation of the active site. Activation of the functional domain usually involves displacement of the regulatory domain or conformational changes in the functional domain.

Through intramolecular interactions, multi-domain systems display complex signal intergration behaviors. For example, the actin nucleation

promoting factor (NPF) N-WASP protein is autoinhibited because of the interactions between the GTPase binding domain (GBD) and the functional C-terminal VCA region (Kim, Kakalis et al. 2000). Binding to small GTPase Cdc42 or the acidic phospholipid phosphatidylinositol 4,5-biphosphate (PIP_2) activates N-WASP activity (Kim, Kakalis et al. 2000; Buck, Xu et al. 2001; Padrick, Cheng et al. 2008). These two inputs, Cdc42 and PIP_2, are poor activators alone, but together potently activate N-WASP. *In vitro* quantitative activity analysis indicated that Cdc42 and PIP_2 act in a highly cooperative fashion with a cooperativity factor close to 300 (Prehoda, Scott et al. 2000). Another autoinhibited multi-domain protein, the cell cycle regulatory protein Cyclin-dependent kinase 2 (Cdk2), is activated through the T-loop T160 phosphorylation and binding to Cyclin A. These two activating inputs, however, appear to be non-cooperative, since the binding affinity between Cdk2 and Cyclin A is not affected by the T-loop phosphorylation state (Brown, Noble et al. 1999). Understanding how the multi-domain systems integrate the signaling inputs will be important to understand the complex cellular networks.

The Dbl family Guanine nucleotide exchange factor (GEF) Vav is a typical multi-domain signaling protein regulated by autoinhibition. One feature of its autoinhibition is the cooperative construction by which the functional Dbl homology (DH) domain is strongly inhibited through the collective actions of the Calponin homology (CH) domain and the Acidic (Ac) region. The strong

autoinhibition relies on multiple interactions, and such architecture is likely to be widespread in multi-domain signaling proteins. However, the structural and thermodynamic basis of the cooperative suppression is unclear. In addition, the mechanism through which strong suppression is relieved in fast activation kinetics remains obscure.

This first part of this thesis (chapter 1-3) is devoted to understanding the structural and energetic basis of Vav cooperative suppression and activation. In this chapter, I will describe the biological importance of Vav proteins. Then I will review Vav biochemistry and the progress that has been made to uncover the regulation mechanism of Vav. Last, I will compare Vav to other multi-domain signaling proteins likely regulated in similar manner. In the second chapter, I describe my efforts to crystallize the regulatory apparatus of human Vav1: the N terminal five domains (CADPZ). The structure of CADPZ was solved through crystallography, revealing two interaction interfaces that are far apart yet both contribute to Vav autoinhibition, and positions of three modulatory tyrosine residues making distinct contacts with the CH domain or the DH domain. I employed NMR chemical shift and functional analyses to quantify the energetic coupling between the two interactions in CADPZ. My biochemical and NMR data together with kinetic studies performed by Ilidio Martins demonstrated that Vav activation involves step-wise phosphorylation events. These studies reveal how a

strongly-suppressed multi-domain system can undergo rapid activation. The principles described herein likely apply to other signaling proteins.

Vav biology

The first vav gene was discovered in 1989 in the laboratory of Mariano Barbacid (Katzav, Martin-Zanca et al. 1989). It was characterized as an oncogene due to its exhibition of transforming activity during gene transfer assays. The vav gene was then mapped by the same group to a region of human chromosome 19 which is involved in karyotypic abnormalities in some hematopoietic malignancies (Martinerie, Cannizzaro et al. 1990). Shortly after that, the full length species of mouse vav1 oncogene and other isoforms in mammalian vav family were isolated, including vav2 and vav3 (Coppola, Bryant et al. 1991; Schuebel, Bustelo et al. 1996; Movilla and Bustelo 1999). Vav in nematodes "*CelVav*" was found in 1994 during the characterization of *Caenorhabditis elegans* genome (Wilson, Ainscough et al. 1994). While Vav1 was found to be expressed in hematopoietic cells, Vav2 and Vav3 were indicated to be ubiquitously expressed (Hornstein, Alcover et al. 2004; Sauzeau, Sevilla et al. 2006).

Although originally identified as an oncoprotein, Vav has been demonstrated as a key player in mediating cellular signaling from membrane receptors to the interior of the cells. They control a variety of cellular processes

including T cell and B cell maturation and activation, neuronal system development and mitogenesis (Bustelo 2001; Tybulewicz, Ardouin et al. 2003; Tybulewicz 2005). Vav proteins are tyrosine phosphorylated upon activation of a variety of receptors, including T-cell receptor (Margolis, Hu et al. 1992), B-cell antigen receptor (Bustelo and Barbacid 1992), cytokine receptors (Evans, Howard et al. 1993), chemokine receptors ((Vicente-Manzanares, Cruz-Adalia et al. 2005)), integrins (Zheng, Sjolander et al. 1996), and growth factor receptors (Alai, Mui et al. 1992). Vav proteins transduce the receptor signals to a number of downstream pathways, including Ras, JNK, ERK, NF-κB and NFAT pathways (Hornstein, Alcover et al. 2004). Among these, the T-cell activation signaling pathway is the most extensively studied. Upon TCR activation, the immunoreceptor tyrosine activation motifs (ITAM) are phosphorylated by Src family kinases such as Lck and Fyn. SH2 containing proteins like tyrosine kinase ZAP-70 are recruited to ITAMs, leading to the phosphorylation and activation of Vav proteins.

Activated Vav proteins function as guanine nucleotide exchange factors (GEFs) for Rho (Ras homology) family GTPases, and catalyze the transition of Rho/Rac/Cdc42 GTPase from GDP bound state to GTP bound state (Adams, Houston et al. 1992; Galland, Katzav et al. 1992) (Crespo, Schuebel et al. 1997; Han, Das et al. 1997; Schuebel, Movilla et al. 1998; Movilla and Bustelo 1999). In 1992, several groups identified a central domain in Vav that is highly

homologous to Dbl Homology domain of the human oncoprotein and GEF, Dbl, the Cdc24p protein of *saccbaromyces cerevisae* and the human Bcr protein (Adams, Houston et al. 1992; Cen, Papageorge et al. 1992; Galland, Katzav et al. 1992). In 1991, Hart and colleagues found that dbl oncogenic product catalyzed nucleotide exchange for Cell division cycle 42 (Cdc42) protein (Hart, Eva et al. 1991). It was realized later that Dbl Homology domain contains ~ 150 amino acids and catalyzes the displacement of nucleotide from Rho GTPases (Worthylake, Rossman et al. 2000). Since the concentration of GTP inside the cell is much higher than GDP (about 10-fold) and GTPase has similar affinities with GTP and GDP, this displacement tends to lead to the binding of GTPase with GTP.

The GEF activity of Vav is the key for relaying T cell activation signals to the actin cytoskeleton. GTP bound Rho proteins were shown to coordinate cell actin cytoskeleton rearrangment and gene trancription activation, and regulate numerous cellular events including cell motility, morphogenesis, gene expression, membrane trafficking, growth factor signaling, and cellular transformation (Van Aelst and D'Souza-Schorey 1997; Mackay and Hall 1998). Rac and Cdc42 activate the actin nucleation promoting factors WASP-family VErprolin homolog (WAVE) and WASP respectively, and mediate actin polymerization through activating actin related protein 2/3 (Arp2/3) complex. Dbl proteins therefore regulate these cell processes through activating Rho GTPases.

In addition to the key function as a GEF, Vav has also been recognized to act as a scaffold protein to recruite PLC-γ1 in the complex of immunological synapse and modulate TCR-dependent calcium signaling (Braiman, Barda-Saad et al. 2006). Its C-terminal first SH3 domain interacts with adaptor potein Grb2, and this interaction is critical for the translocation of Vav to cell membrane and subsequent activation by upstream tyrosine kinases (Tybulewicz 2005). The second SH3 domain forms protein complexes with a variety of proteins, including cytoskeletal regulator Zyxin, RNA binding proteins, transcriptional modulators, viral proteins, ubiquitination factors and Dynamin (Tybulewicz 2005). These various interactions allow Vav proteins to function in multiple signaling pathways.

N-terminal truncation of Vav was found to lead to cellular transformation (Katzav, Martin-Zanca et al. 1989; Abe, Whitehead et al. 1999) (Zeng, Sachdev et al. 2000) (Booden, Campbell et al. 2002). The N-terminal region does not bear catalytic activity, but negatively regulates Vav DH domain (Abe, Whitehead et al. 1999). Truncation of this region thus activates Vav GEF activity. The morphology of cells transformed with oncogenic Vav is reminiscent of those transformed with dbl oncogene, suggesting the GEF function of Vav is important for the transforming activity (Khosravi-Far, Chrzanowska-Wodnicka et al. 1994). Mutation of DH domain abolished the focus formation activity of Vav, confirming the importance of GEF function (Zugaza, Lopez-Lago et al. 2002). In addition to activating Rho family GTPases, oncogenic Vav may increase cellular

transformation by enhancing the expression of several genes such as osteoponin (Schapira V. Lazer G. 1996). Increased osteoponin expression has been shown to be associated with tumor invasion, progression, and metastasis in cancers of stomach, liver, breast, prostate, and colon (Rittling and Chambers 2004). Osteoponin is highly expressed in oncogenic Vav transformed cells, but it has very limited expression in cells transformed with wild type Vav (Schapira, Lazer et al. 2006), suggesting the relationship between Vav and osteoponin.

Aberrant function of Vav proteins is implicated in a variety of cancer cells. For example, upregulation of vav was found to associate with decreased patient survival in pancreatic adenocarcinoma (Fernandez-Zapico, Gonzalez-Paz et al. 2005). Vav1, whose normal expression is restricted to hematopoietic cells, was found to be aberrantly expressed in lung cancer cells. Knockdown of Vav1 in these cells results in reduction of tumor growth in nude mice, suggesting Vav1 plays a critical role in the tumorigenicity of lung cancer cells (Katzav 2009).

Vav has also been identified as the specific binding partner of Nef proteins from HIV-1 (Fackler, Luo et al. 1999). Coexpression and binding of these partners initiates profound morphological changes, cytoskeletal rearrangements and the JNK/SAPK signaling cascade, leading to increased levels of viral transcription and replication. Vav deficiency, on the other hand, also results in cellular abnormalities and diseases. Impaired T cell development was observed in the absence of Vav1, and defective Vav1 expression was found in a subset of

patients with common variable immunodeficiency (Paccani, Boncristiano et al. 2005; Raberger, Boucheron et al. 2008). Loss of Vav2 was thought to be correlated with tachycardia and cardiovascular diseases in mice (Sauzeau, Jerkic et al. 2007), and Vav3 deficiency also leads to sympathetic hyperactivity and cardiovascular dysfunction (Sauzeau, Sevilla et al. 2006).

Taken together, Vav proteins play critical roles in cellular actin cytoskeleton signaling and have been implicated in a variety of severe human diseases. The autoinhibition and activation mechanisms of Vav are important to understand not only to extend our knowledge of Vav regulation, but also to cast light on the strategies of disease diagnosis and drug treatments.

The structural elements of Vav

The cDNA nucleotide sequence of vav oncogene predicts a protein product of 845 amino acids and 95 kilo Daltons. The domain structure of Vav proteins is composed of a leucine rich Calponin Homology (CH) domain, a 45 residue long Acidic (Ac) region containing mostly acidic (23 Glu or Asp) residues, a central Dbl homology (DH) domain, a Pleckstrin homology (PH) domain which is present in many kinases, GTPases and nucleotide exchange factors, followed by a cysteine rich Zinc Finger (ZF) domain similar to those found in protein kinase C family, and two Src Homolgy (SH) 3 domains flanking an SH2 domain (fig. 1-1).

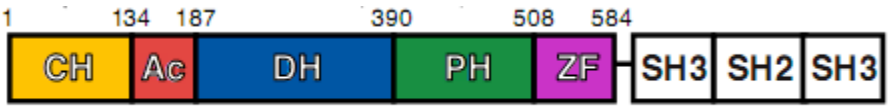


Fig. 1-1 Domain structure of Vav proteins. Domains of Vav1 are colored: CH (gold), Ac (red), DH (blue), PH (green), ZF (magenta), SH3 (white), SH2 (white).

The relevant domains are present in all Vav isoforms, except CelVav protein is lacking the most C terminal SH3 domain.

Interactions between different regions of Vav are responsible for the regulation of its biological activity. Vav was found to be autoinhibited in its resting state, and to be activated by an N-terminal sequence deletion or by Lck phosphorylation (Coppola, Bryant et al. 1991; Katzav, Cleveland et al. 1991; Gulbins, Coggeshall et al. 1993). In addition to post-translational modification, interactions with other proteins or lipids also appear to contribute to the regulation of Vav function. Vav CH domain has been shown to associate with a variety of signaling proteins, including Calmodulin (Zhou, Yin et al. 2007), Ly-GDI (guanine nucleotide inhibitor) (Groysman, Russek et al. 2000), Ezh2 (Nolz, Gomez et al. 2005) and APS (Yabana and Shibuya 2002). These interactions, direct or indirect, were suggested to be important for Vav mediated cell signaling.

In the Vav Acidic region, there are three conserved tyrosine residues (Y142, Y160, Y174) that are phosphorylated by the Syk and Lck kinases. Of these three tyrosines, Tyr174 plays the most important role in the negative regulation of vav activity and was also the most extensively studied (Bustelo and Barbacid 1992; Bustelo, Ledbetter et al. 1992; Aghazadeh, Lowry et al. 2000; Amarasinghe and Rosen 2005). Phosphorylation or phosphomimic mutations of Y174 results in great enhancement in Vav GEF activity *in vitro* and transforming activity *in vivo* (Aghazadeh, Lowry et al. 2000; Zugaza, Lopez-Lago et al. 2002).

Furthermore, mutation of Y174 to phenylalanine also potentiates Vav mediated cellular transformation (Lopez-Lago, Lee et al. 2000). In summary, the N terminus of Vav inhibits its GEF activity, and phosphorylation of Y174 relieves the negative regulation.

The DH and PH modules are invariably coupled in Dbl family Rho GEFs, implying a conserved function for PH domain. The PH domain is not required for GEF activity of these proteins, nevertheless the PH domain plays diverse roles in the GEF proteins. It can act as stabilizer of the DH/GTPase complex, or as localization signal to ensure the compartmentalized action of the GEFs in specific regions of cells (Hoffman and Cerione 2002). In the specific case of Vav, it has been shown that the PH domain is essential for its transforming activity (Zheng, Zangrilli et al. 1996), and can function as a phospholipid-dependent regulator of its GEF activity (Han, Luby-Phelps et al. 1998). Broek and colleages demonstrated that the PH domain of Vav binds to phosphatidylinostol 4,5-biphosphate (PIP_2) and phosphatidylinostol 3,4,5-triphosphate (PIP_3) like other PH domains (Han, Luby-Phelps et al. 1998; Das, Shu et al. 2000). The phosphatidylinostol 3-kinase (PI3K) substrate PIP_2 inhibits, whereas its product PIP_3 activates the phosphorylation and activation of Vav. Moreover, PI3K activation was suggested to promote Vav membrane association through binding to PIP_3 by PH domain, and a functional PH domain is critical for Vav transforming activity (Palmby, Abe et al. 2002). Structural alignement of Vav

DH-PH-ZF module with PDK1 PH domain suggests binding to lipid and membrane associated Rac of Vav can happen simultaneously (Rapley, Tybulewicz et al. 2008).

The ZF domain is present in a variety of signaling proteins. The Vav1 Zinc Finger domain belongs to the C1 domain family of proteins with two Zn^{2+} ions coordinated. It was named a "Zinc Finger" domain because of its resemblance to the DNA binding regions of transcription factors (Colon-Gonzalez and Kazanietz 2006). However, C1 domains do not share structural homolgy with transcription factors and do not bind DNA. In general, C1 domains can be divided into two categories: the "typical" C1 domain and the "atypical" C1 domain. The typical C1 domain is capable of binding to lipids such as diacylglycerol (DAG) and phorbol esters, and is thought to be important for regulating the association between the protein and the cell membrane. The atypical C1 domain, on the contrary, does not bind to lipids, but can be involved in protein protein interactions. The ZF domain of Vav is classified as an atypical C1 domain based on its sequence. Mutations in the Vav1 ZF domain have been found to lead to decreased GEF activity (Booden, Campbell et al. 2002). Furthermore, it has been shown that the ZF domain is required for the transforming activity of Vav1 (Booden, Campbell et al. 2002). In order to explain this observation, it was proposed that the ZF domain in Vav1 might make direct contacts with Rac1 substrate, based on Nuclear Magnetic Resonance (NMR) chemical shift mapping

and Glutathione S-transferase (GST) pull down experiments (Movilla and Bustelo 1999; Heo, Thapar et al. 2005).

A recent crystal structure of Vav1 (residues 170-575) and Rac1 complex was determined by two groups (Rittinger and Kuhn),(Chrencik, Brooun et al. 2008; Rapley, Tybulewicz et al. 2008). It revealed that Vav1 PH domain and ZF domain form a structural unit and pack against the C terminal helix of the DH domain (fig. 1-2). The structural unit, when mutated, has dramatic decrease in terms of protein stability and Rac1 activation, suggesting it might be crucial for the structure integrity of Vav DPZ module. It was proposed that this organization stabilizes the conformation of the DH domain and contributes to efficient GEF activity. The ZF domain is far away from the Rac1 substrate, making no direct interactions between them (Chrencik, Brooun et al. 2008; Rapley, Tybulewicz et al. 2008). The requirement of the ZF domain for potent GEF activity and transforming activity can be explained by the stablization of α6 conformation in DH domain by PH and ZF domain, which is necessary for optimum interactions between Vav1 and Rac1.

The role of the C terminal SH3-SH2-SH3 module in Vav GEF regulation is unclear yet, however it has been suggested that it can mediate cellular functions independently of Vav exchange activities by working as adaptor protein (Braiman, Barda-Saad et al. 2006). Bustelo and colleagues found that a Vav protein lacking the SH3-SH2-SH3 module and targeted to the cell membrane is incapable of

stimulating NF-AT signaling pathway despite the intact CH, DH, PH, ZF regions (Bustelo 2001). It has also been demonstrated that an intact SH3 domain is required for cellular transformation (Groysman, Nagano et al. 1998).

In pervanadate treated Jurkat cells, it was found that Vav could be pulled down by a variety of GST-SH2 proteins, such as Crk, Nck, GAP and PLCr, which all bind to phosphotyrosine containing peptides with different specificities, suggesting Vav may be phosphorylated on multiple tyrosine sites (Ramos-Morales, Druker et al. 1994). Based on analysis of the consensus amino acid sequence of tyrosine phophorylation sites: ΦYXXΦ (Φ=I, L, or V and X=any amino acid) (Songyang, Shoelson et al. 1993; Songyang and Cantley 2004), four tyrosine residues were identified later in the Vav C-terminus that potentially can be phosphorylated, including Y745, Y826, Y841 and Y844 (Romero and Fischer 1996). A recent study regarding the importance of these tyrosine residues demonstrated that single site mutations of Y826 and Y841 to phenylalanine lead to great loss of Vav transforming activity, subsequent Rac activation and GEF activity, implying that the C-terminus also plays a role in regulating Vav function (Lazer, Pe'er et al.). However, the mechanism of this regulation remains to be explored in the future.

Mechanism of Vav autoinhibition by Ac region

A Vav contruct lacking the first 186 amino acids was found to be up to 100-fold more potent in focus transformation assays when compared with the wild type Vav molecule, indicating the N-terminus of Vav is autoinhibitory (Abe, Whitehead et al. 1999). The molecular mechanism of this autoinhibition was not well understood until a solution structure of mouse Vav1 was solved through NMR spectroscopy, which contains the DH domain and part of the Ac region (residues 170-375, termed AD hereafter) (Aghazadeh, Lowry et al. 2000). The structure shows that the Vav1 DH domain (residues 196-375) forms an α helical bundle with 11 α helices and is structurally similar to other DH domains. The N-terminal extension containing residues 170 to 177 forms an additional amphipathic helix and binds to the active site of Vav DH domain which interacts with the Rac GTPase (Fig. 1-3). This causes steric hindrance for interactions with Rac and therefore results in inhibition of Vav GEF activity. Tyr174 in the Ac region, the key regulatory site for Vav DH domain, is buried in the binding interface, which is mainly hydrophobic, and makes contacts with the side chains of residues in the Rac1 binding region: Y209, T212, P320, L325 and V328. Further chemical shift analysis combined with biochemical studies indicated that phosphorylation of Tyr174 results in melting of the inhibitory helix, which disrupts the binding between the Ac region and the DH domain, exposes the GTPase binding site and releases autoinhibition (Aghazadeh, Lowry et al. 2000).

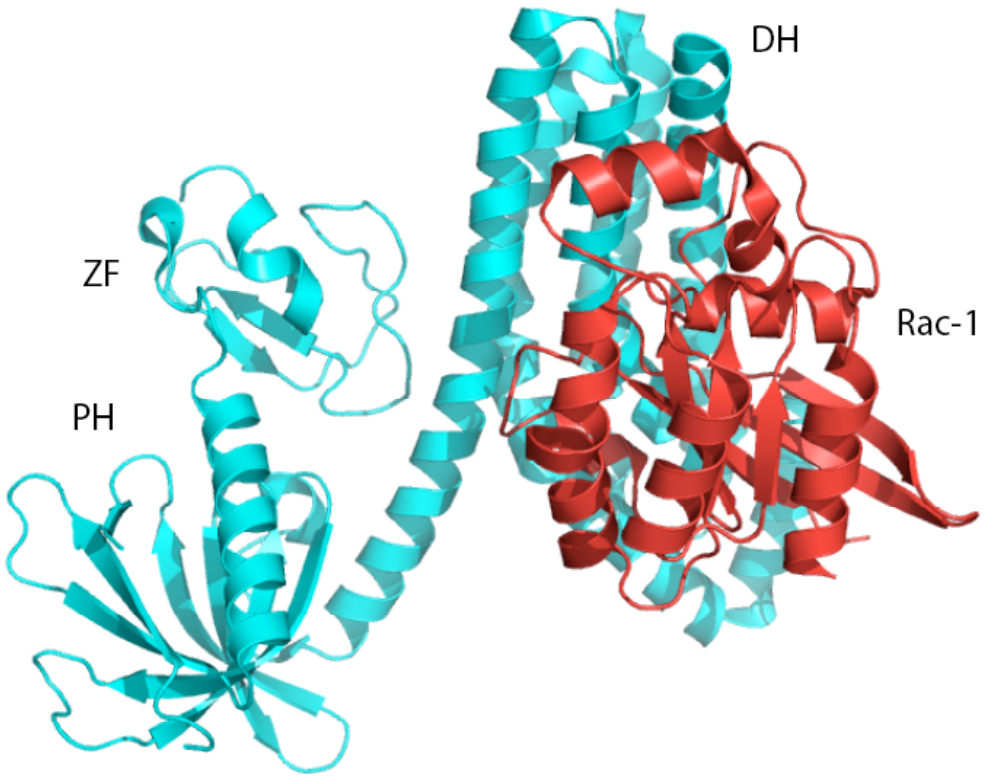


Fig. 1-2 Structure of DPZ-Rac complex. Vav DPZ is displayed in cyan, Rac GTPase is displayed in red.

Internal dynamics control activation and activity of the autoinhibited Vav DH domain

In the solution structure of the AD protein, Tyr174 is completely buried in the helix-DH interface (Aghazadeh, Lowry et al. 2000). However, biochemical studies revealed that the rate of Tyr174 phosphorylation by Lck kinase was still appreciable (Amarasinghe and Rosen 2005). This phenomenon can be envisioned in two scenarios: first, the AD protein may fluctuate between an inactive state and an active state, with a strong preference toward the inactive state. However there is still a significant population of Vav in the active state, which could be recognized and activated by Lck kinases. Second, AD protein only exists in an inactive conformation, whereas kinase binding to Vav might initiate a structural change of AD, and it results in a conformation that can be accessed by kinase. In order to distinguish these two scenarios, Li et al. performed NMR dynamic and biochemical studies of the Vav AD protein. This work showed that the AD module actually exists in an equilibrium between a ground state, where the inhibitory helix is bound to the DH domain, and a weakly populated excited state, where the helix is released and unfolded (Li, Martins et al. 2008, Fig. 1-4). The closed state is favored by a 10:1 ratio compared with the open state. The rate of Tyr174 phosphorylation by Lck was found to be linearly related to the population of the excited state, reflecting that the accessibility of Y174 is directly related to the population of AD in inhibitory helix dissociated state. Furthermore, the

catalytic GEF activity of AD also is linearly related to the population of the excited state. All together, these studies established a deeper understanding of Vav1 autoinhibition and activation mechanisms, and established that the internal dynamics of Vav1 control both the basal activity and the activation rate of the autoinhibited DH domain.

Cooperative suppression of Vav1 transforming activity through the CH domain

In addition to the basal autoinhibition of the Vav DH domain exerted by the acidic α helix, there is another level of negative regulation of Vav biological activity from the CH domain. It has been shown that a Vav construct with the N-terminal 65 amino acids truncated (giving VavΔ65, in which the N-terminal half of CH domain is missing) has increased Rac activation and transforming activity, implying this sequence take part in a negative regulation of Vav (Han, Luby-Phelps et al. 1998; Lopez-Lago, Lee et al. 2000; Bustelo 2001; Llorca, Arias-Palomo et al. 2005). Vav construct lacking the N-terminal 186 amino acids (giving VavΔ186, in which both CH domain and the whole Ac region are missing), gains more potent transforming activity (Crespo, Schuebel et al. 1997; Abe, Whitehead et al. 1999; Bustelo 2001). Since in this construct the Ac region is absent, the transforming activity is independent of phophorylation (Katzav, Cleveland et al. 1991; Llorca, Arias-Palomo et al. 2005).

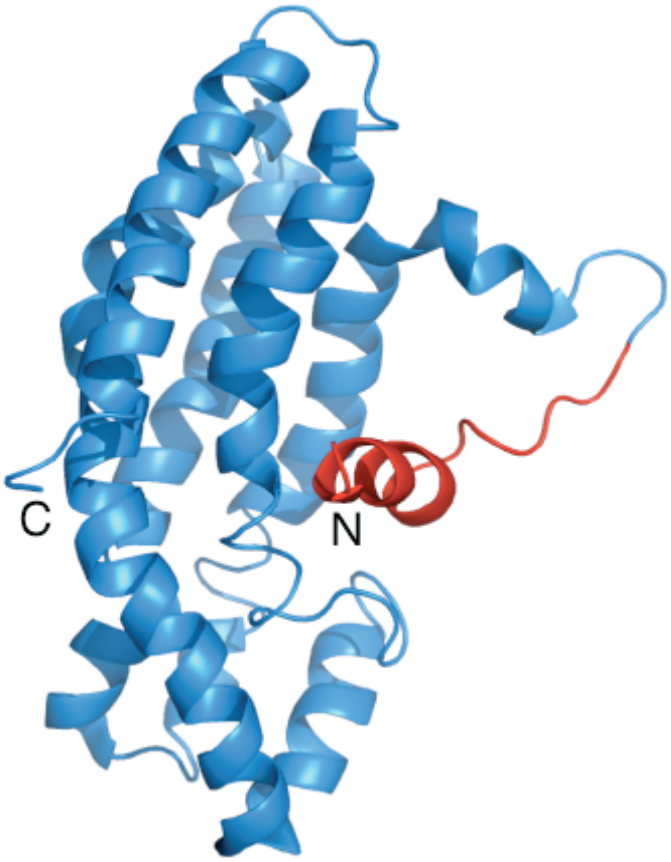


Fig. 1-3 Solution structure of Vav1 AD module (pdb ID: 1F5X). The N terminus and C terminus of the protein are denoted as "N" and "C". The Acidic region is colored in red and the DH domain is colored in light blue. Figure prepared by pymol (Delano 2002).

In VavΔ65, the inhibitory module (Ac and DH domain) is intact. The observation that VavΔ65 is more potent than wild type Vav in transforming assays suggest CH domain suppresses DH domain activity in synergy with the Ac region. The mechanism of this cooperative suppression has been unknown. To understand the mechanism structurally, Llorca et al. presented an electron microscopy (EM) reconstruction of autoinhibited Vav3 at a medium resolution of 22 Å (Llorca, Arias-Palomo et al. 2005). The reconstruction suggested an elongated structure composed of a massive head and a triangular tail (fig. 1-5). The authors then docked each of the eight domains of Vav into the reconstructed volume by using fitting tools in SITUS. The docking suggested that CH domain $\alpha5 - \alpha6$ loop contacts the ZF domain. The active site of the DH domain is buried under a bulky mass which is composed of the CH domain and Ac region, making the protein incapable of binding to Rac (Heo, Thapar et al. 2005).

Based on the result of EM reconstruction, CH domain may directly block Rac1 binding site together with Ac region. Therefore it provided a plausible mechanism for the CH domain cooperative autoinhibition. However due to the low resolution of the structure, the detailed orientation and residue contacts between different domains are not certain and therefore need further validation from higher resolution structural information and functional analysis.

It was realized later that the relative orientations of DH and ZF domains in the EM structure are clearly different from that in the DPZ/Rac complex structure

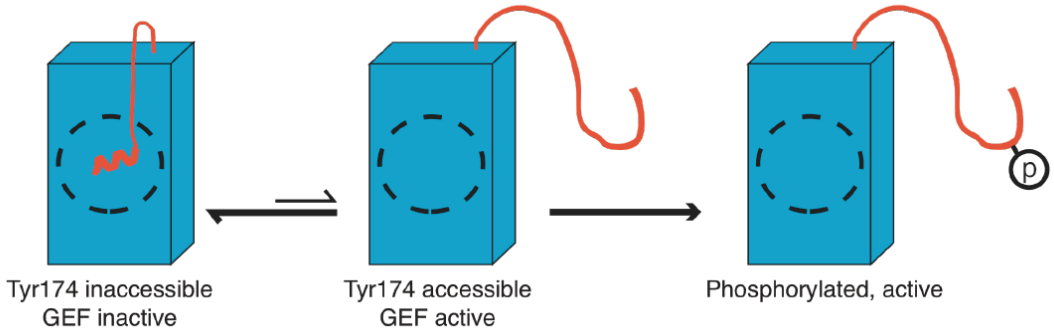


Fig. 1-4 Model for the regulation and activation mechanism of Vav AD module

(Chrencik, Brooun et al. 2008; Rapley, Tybulewicz et al. 2008). In order to clarify the discrepancies, Kuhn and colleagues performed Small Angle X-ray Scattering (SAXS) experiments to confirm the conformation of the DPZ/Rac complex in solution. The SAXS analysis indicated that the crystal structure of the protein complex fits the data much better than the elongated configuration of Vav3 proposed in the EM reconstruction. In addition, mutational analyses of the Vav1/Rac1 switch 1 and switch 2 interfaces performed by Kuhn and colleagues confirmed the binding sites seen in the crystal structure between DH and Rac1 (Chrencik, Brooun et al. 2008). These studies together suggest that the EM structure remains questionable and a high resolution structure of Vav1 is needed to elucidate the mechanism of its cooperative autoinhibition.

The Vav1 AD module, as described earlier, exists in an equilibrium between a helix-bound ground state and a helix-released excited state. In the context of full length Vav1 protein, this module appears to be retained, as suggested by the studies performed by Li et. al. (fig. 1-6). This linearity of chemical shifts sensing the DH-Ac equilibrium from the AD and CADPZ proteins indicates that CADPZ also exists in an equilibrium analogous to that in AD, fluctuating between helix-bound and helix-released states. Importantly, interactions between the AD module and other regions of the full length Vav1 protein seem to modulate the DH-Ac equilibrium toward the closed state. In full

Fig. 1-5 Structure of Vav3 after EM reconstruction and docking. Position of each domain is denoted, adapted from (Llorca, Arias-Palomo et al. 2005)

length Vav1, the observed chemical shifts of AD indicate the equilibrium is shifted toward the closed conformation by ~ 10-fold, resulting a population of~1% in the active state. These NMR data suggest a thermodynamic mechanism for the cooperative suppression of Vav1. However, how this energetic regulation is achieved remains unclear. A more complete understanding of Vav1 cooperative suppression will require the discovery of how the structure and thermodynamics of AD module are regulated by other domains of the protein.

Cooperative autoinhibition in other multi-domain signaling proteins

The construction of cooperative suppression is not unique to Vav proteins. The Tim protein, for example, is another Dbl family GEF regulated by autoinhibition (Yohe, Rossman et al. 2007; Yohe, Rossman et al. 2008). Analogous to Vav, it contains an N-terminal inhibitory helix whose truncation, phosphorylation or mutation will lead to activation of GEF activity. In addition to this basal inhibition, Tim contains a C-terminal SH3 domain, which cooperatively inhibits DH domain activity with the helical motif. It has been proposed that this SH3 domain binds to an internal poly-proline region between the helical motif and the DH domain, modulating the direct inhibitory interactions and further suppressing DH activity (Yohe, Rossman et al. 2008). Although the physical basis for autoinhibition by SH3 domain remains to be understood, it may well be analogous to the inhibitory effect by CH domain in Vav protein.

Src family kinases also exhibit hierarchical regulation. Autoinhibition of the Src kinase domain arises through two distinct intramolecular interactions: 1. contacts between SH2 domain and a C terminal phospho-tyrosine motif and 2. contacts between SH3 domain and a poly-proline motif (Xu, Harrison et al. 1997; Xu, Doshi et al. 1999). Although quantitative analyses of the inhibition from the two interacting equilibria are not available yet, dynamic simulation analyses suggest that both SH2 and SH3 domains suppress Src kinase activity through shifting the kinase domain to an inactive conformation (Young, Gonfloni et al. 2001; Faraldo-Gomez and Roux 2007). Destabilization of either of these two interactions, either by truncation or by allosteric activation, results in increased in kinase activity (Briggs and Smithgall 1999); whereas a peptide fragment containing the ligands for both SH2 and SH3 domains strongly activates Src activity (Alexandropoulos and Baltimore 1996). Similarly, these two inhibitory interactions can be considered analogous to the contacts made by Ac region and CH in Vav protein, destabilization of the two interactions can be considered as disrupting the DH-Ac interaction and truncation of CH in Vav, which have both been shown to increase Vav transforming activity. Understanding the physical basis of Vav cooperative suppression will therefore help us to learn how the modulatory interactions collectively generate strong inhibition in these proteins.

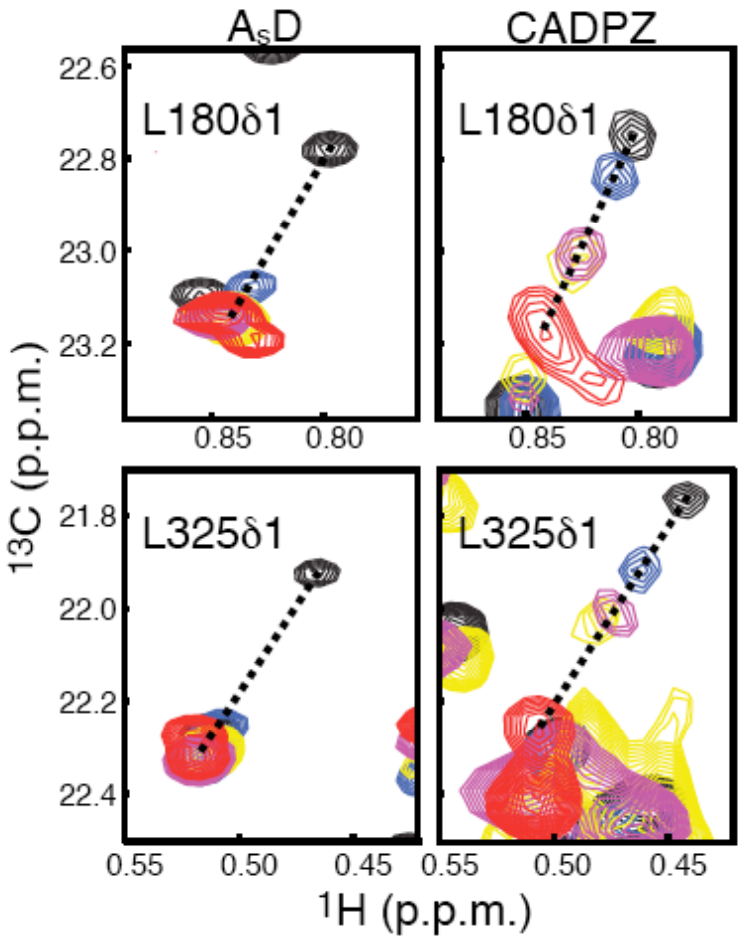


Fig. 1-6 Overlaid $^{1}H/^{13}C$ methyl TROSY spectra of murine A_sD and CADPZ proteins (left and right panels, respectively), showing Leu180δ1 and Leu325δ1 signals (top and bottom panels, respectively). Colors are as follows: wild-type (WT) (black), K208A (blue), K208S (yellow), K208G (magenta) mutants and pAsD or CADPZY174D (red). Dotted lines connect resonances from phosphorylated and non-phosphorylated proteins.

Activation of Vav through tyrosine phosphorylation

Protein autoinhibition is generally relieved in two ways: either through covalent modification or through binding to an allosteric activator. Vav activation can occur through the phosphorylation of three conserved tyrosine residues in the Ac region (Bustelo and Barbacid 1992; Crespo, Schuebel et al. 1997; Han, Das et al. 1997; Lopez-Lago, Lee et al. 2000): Y142, Y160 and Y174. In particular, Y174 was found to be a key regulatory site. Phosphorylation or mutation of Y174 leads to in vitro activity enhancement and cellular transformation in vivo (Bustelo and Barbacid 1992; Crespo, Schuebel et al. 1997; Han, Das et al. 1997; Mosteller, Han et al. 2000). In addition to Y174, the sites for Y142 and Y160 also contain the consensus amino acid sequence for Src family kinases: ΦYXXΦ (Φ=I, L, or V and X=any amino acid) (Songyang, Shoelson et al. 1993; Songyang and Cantley 2004). Upon receptor stimulation, these two tyrosines also get phosphorylated (Lopez-Lago, Lee et al. 2000; Tamas, Solti et al. 2003; Miletic, Graham et al. 2007). Single site mutation of Y142 or Y160 to F has been shown to result in a small increase in Vav transforming activity (Lopez-Lago, Lee et al. 2000), implying these two tyrosines also play a role in the regulation or activation of Vav function. The underlying mechanism by which Y142 and Y160 regulate Vav activity is not well understood.

Autoinhibited systems usually exhibit rapid activation kinetics (Huse and Kuriyan 2002), however the kinetic and thermodynamic basis of the activation process are not well understood. The kinetic pathway for Vav activation is therefore of particular interest. A rigorous quantitative understanding of Vav activation will elucidate the basic kinetic and thermodynamic principles involved in relief of autoinhibition in other systems. Amarasinghe et al. demonstrated through biochemical and structural analyses of the Vav AD module that Y174 but not Y142 or Y160 in the Ac region is protected from phosphorylation by the Src-family kinas, Lck (Amarasinghe and Rosen 2005). In Vav AD module, Y142 or Y160 does not make significant contacts with the DH domain and is likely to be free in solution. Phosphorylation of Y142 leads to a four-fold increase of k_{cat}/K_M in phosphorylation of Y174, likely because phosphorylated Y142 creates a docking site for the Lck SH2 domain and promotes efficient phosphorylation reaction. This study suggests there might be a stepwise phosphorylation event during Vav activation in which access point for kinase are generated to accelerate the phosphorylation reaction. In full length Vav, however, whether Y142 and Y160 are buried or solvent exposed is unclear. Addressing this question will require further structural and kinetic studies. Through my thesis work on Vav, which is described in the following chapter, we now understand that Y142 and

Y160 are not solvent exposed. Dynamic properties of the Ac region may account for the exposure of Y142 and Y160 to Src family kinases.

Chapter 2

Structural and Energetic Mechanisms of Cooperative Autoinhibition and Activation of Vav1[1]

As a guanine nucleotide exchange factor that plays a critical role in actin regulation pathways, Vav is strongly autoinhibited in the resting state. Although attempts have been made to uncover the mechanism of cooperative suppression of Vav1 by CH and Ac domains through EM reconstruction (Llorca, Arias-Palomo et al. 2005), the resulting conclusions have not been validated through atomic resolution structural analysis. Moreover, whether and how the dynamic process of Ac helix binding/release in the AD module is modulated in the context of full length protein are unknown. In fact, due to technical challenges, quantitative measurements of protein dynamics in a multi-domain system have never been achieved. The kinetic pathway of Vav tyrosine phosphorylation and activation is another important question for us to address. A deeper understanding of Vav

[1] This chapter is mostly adapted from Yu et al., *Cell*, 2010, Jan 22;140(2):246-56.

regulation may reveal a general mechanism that is adopted by other multi-domain proteins. To address these issues, during the first half of my thesis work, I concentrated on obtaining crystals of human Vav1 that diffracted to high resolution and on performing biochemical and NMR dynamic studies on Vav mutant proteins. The structure of human Vav1 was determined from these crystals by our colleagues Diana Tomchick and Mischa Machius in the UT Southwestern Structural Biology Laboratory. Another student in the Rosen lab, Ilidio Martins, together with Pilong Li, performed NMR studies on mouse Vav1 constructs (which complemented my work on the human protein), and conducted kinetic studies to understand how the strong autoinhibition of Vav1 can be rapidly released by phosphorylation. Martin Fernandez-Zapico[2] and Daniel D. Billadeau[2] performed the work to test the activities of Vav mutant proteins *in vivo*.

Results

Structure of Autoinhibited Vav1

To understand how interdomain contacts act together to suppress GEF activity, I sought to detemine the structure of the CADPZ element of Vav1 at

---

[2] Martin Fernandez-Zapico and Daniel D. Billadeau are at Department of Immunology and Division of Oncology Research, Schulze Center for Novel Therapeutics, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

atomic resolution. I started with mouse Vav1 (residues 1-584, 1-590 and 1-584 with all cysteines mutated to alanine in the CH domain) constructs and tried to crystallize the corresponding purified proteins. After months of trials, I obtained needle-like crystals of construct 1-584 in 20 mM Tris (pH 7.5), 2 M ammonium sulfate . The crystal grew to a size of ~50 x 10 x 5 μm after seeding. However, these crystals only diffracted to 8 Å at the Advanced Photon Source 19ID beamline. I then moved to test the human Vav1 (1-584) construct termed as CADPZ from now on, which shares 93.3% amino acid sequence identity with mouse Vav1. The human protein could crystallize in test tubes at a concentration of higher than 10 mg/ml after protein purification. Low temperature (at or below 4°C) and low concentration of salt (at or below 50 mM) were found later to be the critical factors for crystallization. After optimization the crystals diffracted to ~3.5 Å in house, but with high mosaicity. I suspected that high mosaicity might be because the crystals were partially dissolved in the cryosolution that contained 50 mM sodium chloride. The high mosaicity problem was gone when the cryo-solution contained no salt. After further optimization, the crystals grew to a size of ~50 x 100 x 200 μm (Fig. 2-1) and diffracted to 2.73 Å at the Advanced Photon Source 19ID beamline.

Initially, we tried to determine the phases by molecular replacement using the Vav AD structure (residue number: 187-375, pdb ID: 1F5X) as the search

model. But this strategy failed. I then grew selenium methionine derivative crystals under the same conditions by seeding with native crystals. These derivative crystals enabled phases to be determined using the single-wavelength anomalous dispersion method. The crystal structure of the CADPZ element of human Vav1 (Table 2-1, Figs. 2-2) was solved by Diana Tomchick and Mischa Machius using diffraction data to a resolution of 2.73 Å. The structure was deposited into the protein data bank with ID code: 3KY9. One feature of CADPZ structure is that the AD module, consisting of the Y174-containing inhibitory helix and DH domain, closely resembles the previously reported solution structure of this fragment (Aghazadeh, Lowry et al. 2000, Fig. 1-3). The inhibitory helix lies perpendicular to the long axis of the DH domain in the bottom of the active site. The PH and ZF domains dock onto the C-terminus and the center of an extended C-terminal helix of the DH domain, respectively, and make extensive contacts with each other as well. The organization of the DPZ portion of the structure closely resembles that seen in recently reported structures of the DPZ-Rac complex (Chrencik, Brooun et al. 2008; Rapley, Tybulewicz et al. 2008). However, a small bend near the center of the C-terminal DH helix somewhat changes the relative orientations of the DH domain and PZ unit in the two structures (Fig. 2-3). The CH domain binds the PH domain and the extended C-terminal helix of the DH domain, but does not contact the ZF domain, in contrast to previous inferences from EM and biochemical data (Llorca, Arias-Palomo et al.

2005). The CH domain contacts the helix-DH element only through R58 (CH), whose sidechain is near that of E169 (helix).

The Ac element emerges from the side of the CH domain opposite the DH domain. A short N-terminal helix (residues 143-150) wraps around the CH domain and also contacts the PH domain. This is followed by a long extended element (residues 151-166) that connects to the inhibitory helix packed in the DH active site (Figs. 2-2). Y142 is located immediately before the N-terminal Ac helix; Y160 is located in the extended element. Both sidechains lie on the surface of the CH domain, with their aromatic rings contacting primarily hydrophobic residues (Y142: V30, G106, I109, L145; Y160: V41, L159, R63) (Fig. 2-4 A). The hydroxyl group of Y160 forms a hydrogen bond with the sidechain of D39 (Fig. 2-4 B).

Validation of the Structure by Mutagenesis

The overall domain organization of the structure here is quite different from the previously reported EM reconstruction of Vav3, which is 58 % identical to Vav1 in amino acid sequence. Most significantly, the contacts of the CH domain to the PH domain and extended C-terminal helix of the DH domain are not possible based on the previous description. To validate the domain organization

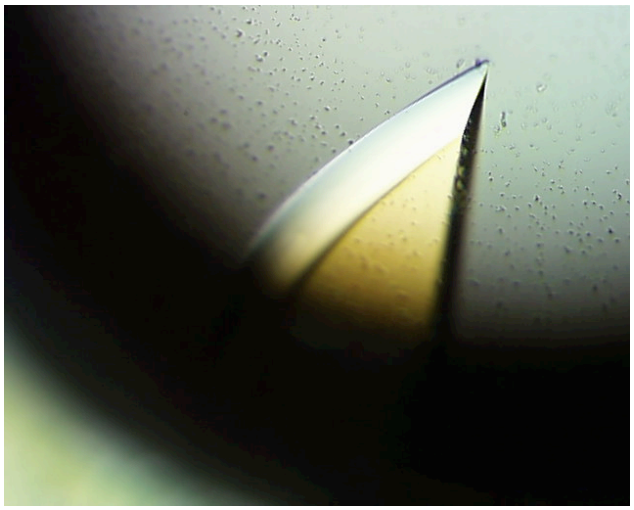


Fig. 2-1 Crystal of Vav1 CADPZ (~50 x 100 x 200 μm, obtained in 20 mM Tris (pH 7.5), 50 mM NaCl, 5% (w/v) glycerol, 2 mM TCEP)

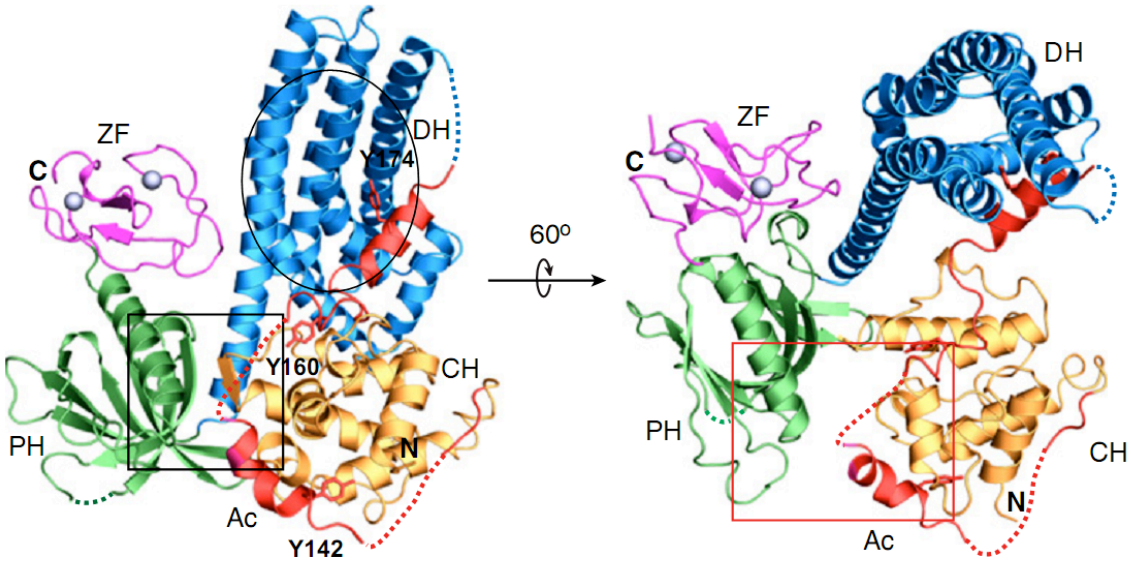


Fig. 2-2 Ribbon diagram of Vav1 CADPZ. Positions of each domain as well as the N terminus and C terminus is denoted. Domains are colored as fig. 1-1. Dashed lines indicate regions not observed in the electron density map. Sidechains of tyrosines 142, 160 and 174 are shown as sticks. DH active site is circled. Boxed areas in left and right panels are expanded in Figures 2-6A and 2-6B, respectively.

Data collection		
Crystal	**Native**	**SeMet[a] *peak***
Energy (eV)	**9,669.3**	**12,667.7**
Resolution range (Å)	**47.30-2.73 [2.94-2.88] (2.78-2.73)**	**42.67-2.80 (2.83-2.80)**
Unique reflections	**42,421 [2,128] (2,089)**	**75,280 (2,154)**
Multiplicity	**14.3 [14.6] (12.8)**	**9.7 (7.6)**
Data completeness (%)	**100.0 [100.0] (99.9)**	**98.9 (87.6)**
R_{merge} (%)[b]	**7.6 [59.8] (95.9)**	**9.3 (77.5)**
R_{meas} (%)[c]	**7.8 (84.2)**	
I/σ(I)	**44.8 [5.2] (2.0)**	**51.9 (2.2)**
Wilson B-value (Å^2)	**95.2**	**98.8**
Phase determination		
Anomalous scatterer	**selenium, 30 out of 36 possible sites**	
Figure of merit (42.7-2.8 Å)	**0.27**	
Refinement statistics		
Resolution range (Å)	**47.29 -2.73 (2.81-2.73)**	
No. of reflections R_{work}/R_{free}	**39,835/1,683 (2,439/100)**	
Data completeness (%)	**99.5 (93.4)**	
Atoms (non-H protein)	**8,831**	
R_{work} (%)	**23.1 (30.0)**	
R_{free} (%)	**27.1 (35.1)**	
R.m.s.d. bond length (Å)	**0.007**	
R.m.s.d. bond angle (°)	**1.12**	
Mean B-value (Å^2) (protein)	**112.9**	
Ramachandran plot (%) (favored/additional/disallowed)[d]	**94.6/5.4/0.0**	
Missing residues	**Chain A: 1, 130-141, 151-155, 181-188, 479-481, 565-584; Chain B: 130-137, 151-155, 181-188, 566-584.**	

Table 2-1. Data collection and structure refinement statistics of Vav1. Data for the outermost shell are given in parentheses. Data for the 2.94-2.88 Å shell are given in brackets. Although data between 2.88 and 2.73 Å resolution had high values of Rmerge, inclusion of these reflections improved the quality of the electron density map, likely

because of the very high redundancy of the dataset. Thus, they were included in structure calculations.

[a]Bijvoet-pairs were kept separate for data processing.

[b]$R_{merge} = 100\ \Sigma_h\Sigma_i|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{h,i}$, where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.

[c]$R_{meas} = 100\ \Sigma_h[n_h/(n_h-1)]^{1/2}\Sigma_i|\hat{I}_h - \langle I_{h,i}\rangle|/\Sigma_h\Sigma_i\langle I_{h,i}\rangle$, where $\hat{I}_h = 1/n_h\Sigma_i I_{h,i}$ and the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection. These values were calculated using the program *SCALA* (Evans 2006)

[d]As defined by the validation suite MolProbity (Davis, Leaver-Fay et al. 2007).

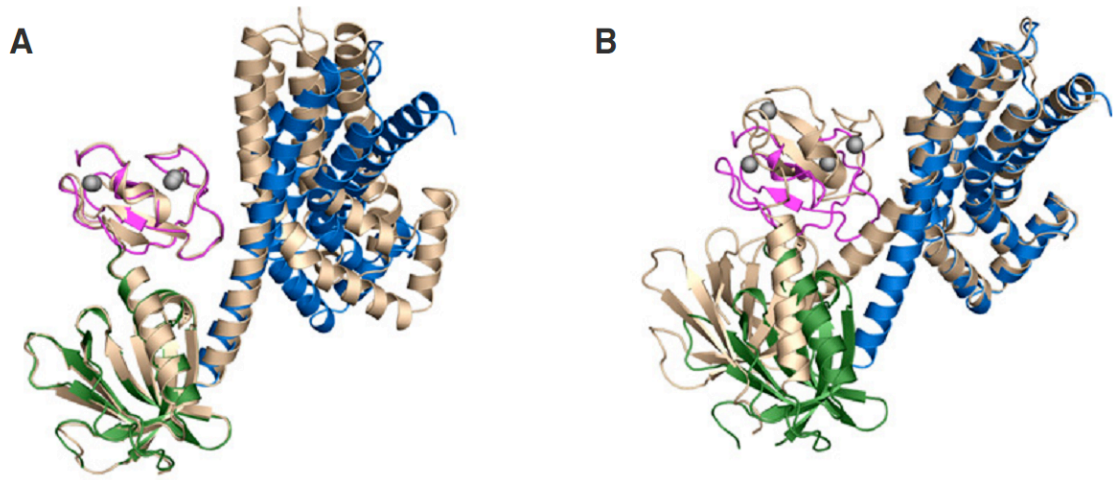


Fig. 2-3 Overlay of DPZ elements from the CADPZ and DPZ/Rac complex. (2vrw) structures. The two structures are related by a bend in the center of the C-terminal helix of the DH domain. DPZ element from the DPZ/Rac complex is colored tan. (A) Alignment based on the PH and ZF domains (pairwise backbone root-mean-squared-deviation (r. m. s. d.) for residues 391-417, 419-456, 465-564 of PZ = 0.91 Å). (B) Alignment based on DH domains (backbone r. m. s. d. for residues 190-368 = 0.88 Å).

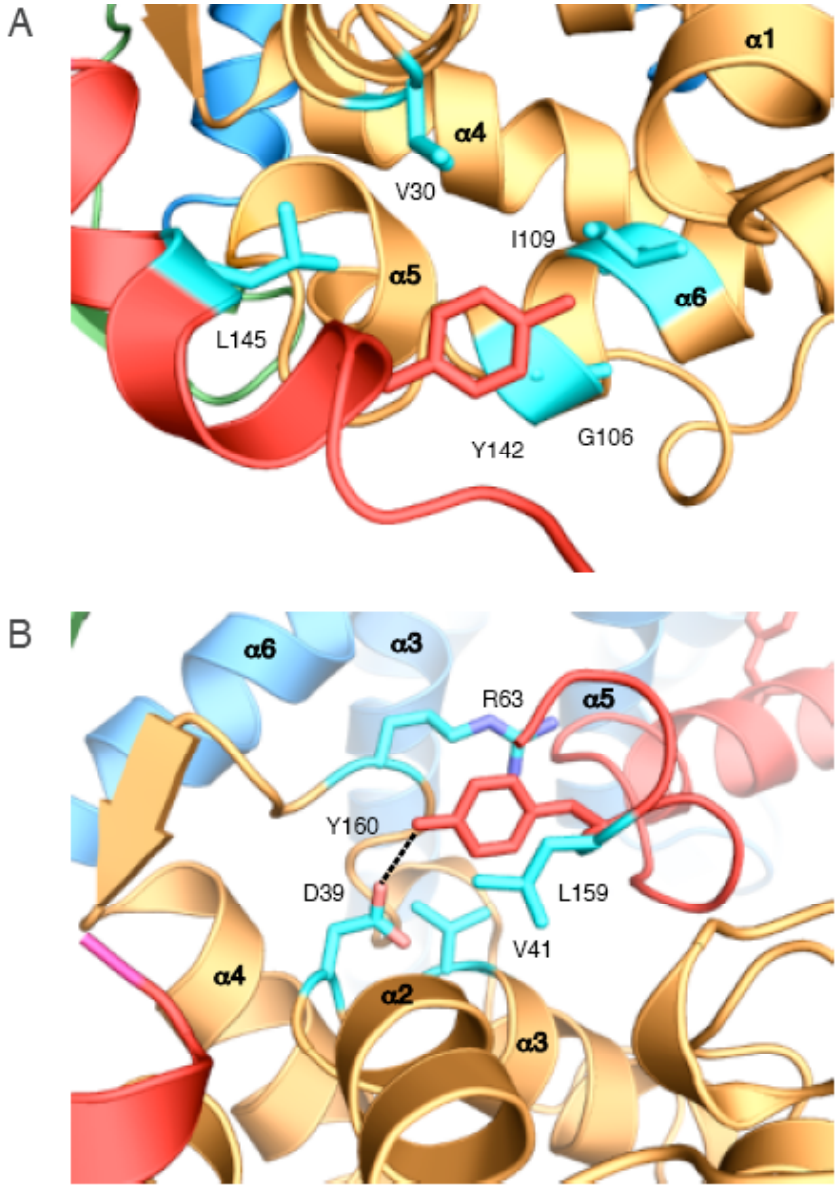


Fig. 2-4 Contacts of Y142 (A) and Y160 (B) to other residues in the CADPZ crystal structure. Y142 and Y160 are red and interacting residues are cyan. Dashed line represents a hydrogen bond.

observed in the crystal, I mutated a series of residues at different contact surfaces of the CH domain, and compared the GEF activity of these proteins to several benchmarks. In the wild type CADPZ protein, GEF activity is inhibited by the effects of the Acidic helix on the DH active site and by the modulation of this core interaction by the interdomain contacts of the CH domain (referred to as modulatory interactions hereafter). Thus, as shown in Fig. 2-5A, CADPZ has extremely low GEF activity toward Rac. Truncation of the CH domain (giving the ADPZ protein) results in partial activation, consistent with the absence of modulatory interactions of the CH domain, but the remaining presence of the core inhibitory interaction of the Acidic helix with the DH active site. GEF activity of fully active (mimicked by $ADPZ_{Y174D,}$ in which both modulatory and core repression are gone), partially active (ADPZ) and maximally inhibited (CADPZ) proteins parallels affinity for Rac, as qualitatively assessed in pull-down assays (Fig. 2-5B), further supporting the idea that activity reflects accessibility of the DH active site (Li, Martins et al. 2008)

To probe the conformation of phosphorylated CADPZ, I acquired $^{1}H/^{13}C$ methyl TROSY (HMQC) spectra (Tugarinov, Hwang et al. 2003) of the isolated CH domain, pCADPZ, and CADPZ proteins. HMQC experiments can be as much as 3-fold more sensitive than the traditional HSQC experiments, and are especially useful for large proteins (Tugarinov, Hwang et al. 2003). Proteins were

also deuterated except for methyl groups to gain better sensitivity and resolution through the methyl TROSY effect. For large sized proteins (above 23kD), a 25-50 fold gain in sensitivity can be gained when using TROSY-type experiment compared with the conventional experiments (Fernandez and Wider 2003). In overlaid spectra, all methyl resonances of CH have corresponding matching resonances with those of phosphorylated CADPZ (giving pCADPZ, with all three tyrosines in the Acidic region phosphorylated), but do not match with those of unphosphorylated CADPZ (Fig. 2-5C). This result indicated that phosphorylation of CADPZ disables both the modulatory and core elements of the autoinhibitory apparatus, leading to the maximal GEF activity shown in Fig. 2-5A.

From the studies performed above, we realized the activities of wild type CADPZ (fully inhibited by core plus modulatory interactions), ADPZ (only core inhibition) and pCADPZ (fully active) provide reference points for analysis of a series of CADPZ mutants. We designed the mutations based on the structure of the CH-Ac interface with the DPZ element (Fig. 2-6). The CH-Ac element contacts DPZ through two groups of residues. In the first group, the α3-α4 loop of the CH domain (residues 66-68) makes parallel β-sheet-like interactions with the first β-strand of the PH domain (residues 406-408); inter-strand hydrogen bonds are observed between the sidechain of D406 and mainchain NH of F69, and S67 is immediately adjacent to D406 (Fig. 2-6). Additional contacts in this region

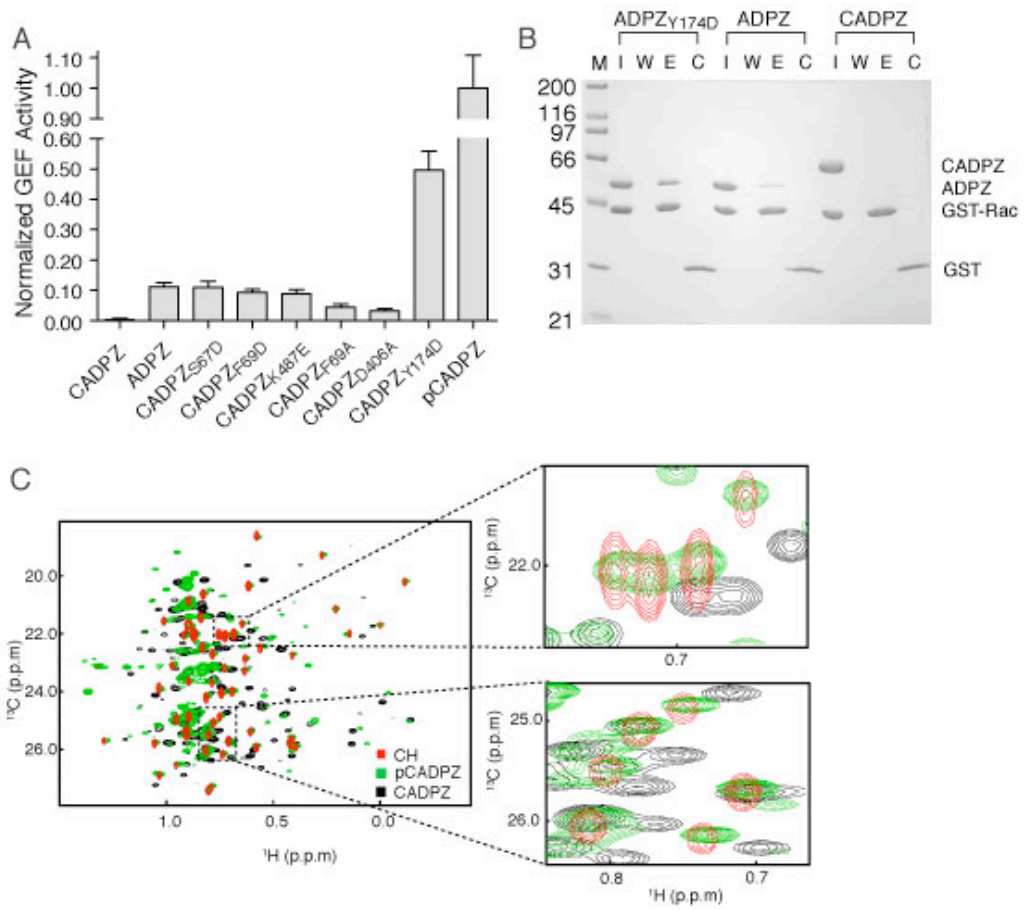


Fig. 2-5 Modulatory contacts contribute to autoinhibition of Vav activity. (A) Normalized GEF activity of human Vav1 proteins. Error bars show standard deviation from three independent measurements. (B) Immobilized GST-Rac(GDP) was used to pull-down the indicated Vav1 proteins, which were separated by SDS-PAGE and visualized by Coomassie Blue staining. M, I, W, E, C represent molecular weight markers, input, final wash, elution and elution from control (GST) beads, respectively. (C) Overlaid $^{1}H/^{13}C$ methyl TROSY spectra (Tugarinov, Hwang et al. 2003) of the isolated CH domain, pCADPZ, and CADPZ, colored red, green and black, respectively.

CH signals superimpose on those of pCADPZ, and are often different from those of CADPZ. These data suggest that when all three tyrosine sites are phosphorylated in CADPZ the CH domain is dissociated from the remaining elements of the protein. Regions zoomed in are shown on the right. Spectra in Figures 2-11 B-D similarly show that the inhibitory helix is largely dissociated from the DH domain in pCADPZ (compare pCADPZ and $CADPZ_{Y174D}$ signals). Together, these data indicate that both the modulatory and core elements are fully disrupted upon full phosphorylation of CADPZ.

occur within a hydrophobic cluster that contains F69 (CH), F425 (PH), Y441 (PH) and F386 (DH extended helix). The second interaction group is immediately adjacent, with the N-terminal helix of the Ac element contacting the basic β5-β6 and β7-α1 loops of the PH domain. Several oppositely charged sidechains are in close proximity across this interface, including D150 (Ac) and K487 (PH) (Fig. 2-6B). CADPZ mutations S67D, F69D and K487E were designed to strongly destabilize the β-strand, hydrophobic cluster and Ac-PH interactions, respectively. These CADPZ mutants had GEF activity ~10% of pCADPZ, very similar to that of ADPZ, which lacks the CH domain entirely. F69A and D406A mutations caused smaller increases in GEF activity, consistent with less severe perturbations of the CH interactions (Fig. 2-5A). Mutation of M66, a residue near the CH-PH interface but making no interdomain contacts had no effect on GEF activity (not shown).

The GEF activity of Vav1 is related to the signaling strength in cells. Previous reports have shown that overexpression of Vav1 proteins lacking the CH domain leads to increased Rac activation and cell transformation (Han, Luby-Phelps et al. 1998; Lopez-Lago, Lee et al. 2000; Bustelo 2001; Llorca, Arias-Palomo et al. 2005), consistent with the increased GEF activity of the ADPZ protein. Further truncation to remove the Ac domain inhibitory helix increases

cellular activity further (Abe, Whitehead et al. 1999), consistent with loss of the core inhibitory element.

To assess the activity of mutant Vav proteins in vivo, Drs. Fernandez-Zapico and Billadeau performed the following Rac activation and transformation assays. As shown in Figure 2-7A, destabilizing the modulatory interactions through the S67D mutation also increases Rac activation in cells. Rac activation is higher in cells expressing the Y174D helix mutant, consistent with higher GEF activity of the corresponding CADPZ protein (Fig. 2-5A), and in the S67D/Y174D double mutant. Similar behavior is observed in focus forming assays, where wild type Vav1 is much less active than the S67D, Y174D and double mutants (Fig. 2-7B). The similarity between the three mutants in this assay likely arises because the long timecourse of the assay flattens differences between partially and fully active proteins. Nevertheless these data suggest that the additional modulatory interactions are relevant for the control of Vav activity under non-stimulated conditions. The combined biochemical and biological data validate the crystal structure and illustrate the importance of the modulatory contacts in controlling GEF activity *in vitro* and *in vivo*.

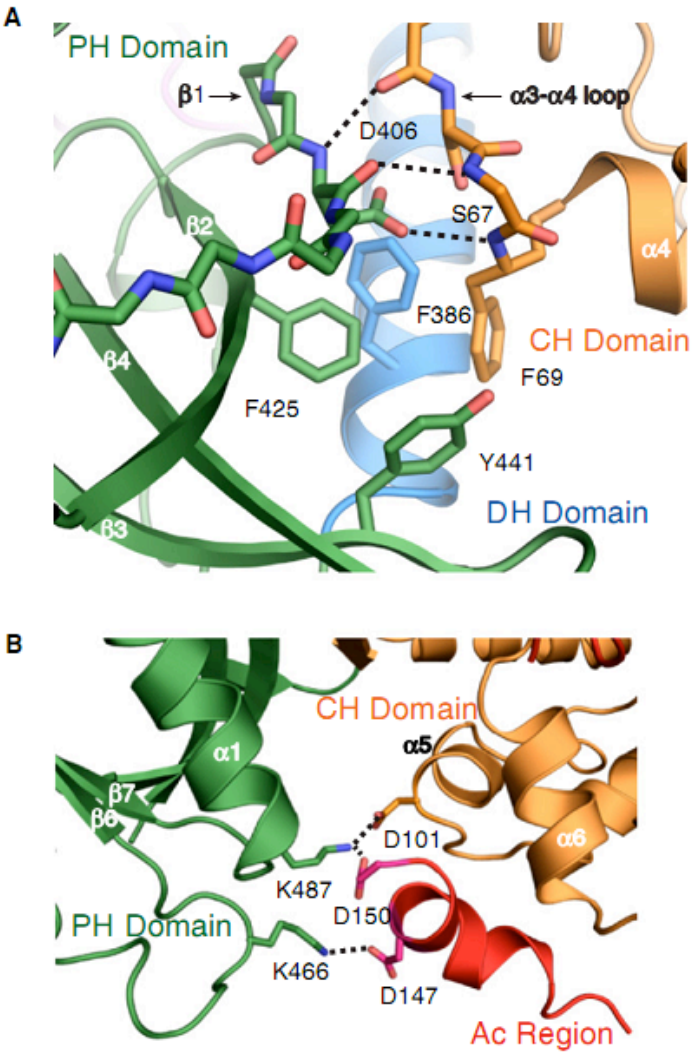


Fig. 2-6 Interdomain contacts in Vav1. Domains colored as in Fig. 1-1. (A) Contacts between CH, PH, and DH domains. Residues discussed in the text are shown as sticks. (B) Contacts between CH domain, Ac region and PH domain. Potential hydrogen bonds and ionic interactions are shown as dashed lines. Regions shown in panels A and B are boxed in black and red, respectively, in Figure 2-2.

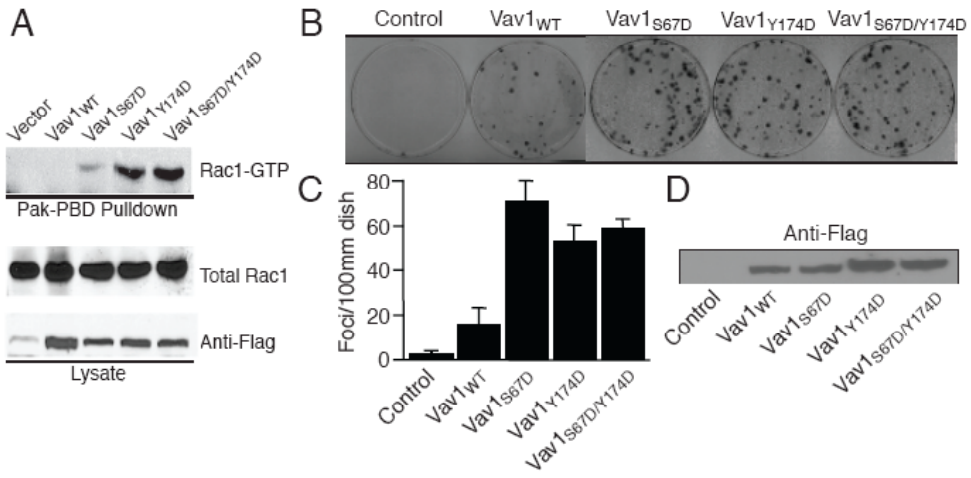


Fig. 2-7 Disrupting modulatory contacts increases Vav activty in vivo. *performed by Martin Fernandez-Zapico and Daniel D. Billadeau. (A) Panc04.03 cells were transfected with a control vector or the indicated full-length Vav1 expression constructs and analyzed for Rac1(GTP) using GST-Pak-PBD. Proteins were resolved by SDS-PAGE and detected by immunoblotting with the indicated antibodies. (B) Neoplastic transformation in NIH3T3 cells was assayed using the foci formation assay. Foci that formed in vector control, wild type Vav1 or the Vav1 mutants (S67D, Y174D and double S67D/Y174D) were stained with crystal violet. Images from representative plates are shown. (C) Bar chart showing that Vav1 mutants have increased number of foci. Each experiment was performed in triplicate, and the mean standard deviation was calculated. (D) Control Western blot analysis show expression of the Vav1 WT and mutants.

Energetic and Structural Mechanisms of Autoinhibition

The data above demonstrate that the modulatory interactions of the CH and Ac elements with the PH domain further suppress the GEF activity of the helix-DH core. The physical basis for this suppression was demonstrated by Ilidio Martins and Pilong Li before my work, and that is the interactions of Vav CH domain shift the helix-DH equilibrium to the helix bound closed state by an additional 10-fold (see Fig. 1-6). With the structure of CADPZ, we now understand it is the coupling of CH-PH and DH-Ac equilbrium that determines the collective suppression. How is coupling between these equilibria achieved structurally? In this regard, the paucity of contacts between the CH domain and the helix-DH core element is striking (Fig. 2-2, right panel). This suggests that the CH domain acts on GEF activity in an indirect manner, perhaps by restraining the N-terminus of the Acidic domain against the DPZ element, and consequently entropically disfavoring the dissociated state of the inhibitory helix. One prediction of this structural mechanism is that cutting the Acidic element immediately before the helix should destroy the mechanical linkage between the modulatory and core elements, preventing the CH domain from affecting GEF activity.

To test this hypothesis, I performed two titrations, monitoring GEF activity (Fig. 2-8). In the first experiment, we added CH (residues, 1-134) to ADPZ (residues 135-584), an interaction with K_D of 32 μM based on independent NMR analyses (not shown, measured by Ilidio Martins). As illustrated in Figure 2-8, the CH domain can act in trans to suppress GEF activity of ADPZ, with activity decreasing linearly with the fraction bound. Thus, a protein cleaved at the CH-Acidic junction can manifest coupling between the modulatory and core equilibria (see also Fig. 2-9). We used resonances of L409δ on PH domain and L325δ1 on DH domain to probe the modulatory and core equlibria respectively. In a fast exchange regime where a system fluctuates rapidly between two states, then the chemical shift of any NMR signal will be a population-weighted average of the chemical shifts of the pure states. Thus, when the chemical shifts of the pure states are roughly known, the chemical shift observed will reflect the relative populations of the two states in equilibrium. In the titration experiments, ADPZ and CADPZ represent the two extremes for open and closed states, both for the CH-PH equlibrium and helix-DH equilibrium. As in fig. 2-9B, titration of CH intro ADPZ shifts the resonance of L409δ closer to the position of CADPZ, indicating the population of PH domain that is bound to CH increases. The same principle applies to L325δ1 resonance (Fig. 2-9C), which shows that the CH titration also shifts the helix-DH equilibrium of ADPZ to a more closed state.

This in trans inhibitory effect is specific, as titrations of CH_{S67D} into ADPZ or CH into $ADPZ_{K487E}$ produced much less inhibition (Fig. 2-9D). In contrast, addition of a CA construct lacking the inhibitory helix (residues 1-168) to A_sDPZ (residues 169-584, K_D = 290 μM) had no effect on GEF activity. Thus, cleaving the Ac element before the inhibitory helix prevents the modulatory equilibrium from communicating to the core equilibrium. In an analogous titration of CA into deuterated, methyl protonated A_sDPZ, monitored by methyl TROSY spectra of A_sDPZ, resonances at the CH-PH interface moved linearly from their positions in free A_sDPZ toward their positions in CADPZ (Figs. 2-9E-F). Thus, the complex reproduced many of the packing interactions of the intact protein. In the same titration, resonances at the helix-DH interface were not affected, indicating that the helix-DH equilibrium is not affected by CH-PH interactions when the Ac element is broken (Fig. 2-9G, compare to Fig. 2-9C). Thus, the decrease in GEF activity tracks with modulation of the helix-DH equilibrium, not simply binding of the CH domain to the PH domain. The data are consistent with the idea that the modulatory interactions act on the helix-DH equilibrium by restraining the N-terminus of the Ac element, disfavoring dissociation of the helix.

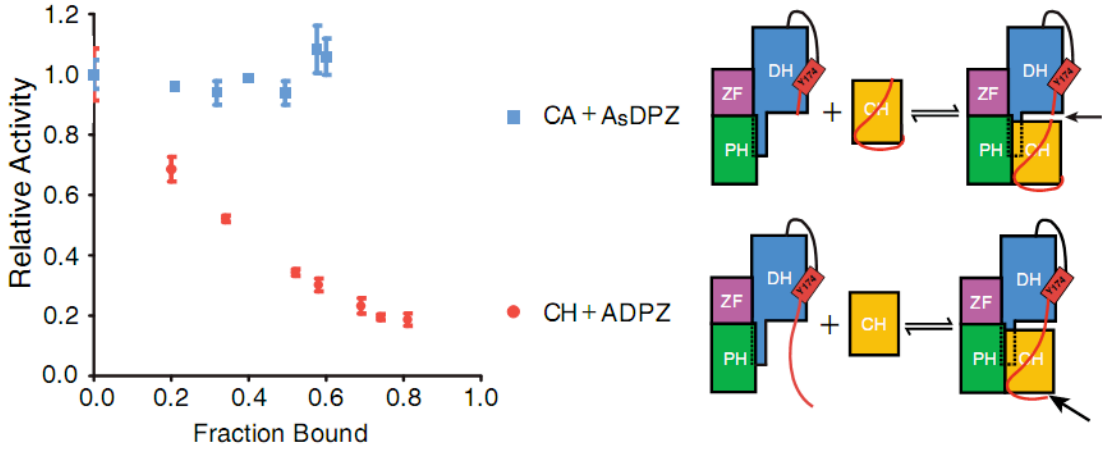


Fig. 2-8 Inhibitory Effects of the CH Domain Require an Intact Ac Element. Relative GEF activity in titrations of CH proteins into ADPZ proteins: CH+ADPZ (red circles), CA+AsDPZ (blue squares). Activities were normalized to those of free ADPZ or AsDPZ. Error bars represent SD from three independent measurements. Cartoon on the right depicts the corresponding titration reaction. Arrow indicates the position where the Ac element is cut.

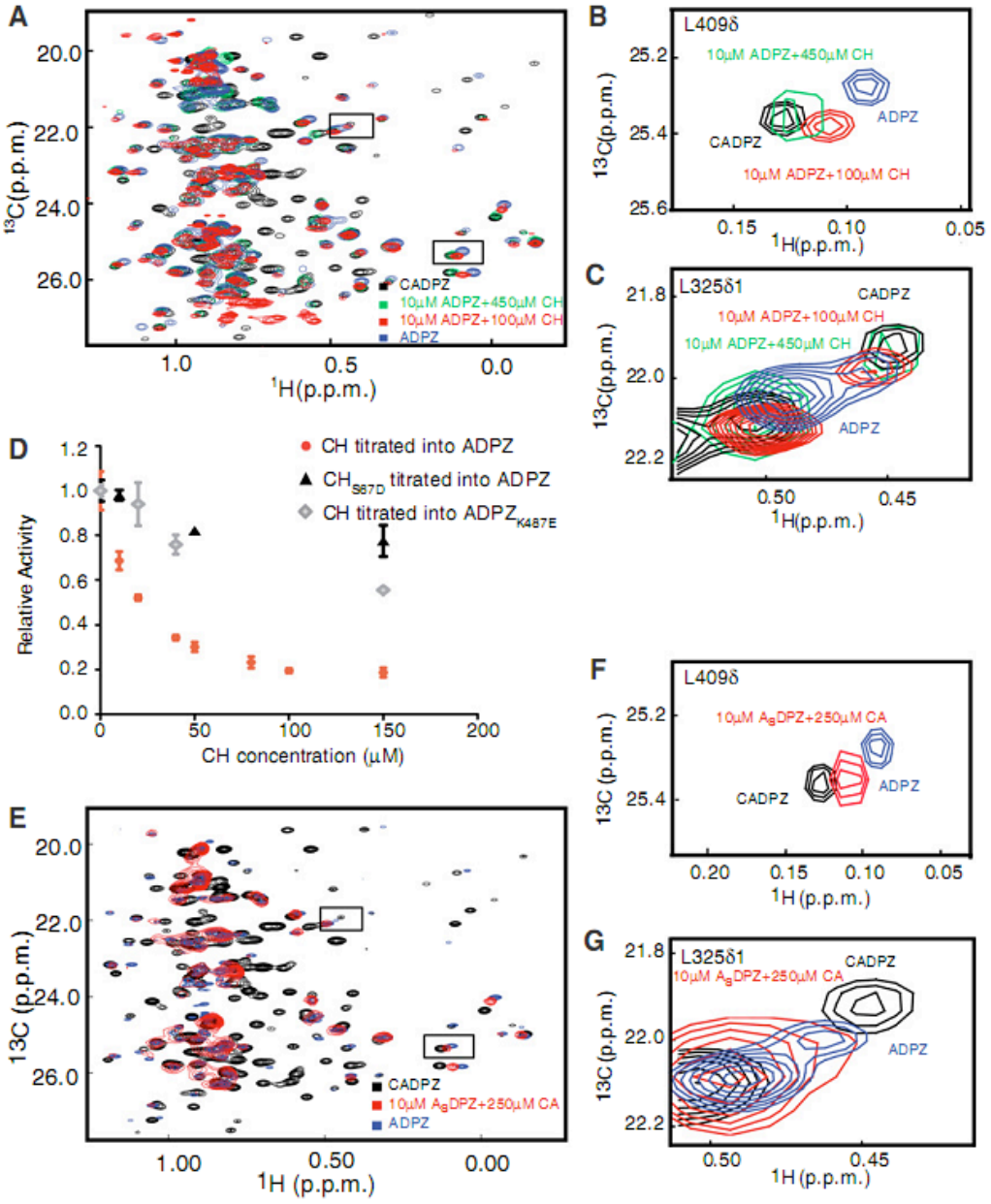


Fig. 2-9 Titrations of CH/CA into ADPZ/A$_s$DPZ (A) Overlaid ^{1}H/^{13}C methyl TROSY spectrum of CADPZ (black), ADPZ (blue) and 10 μM ADPZ + 100 μM CH domain (red), 10 μM ADPZ + 450 μM CH domain (green). Regions enlarged in (B) and (C) are boxed.

(B and C) Signals of L409δ and L325δ1, respectively. Note that when the Ac element is intact, signals at both the CH-PH and helix-DH interfaces shift to the CADPZ positions as the CH domain is added to ADPZ. (A), (B), and (C) are contoured at different levels for clarity. (D) Inhibition of ADPZ GEF activity by CH protein in trans is specific. CH titration in ADPZ is red (same data as in Figure 2-8). CH_{S67D} was titrated into ADPZ (black), CH was titrated into $ADPZ_{K487E}$ (gray). GEF activities were normalized to free ADPZ or ADPZK487E activity. Normalized GEF activities are plotted against CH or CH_{S67D} concentrations. Error bars indicate standard deviation for three independent measurements. (E) Overlaid $^{1}H/^{13}C$ methyl TROSY spectrum of CADPZ (black), ADPZ (blue), and 10 μM A_sDPZ + 250 μM CA (red). Regions enlarged in (F) and (G) are boxed. (F) and (G) Signals of L409δ and L325δ1, respectively. Note that in contrast to the ADPZ + CH titration in Figures 2-9A–2-9C, when the Ac element is cut (in the A_sDPZ + CA titration here) although the signal at the CH-PH interface (L409δ) shifts toward the CADPZ position when CA is added, the signal at the helix-DH interface (L325δ1) does not. Thus, when the Ac element is cut, coupling between the modulatory and core equilibria is lost. (E), (F), and (G) are contoured at different levels for clarity.

Validation of the Energetic Mechanism by Mutagenesis

To investigate the idea of coupled equilibria in greater detail, I examined the NMR properties of the proteins described above, which are impaired in one or both elements of the inhibitory apparatus. For each protein I examined two sets of NMR resonances. The first set consists of methyl groups located at the CH-PH interface including L409δ1,2, L491δ from the PH domain and V102γ from the CH domain (see Figs. 2-10). As described earlier, the chemical shifts of these methyl signals should reflect the populations of CH bound or dissociated states. The bound extreme is represented by the CADPZ chemical shifts, and the dissociated extreme by the chemical shifts of ADPZ or pCADPZ, where the modulatory contacts are absent or fully destabilized, respectively (Figs. 2-10 A-D). For each protein I also examined NMR signals of methyl groups at the helix-DH interface (Figs. 2-10 E-G). As in Fig. 2-10, these signals report on populations across the core inhibitory equilibrium. By comparing CH-PH and helix-DH signals for each mutant, I can observe how changes to one process affect the other. We can also correlate both equilibria to GEF activity. The behavior of ADPZ is demonstrative. In this protein the CH-PH signals are strongly shifted from their CADPZ positions, reflecting the absence of modulatory interactions. In contrast, since the core inhibitory module is intact, the helix-DH signals in ADPZ are

shifted only partially from the CADPZ positions toward the pCADPZ positions (Fig. 2-10 A-D). The GEF activity of ADPZ, which is dependent on exposure of the DH active site, parallels the helix-DH signals (Fig. 2-5). In the CADPZ mutants, S67D, F69D and K487E, which were designed to destabilize the modulatory interactions, the CH-PH signals shift appreciably toward their open positions (Fig. 2-10A-D). The helix-DH signals shift toward the open position but remain in the range between the CADPZ and ADPZ positions (Fig. 2-10E-G). These changes are consistent with the GEF activities of these mutants, which are at or below that of ADPZ (Fig. 2-5). The less perturbing mutations F69A and D406A show smaller shifts in the CH-PH signals, much smaller shifts in the helix-DH signals and lower GEF activity. Thus, shifting the modulatory equilibrium changes the helix-DH equilibrium and consequently GEF activity. Perturbations can also be observed in reciprocal fashion: introducing the Y174D (helix) mutation into the F69A background shifts the CH-PH signals toward their positions in ADPZ (Fig. 2-10A-D). Thus, destabilizing the helix-DH interactions also destabilizes the modulatory interactions. These data support our model that the core and modulatory equilibria are thermodynamically coupled, and that this coupling is a major contributor to the cooperative suppression of GEF activity by interactions of the CH domain.

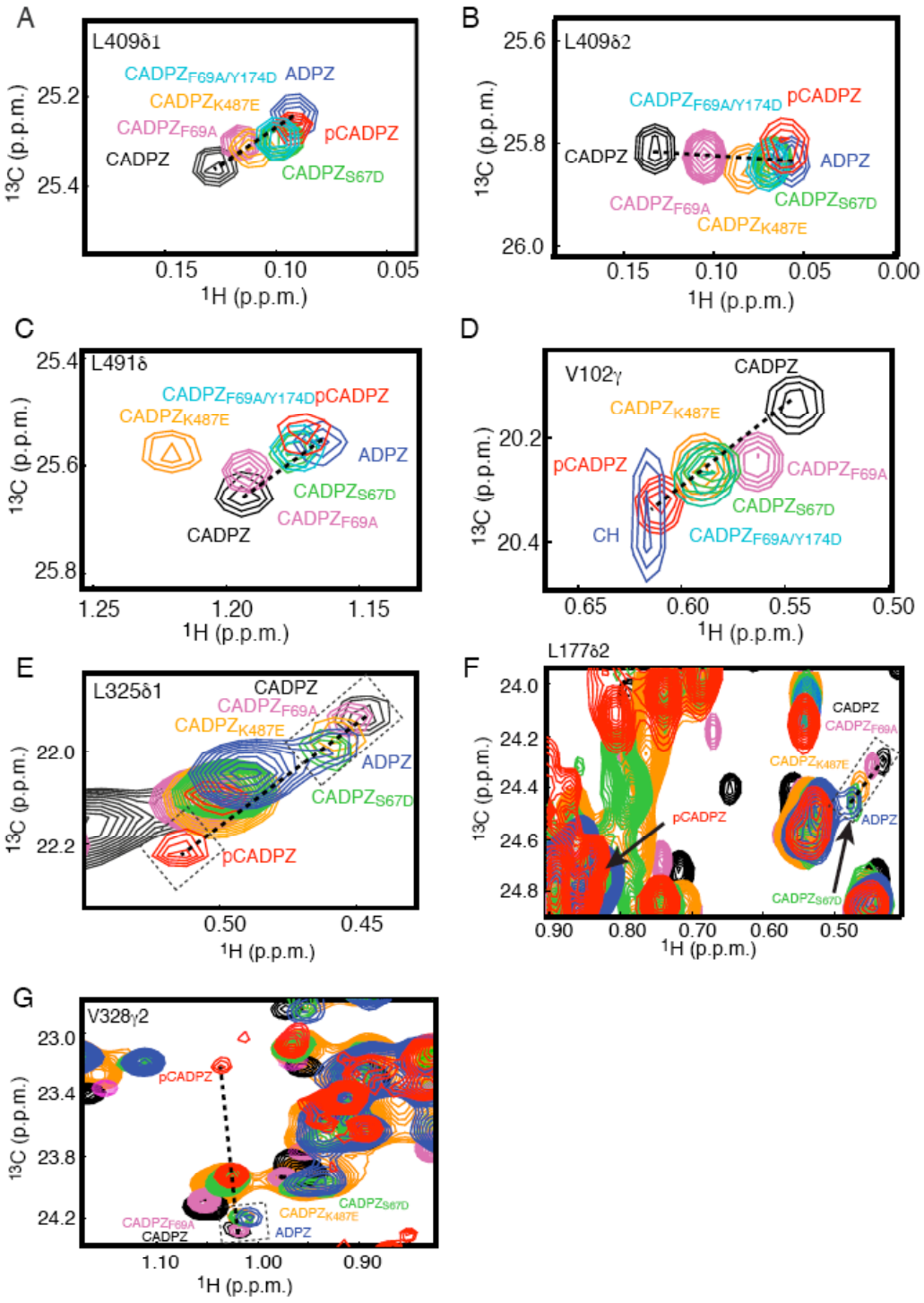

A
L409δ1
CADPZF69A/Y174D
ADPZ
CADPZK487E
CADPZF69A
pCADPZ
CADPZ
CADPZS67D
13C (p.p.m.)
1H (p.p.m.)
B
L409δ2
CADPZF69A/Y174D
pCADPZ
CADPZ
ADPZ
CADPZF69A
CADPZS67D
CADPZK487E
C
L491δ
CADPZF69A/Y174D
pCADPZ
CADPZK487E
ADPZ
CADPZS67D
CADPZF69A
CADPZ
D
V102γ
CADPZ
CADPZK487E
pCADPZ
CADPZF69A
CADPZS67D
CH
CADPZF69A/Y174D
E
L325δ1
CADPZ
CADPZF69A
CADPZK487E
ADPZ
CADPZS67D
pCADPZ
F
L177δ2
CADPZ
CADPZF69A
CADPZK487E
pCADPZ
ADPZ
CADPZS67D
G
V328γ2
pCADPZ
CADPZS67D
CADPZK487E
CADPZF69A
CADPZ
ADPZ

Fig. 2-10 CH-PH and helix-DH equilibria are thermodynamically coupled. (A)-(D) Overlaid $^{1}H/^{13}C$ methyl TROSY spectra of L409δ1, L409δ2, L491δ, V102γ (all at CH-PH interface) resonances in human CADPZ (black), $CADPZ_{F69A}$(magenta), $CADPZ_{F69A/Y174D}$ (cyan), $CADPZ_{S67D}$ (green),$CADPZ_{K487E}$ (orange), pCADPZ (red), and ADPZ (blue) proteins. (E)-(G) Overlaid $^{1}H/^{13}C$ methyl TROSY spectra of the L325δ1, L177δ2, V328γ2 resonances (DH-helix interface). Spectra colored as in (A). In (E) and (F), reporter signals are boxed; dashed lines connect endpoints.

Stepwise phosphorylation and activation of Vav1

The kinetic pathway of Vav1 activation has been puzzling for us. As it is known to occur *in vivo*, Y174 is rapidly accessed by kinases (Miletic, Graham et al. 2007), but its exposure is strongly suppressed by the combined actions of the core and modulatory equilibria. We hypothesized that residues Y142 and Y160 could be critical initial phosphorylation sites on the kinetic pathway to activation. In order to test this idea, I purified Y142 and Y160 singly or doubly phosphorylated protein through ion exchange chromatography, and made phosphorylation mimicking Y to D protein mutants as well in order to get more material for NMR studies. The Y to D mutants have GEF activity very similar to their pY counterparts (generated and purified by Ilídio Martins), indicating that the mutations act similarly as phosphorylation (Fig. 2-11A). The GEF activity of the doubly phosphorylated pY142/pY160 or $CADPZ_{142D/Y160D}$ protein was found to be similar to ADPZ, and the activity change of the mutants was consistent with a shift in the helix-DH equilibrium toward the ADPZ position as assessed by NMR (Figs. 2-11B-D). These data suggest that initial phosphorylation at Y142 and Y160 could destabilize the modulatory interactions, increasing access of Y174 to kinases.

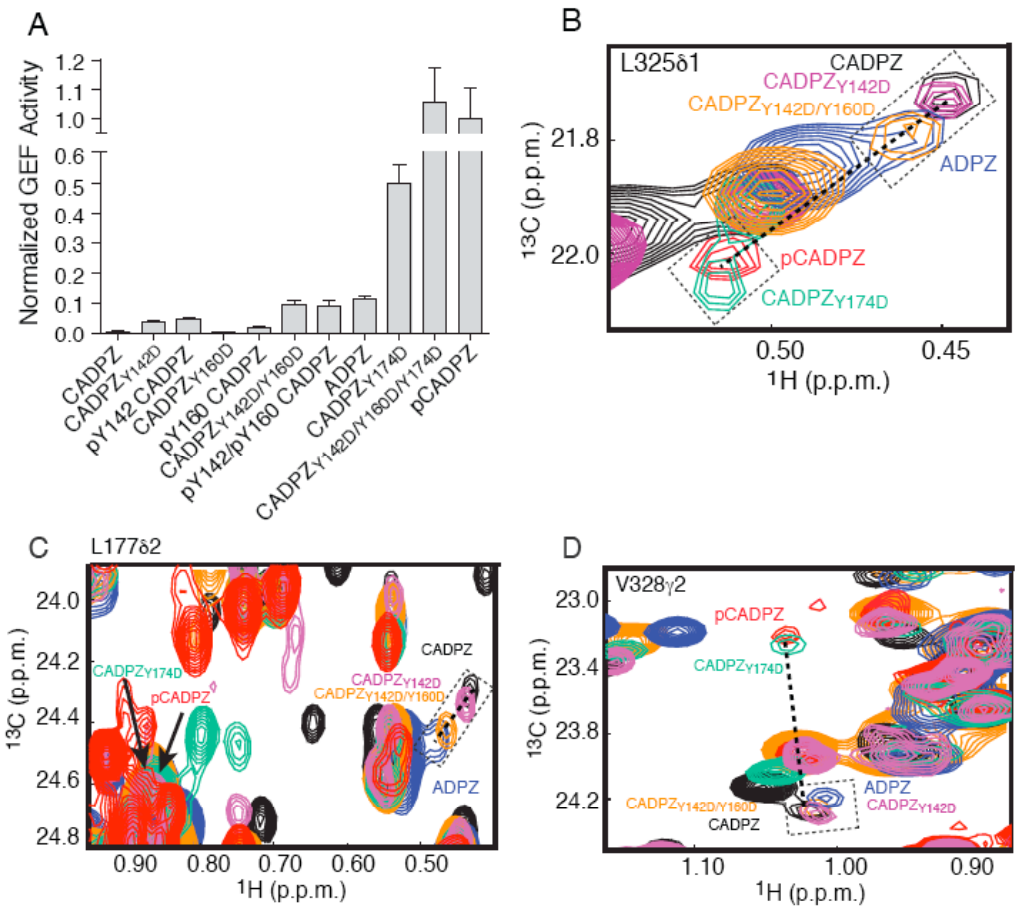


Fig. 2-11 Phosphorylation or mutation of Y142 and Y160 destabilize CH-PH and helix-DH equilibria, increasing kinase access to Y174. (A) Normalized GEF activity of human Vav1 proteins. Error bars represent standard deviations from three independent measurements. (B)-(D) Overlaid $^{1}H/^{13}C$ methyl TROSY spectra of the L325δ1, L177δ2, V328γ2 resonances in human CADPZ (black), ADPZ(blue), and $CADPZ_{Y142D}$ (magenta), $CADPZ_{Y174D}$ (green), $CADPZ_{Y142D/Y160D}$ (gold), pCADPZ (red) proteins.

In order to further validate this model, my colleague Ilidio Martins performed the following phosphorylation kinetic studies. He used an HPLC-based assay to quantify the degree of phosphorylation by [γ-^{32}P] ATP at each site individually (Y142, Y160, Y174) when CADPZ is treated with the Lck kinase domain. He found that Y174 was phosphorylated by the Lck kinase domain much more rapidly in the Y142D/Y160D mutant than in wild type CADPZ, as shown in Fig. 2-12A. More importantly, in wild type CADPZ, Y142 and Y160 are phosphorylated more rapidly than Y174 (Fig. 2-12A). This is striking since in short peptides containing the different Vav1 tyrosine motifs, phosphorylation of Y174 is strongly favored over Y142 and Y160 (Amarasinghe and Rosen 2005). These data suggest a stepwise pathway dictated by the structural organization of CADPZ in which Y142 and Y160 are modified initially, followed by Y174.

Based on the conformation we observed in the crystal, Y142 and Y160 make contacts with CH domain and look buried. Then how could these two sites be accessed by kinase first? NMR relaxation data collected by Ilidio Martins indicated substantial μs-ms timescale fluctuations of methyl groups surrounding these two tyroisines in the Ac element (Fig. 2-12B). These dynamics likely represent transitions to kinase accessible conformations, explaining how Y142 and Y160 can be phosphorylated rapidly. Upon initial phosphorylation of

Y142/Y160, destabilization of modulatory contacts and recruitment of kinase should act synergistically to increase phosphorylation of Y174 by Lck. Thus, the structural and energetic architecture of Vav1 provides a facile kinetic route to activation in the face of substantial suppression of basal activity. Since Y142 and Y160 are conserved in 31 of 33 available Vav sequences (*Aedes aegypti* and *Tribolium castaneum* sequences have only Y142), this mechanism is likely to be general across the family.

Discussion

Our collective data show that the regulatory apparatus of Vav1 is composed of at least two thermodynamically coupled processes (Fig. 2-13). A core process, in which the inhibitory helix blocks the DH active site, provides approximately 10-fold suppression of activity and phosphorylation rate (Li, Martins et al. 2008). Binding of the CH domain to the N-terminus of the Acidic element and to the DPZ fragment restricts the inhibitory helix, thermodynamically coupling these interactions to the core. This coupling shifts the helix-DH equilibrium ~10-fold further toward the inhibited (helix bound) state. Together, the two processes provide the high level of suppression (~100-fold) that is apparently needed for in vivo function (since

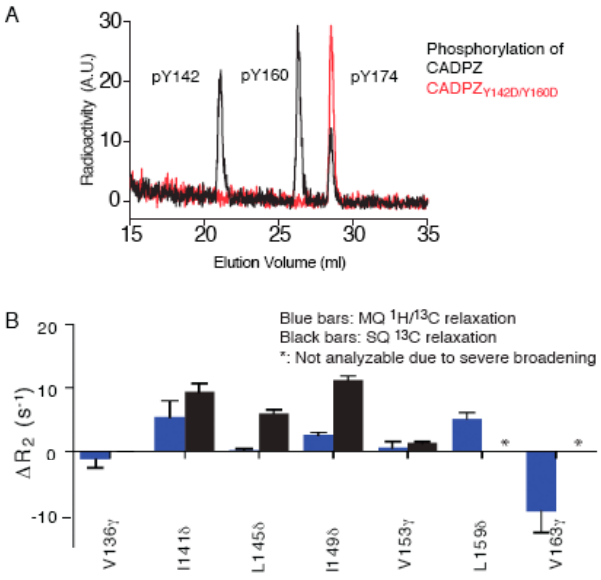


Fig. 2-12 Fluctuations in the Ac Element Enable Stepwise Activation of Vav1. * performed by Ilidio Martins. (A) HPLC analysis (with scintillation detection) of tryptic digests of CADPZ (black) and $CADPZ_{Y142D/Y160D}$ (red) after phosphorylation with [γ-^{32}P] ATP for 7.5 min. Peaks corresponding to pY142-, pY160-, and pY174-containing peptides are indicated. At complete phosphorylation, the peak corresponding to pY174 has virtually identical amplitude for the two proteins. (B) Methyl groups in the Ac element of CADPZ show substantial ms-ms timescale fluctuations. Multiple quantum $^{1}H/^{13}C$ $\Delta R2$ (blue bars) and single quantum ^{13}C $\Delta R2$ (black bars) for methyl groups in the Ac element of CADPZ. Data were recorded on an 800 MHz spectrometer at 15°C. Error bars indicate standard deviations of two independent measurements. Asterisks indicate resonances where $\Delta R2$ could not be analyzed due to severe signal broadening.

CH truncation leads to cell transformation (Bustelo 2001); (Katzav, Cleveland et al. 1991; Llorca, Arias-Palomo et al. 2005). This 10-fold x 10-fold construction of Vav1 (rather than a single 100-fold inhibitory equilibrium) is likely functionally important, as we have shown that inhibition by the core alone can be increased to 50-fold simply by two point mutations (Li et al., 2008). Biochemical data on other multi-domain systems suggest similar (relatively small, 2-20-fold) magnitudes of core suppression and modulatory coupling (DiNitto et al., 2007; Moarefi et al., 1997).

This construction has important implications for the biology of Vav1. It provides a quantitative rationale for the increased activity of CH domain truncations of the protein. In such truncations, even though the core inhibitory module is intact, the inherent thermodynamics of the module produce high levels of the active state (9-33%, Figs. 2-11) when coupling to the modulatory equilibrium is lost. This effect increases both the phosphorylation rate (and likely steady-state phosphorylation level) and basal GEF activity of Vav1, explaining increased Rac activation and cell transformation (Bustelo 2001) (Han, Luby-Phelps et al. 1998; Lopez-Lago, Lee et al. 2000; Llorca, Arias-Palomo et al. 2005). The ability of CH interactions to modulate the AD core also implies that molecules that disrupt these contacts could partially stimulate Vav1 activity, and act cooperatively with kinases to fully activate the protein in vivo.

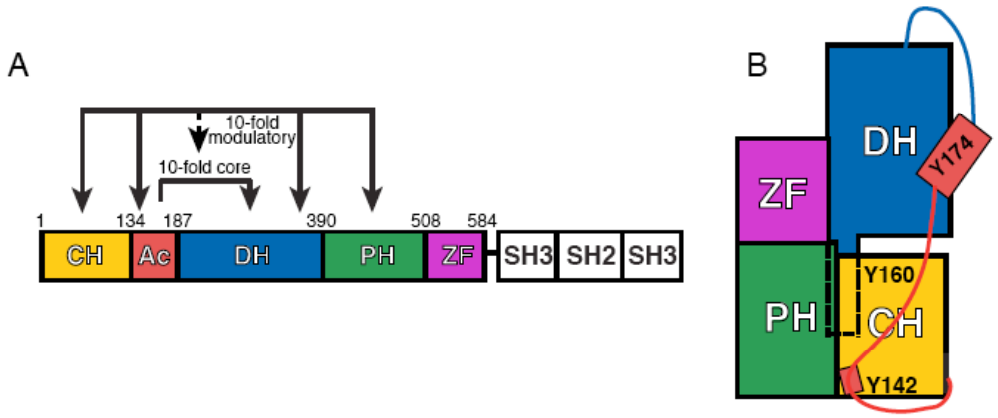


Fig. 2-13 Structural and energetic models of autoinhibition in Vav1. (A) Domain architecture of Vav1 and thermodynamic model for cooperative inhibition. Solid lines represent direct physical contacts; dashed line represents thermodynamic coupling between core and modulatory equilibria. (B) Schematic representation of autoinhibited CADPZ. Dotted lines outline C-terminal helix of DH domain lying behind PH and CH domains.

The creation of suppression from multiple weak interactions also has important implications regarding the kinetics of Vav1 activation upon receptor stimulation. In many allosteric systems, it has been shown that activators gain access to their targets by recognizing the small population of active state present under basal conditions (Henzler-Wildman and Kern 2007). This is the case in the AsD core of Vav1, where kinases can only recognize Y174 in the helix-dissociated state (Li, Martins et al. 2008). However, strong suppression of activity opposes this pathway, since it requires that the accessible population be very small, which would tend to make activation slow. The layered energetic and physical construction of modular proteins provides a likely general solution to this problem, which we have observed for Vav1. That is, the structure and energetics of multi-domain systems can be organized so that the modulatory elements are more accessible to activators (e.g. Y142 and Y160 in the Ac element show significant μs-ms timescale dynamics and can be rapidly phosphorylated). This allows activation to occur in stepwise fashion, with initial events disabling the modulatory interactions, and consequently increasing access to the core. Thus, modularity enables evolution not only of strong inhibition, but also of ready routes to activation.

A variety of biochemical and structural data suggest that the energetic and kinetic mechanisms that we have found for Vav1 are likely to be widespread among multi-domain proteins. The Tim family of GEFs represents a structurally and functionally similar example (Yohe, Rossman et al. 2008). The Tim proteins have a DH domain whose activity is inhibited by an adjacent helix that binds in the active site, and that can be displaced by phosphorylation, analogous to the Vav AD core. The Tim proteins also have a C-terminal SH3 domain that further suppresses the activity of the helix-DH core through binding an internal proline-rich motif (PRM). The mechanism of suppression by the SH3 domain has not been explored, but available data would be consistent with coupling of the helix-DH and SH3-PRM equilibria. Moreover, although the kinetic pathway to Tim activation has not been examined experimentally, it has been speculated that activation may occur stepwise, with initial displacement of SH3-PRM interactions facilitating phosphorylation of the inhibitory helix (Yohe, Rossman et al. 2008). The unrelated Ras GEF, SOS, also appears to use the same regulatory logic, with the DH-PH element shifting an allosteric equilibrium in the Rem-Cdc25 element toward its inactive state (Sondermann, Soisson et al. 2004).

The broadest class of, and best characterized, proteins that appear to behave in this manner are the protein kinases. A large body of data suggests that the isolated kinase domain fluctuates between an inactive state, where the active site

is not properly organized for catalysis, and an active state, where the catalytic cluster and substrate binding site are catalytically competent (e.g. (Levinson 2006; Vogtherr, Saxena et al. 2006; Masterson, Mascioni et al. 2008). Phosphorylation of residues in the so-called activation loop stabilizes the active conformation, increasing catalytic activity. While the active state is common among many kinases, the inactive states differ. Sampling of the inactive conformation has been capitalized upon in the Abl tyrosine kinase inhibitor imatinib, a widely used anti-cancer therapeutic (Schindler, Bornmann et al. 2000). Imatinib distinguishes Abl from other closely related kinases by selectively binding to its unique inactive conformation. NMR analysis of fluctuations between the active and inactive states in isolated kinase domains is an active area of research. However, to date such studies have not yielded a quantitative understanding of kinase thermodynamics (populations of the states) or a detailed view of the equilibrating structures in the absence of stabilizing inhibitors ((Masterson, Mascioni et al. 2008) (Vajpai, Strauss et al. 2008) (Vogtherr, Saxena et al. 2006)). Many kinases also contain additional domains that repress the activity of the catalytic domain (Huse and Kuriyan 2002). Disabling such domains, either by truncation or through regulatory interactions, increases the activity of kinases toward various substrates. A detailed interpretation of the biochemical data in terms of allosteric equilibria are complicated by the fact that phosphorylation of the activation loop locks the kinase in an active state, and that the activity of non-phosphorylated

kinases is very low and thus hard to quantify. However, molecular dynamics simulations of the Src kinase have suggested that, at least in that system, the additional SH3 and SH2 domains suppress activity through biasing the regulatory equilibrium in the kinase domain toward the inactive state ((Young, Gonfloni et al. 2001; Faraldo-Gomez and Roux 2007)). As in Vav, this construction also appears to dictate that kinases are activated in stepwise fashion, with initial events that disable modulatory interactions (e.g. in Src by ligand binding to the SH3 domain) (Moarefi, LaFevre-Bernt et al. 1997) enabling phosphorylation of the activation loop and consequent maximum stabilization of the active state.

These various examples strongly suggest that although multi-domain proteins use a variety of structural mechanisms to achieve regulation, the underlying energetic mechanism that we have illustrated for Vav, and the requirement for an ordered activation process that arises from it, are likely to be widely observed. The NMR approach we developed here provides a means to directly measure the populations of different states across regulatory equilibria and has allowed us to quantitatively characterize the energetic landscape of Vav1. This in turn has established coupled equilibria as a major mechanism of interdomain cooperativity in this system. Application of these methods to other systems should reveal how widespread cooperative inhibition through coupled equilibria is in multi-domain proteins.

Materials and Methods

DNA Cloning and Proteins

Vav genes from mouse or human species were both cloned into pMal vector with maltose binding protein (MBP) and a subsequent tobacco etch virus (TEV) protease cleavage site at the N terminus of the gene. Mutations of Vav constructs were made using quick change site-directed mutagenesis kit from NEB. WT or mutated Vav CH (residues 1-134), CA (1-168), A_sD (169-375), AD (135-375), A_sDPZ (169-584), ADPZ (135-584), and CADPZ (1-584) proteins were expressed in *E. coli* strain BL21(DE3)$T1^R$. Expression in LB media was induced at optical density ($O.D._{600}$) ~ 0.8 with 0.5 mM IPTG. Some CADPZ mutants were co-expressed with groES-groEL chaperones in order to increase protein yield (Takara Bio Inc). Chaperone expression was induced at OD_{600} ~ 0.25 with 200 mg/l of arabinose. Proteins were purified by successive weak anion exchange (DEAE), amylose affinity and strong anion exchange chromatographies (Source15Q). The first DEAE column was used to remove the aggregated CADPZ protein, which comprises around 90% of the total expressed protein. This step is necessary to get CADPZ material in monomeric form. After concentrating protein, MBP tag was cleaved at room temperature for one hour. Successive

strong anion exchange (S15Q) and gel filtration chromatographies were used to further purify Vav from uncleaved protein, MBP tag and TEV enzyme.

U-[^{15}N,^{2}H], Ile-[^{13}C$^{\delta}$H$_3$]-,Leu-[^{13}CH$_3$,^{12}CD$_3$],Val-[^{13}CH$_3$,^{12}CD$_3$]-labeled proteins were produced in bacteria grown in deuterated M9 media (Korzhnev, Kloiber et al. 2004). Transformed bacteria were grown in 5ml LB for overnight, then transferred to 200ml 50% deuterated LB. Around 4 hours later, when O.D.$_{600}$ reached 0.8, the LB media was spun down and bacteria was resuspended and transferred into 1L deuterated M9 media. About 1 hour before induction (when O.D.$_{600}$ reached ~0.4), 50 mg/l 2-keto-3-methyl-d$_3$-3-d$_1$-4-^{13}C-butyrate and 50 mg/l 4-[^{13}C,^{1}H]-3,3-^{2}H-α-ketobutyrate were added to the media. 1mM IPTG was added when O.D. reached 1.0 to induce protein expression, and bacteria were allowed to grow 6-8 more hours in M9 media. All proteins were induced as in LB media, except for CADPZ, ADPZ, and A$_s$DPZ, which were co-expressed with groES-groEL.

To generate CADPZ proteins that are phosphorylated selectively at Y142, Y160 and Y142/Y160, kinase reactions were quenched with 40 mM EDTA after running for a limited time (30 min - timing was optimized to get maximum amount of Y142 and Y160 doubly phosphorylated protein) and loaded onto a 1 ml anion exchange column (MonoQ, GE Healthcare) in 20 mM HEPES pH=7.0, 100 mM NaCl, 10% Glycerol, 2 mM TCEP. Buffer pH was found to be critical for separating different phosphorylated species of Vav. Proteins were eluted with a

shallow gradient of 100 mM to 300 mM NaCl in 80 ml, giving seven separate peaks. The identity of the phosphorylated species in each peak was determined using mass spectrometry based on total mass and LC/MS/MS analyses of digests with chymotrypsin, GluC or AspN.

Crystallization and x-ray diffraction data collection

Crystals of human Vav1 CADPZ (~50 x 100 x 200 μm) were grown at 4 °C using the hanging-drop vapor-diffusion method. Drops contained 2 μl of 3-10 mg/ml protein solution (20 mM Tris (pH 7.5), 50 mM NaCl, 5% (w/v) glycerol, 2 mM TCEP) and were equilibrated against 200 μl of water. Crystals of selenomethionine-labeled protein were obtained under similar conditions when seeded with native crystals. Crystals were cryoprotected by addition of ethylene glycol to 30% (v/v). Native and anomalous dispersion data were collected at 100 K using the Advanced Photon Source 19ID beamline. Vav crystals tend to decay under the x-ray beam. In order to get good completeness and redundancies, data were collected at three positions of the crystal and merged for scaling. The native dataset was collected at the anomalous edge for zinc. Vav1 CADPZ crystallized in space group $P2_1$ with unit cell parameters of a = 85 Å, b = 59 Å, c = 161 Å, β = 97 degrees, containing two molecules in the asymmetric unit with ~55% solvent. Crystals diffracted x-rays to a minimum Bragg spacing of ~2.65

Å. Data were processed using the HKL3000 suite of programs (Minor, Cymborowski et al. 2006).

Phase determination and structure refinement

Phases were obtained from a single-wavelength anomalous dispersion experiment using a selenomethionine-Vav1 CADPZ crystal with data to a resolution of 2.80 Å. Thirty of 34 expected selenium sites were located using the program SHELXD (Schneider and Sheldrick 2002). Phases were refined with the program MLPHARE (Otwinowski 2001), resulting in a figure-of-merit of 0.27. Phases were further improved by density modification and two-fold non-crystallographic averaging with the program DM (CCP4 (Collaborative Computational Project 1996) resulting in a figure-of-merit of 0.88. An initial model containing about 54% of all residues was automatically generated using the program ARP/warp (Morris, Zwart et al. 2004). Additional residues were manually added using the program coot (Emsley and Cowtan 2004). Refinement was carried out with native data to a resolution of 2.73 Å in the program PHENIX (Adams, Grosse-Kunstleve et al. 2002) and consisted of simulated annealing, conjugate-gradient minimization and refinement of individual B-factors and TLS parameters, interspersed with manual revisions of the model. The current model (pdb code: 3KY9) contains two Vav1 CADPZ monomers (labeled A and B,

respectively) in the asymmetric unit with backbone r.m.s.d. of 0.13 Å for residues 2-129, 142-150, 156-180, 189-417, 419-456, 463-478, 482-564.

NMR spectroscopy

NMR analyses were performed in buffer containing 20 mM Tris (hydroxymethyl-d_3) amino-d_2-methane, pH 7.5 (uncorrected), 100 mM NaCl, 2 mM TCEP, and 10% (w/v) d_8-glycerol; 99.9% D_2O. Data were acquired on 600 and 800 MHz Varian Inova spectrometers equipped with cold probes. Methyl-TROSY experiments were performed at 25 ˚C. Data were analyzed using software NMRDraw and NMRView5 (http://www.chem.ucsb.edu/MagRes/protocols/nmrPipe.html).

Guanine nucleotide exchange assays

Human Rac1 gene was cloned into pET21a vector. GDP bound Rac1 protein was expressed in *E.coli* and purified through an ion exchange column (S sepharose) and a second gel filtration chromatography. Before performing guanine nucleotide exchange assays, Rac1-GDP was exchanged against fluorescent N-methylanthraniloyl-GMPPNP (mant-GMPPNP) for 20 minutes. In each assay, 10 μM Vav1 CADPZ, A_sDPZ, or ADPZ was used to catalyze 0.2 μM

Rac1-GDP. I monitored the release of Mant-GMPPNP from Rac1 through the decrease of Mant fluorescence intensity (λ_{exc} = 350 nm; λ_{em} = 432 nm). Experiments were performed using a Jobin Yvon Horiba Fluorimeter (www.jyhoriba.com).

Results were analyzed using Matlab and fitted against a single exponential decay as $\mathbf{F(t) = \Delta F*exp(-R*t)+ F_0}$, where ΔF is total fluorescence decrease, R is decay rate constant and F_0 is the equilibrium fluorescence intensity. Under the assay conditions I used, the decay rate constant, R, scales linearly with GEF concentration. The normalized GEF activity of each protein was calculated from the equation below,

R(normalized)= R(i)-R(control)/[R(open)-R(control)]

R(control) here is the rate measured without any enzyme present, and R(open) used here is the rate measured for pCADPZ. R(i) represent the rate measured for the target protein.

For each assayed protein, three independent measurements were performed to calculate the standard deviation. Figures were prepared later using the software Prism. In the titrations of Figure 2-8, binding affinities used to calculate fraction

bound were determined independently by Ilidio Martins under identical solution conditions through NMR titrations of unlabeled ADPZ proteins into methyl-labeled CH-containing proteins. In this figure, the relative activity was normalized against the rate of free ADPZ or A_SDPZ instead of pCADPZ.

Chapter 3

Concluding Remarks I

Previous NMR-based structural and dynamic studies of Vav1 revealed a core regulatory module (AD) in which the inhibitory helix transits between a DH bound state (90%) and a dissociated state (10%) (Aghazadeh, Lowry et al. 2000) (Li, Martins et al. 2008). The existence of this pre-equilibrium allows Src family kinases to access the modification site and activate the DH domain. In full-length Vav, stronger suppression is required for normal cell behaviors (Bustelo, Ledbetter et al. 1992). Li et. al demonstrated that in the context of the first five domains of Vav1 (CADPZ), the core inhibitory equilibrium is further biased by ~ 10-fold through other domain domain contacts. The detailed interactions coupled to the core inhibitory interactions and the mechanisms through which the coupling is achieved were not known.

Through the work described in the previous chapter, we now understand that it is the coupling between CH-PH (modulatory interactions) and DH-helix equilibria that controls the cooperative suppression. The CH domain binds to both the Ac and DPZ elements, restricting the inhibitory helix toward the DPZ fragment and consequently disfavoring entropically the dissociated state of the

inhibitory helix. Phosphorylation of Y142 and Y160 could destabilize the modulatory interactions, increasing access of Y174 by kinases. Fluctuations around Y142 and Y160 suggest that these two residues might be the initial phosphorylation sites. Thus, upon phosphorylation of Y142 and Y160, destabilization of modulatory contacts and recruitment of kinases could work synergistically to increase the phosphorylation rate of Y174 by Src family kinases.

Despite the considerable progress made to uncover the mechanism of autoinhibition and activation of Vav, there are still questions remaining to be answered. For example, the different conformations that Y142 and Y160 are sampling (whether they represent CH bound or dissociated states) are left to be determined. In addition, the regulation of Vav function may be far more complex than we have realized here. Below are two interesting observations during my studies that suggest additional regulatory roles of Ac region and C-terminal SH3-SH2-SH3 motif.

1. The GEF activity of $ADPZ_{Y174D}$ is about two-fold higher than A_sDPZ_{Y174D}, implying that the N-terminus of the Ac region also contributes to Rac binding and activation.

2. In my *in vitro* GEF assays, $CADPZ_{F69A}$ was found to be slightly active. On the contrary, full length Vav (CADPZ + SH3-SH2-SH3) with mutation F69A was

completely inactive. These results suggest that the C-terminal SH3-SH2-SH3 motif may add another layer of autoinhibition onto DH domain.

Crystal structures of $ADPZ_{Y174D}$ complexed with Rac1 and full length Vav will address the physical basis for these observations. Although I have attempted to crystallize the full length Vav1, the protein displays strong degradation behavior, resulting in very limited amounts of material from bacterial culture. Moreover, robotic screens of full length Vav1 protein did not yield any crystals. It is likely that in the future, a deeper understanding of Vav regulation will be gained through investigating these questions.

The hierarchical architecture of coupling multiple weak interactions in Vav1 has the advantage of achieving both strong suppression and the ability to be rapidly activated. Molecules with such construction are prevalent in nature (listed in Table 3-1). Examples include GEF proteins, kinases, and phosphatases. Studies have been performed to reveal the activities of these proteins at different states, but information regarding the coupled equilibria is still lacking. These examples suggest that although these multi-domain signaling proteins may employ a variety of different structural mechanisms to acquire regulation, it is very likely that the energetic mechanism underlying the cooperative autoinhibition and the activation process are analogous to that of Vav.

Molecule	Function	Regulatory interactions
Vav	GEF	DH-helix: 10-fold core CH-PH: 10-fold modulatory Together: ~100-fold (Yu, Martins et al. 2010)
Tim	GEF	DH-helix: 50-fold core SH3: 3-fold modulatory Together: 75-fold (Yohe, Rossman et al. 2008)
SH-PTP2	Phosphatase	Mono-phospho ligand: 9-16 fold Di-phospho ligand: 37-fold (Pluskey, Wandless et al. 1995)
Abl	Kinase	SH3: 7-fold Myristoylation: 7.5-fold Together: 14-fold (Hantschel, Nagar et al. 2003)
FAK	Kinase	Deletion of FERM: 5-fold Phosphorylation: 20-fold (Lietha, Cai et al. 2007)
Src	Kinase	SH2 ligand: 2.7-fold SH2 and SH3 ligand: 10-fold (Alexandropoulos and Baltimore 1996)
Hck	Kinase	SH2 ligand: 2.5-fold SH2 and SH3 ligand: 6-fold (Moarefi, LaFevre-Bernt et al. 1997)

Table 3-1. Molecules with multiple weak interactions

The generality of Vav regulation mechanism could be assessed in these systems. The NMR-based approach we have used here allows for accurate quantitative determination of the population of different states across the coupled equilibria. Applications of this approach will reveal whether coupled equilibria of multiple weak interactions also exist in other multi-domain systems. Kinetic and thermodynamic studies will further allow us to understand if an ordered activation process also occurs.

Chapter 4

Introduction II

Discovered during 1940s in muscle, actin is one of the most abundant proteins on earth with high cellular concentrations. The dynamics of the actin cytoskeleton play a critical role in many essential cellular processes, including cell migration, division, adhesion, endo/exocytosis, cell polarity establishment and maintenance, and cellular morphogenesis. Tight temporal and spatial regulation of actin filament assembly and turn-over are essential for cell survival.

Spontaneous actin filament nucleation observed in vitro is too slow compared with rates observed in cell responses. To overcome the kinetic barrier of actin nucleation, cells express actin nucleating molecules including actin related protein 2 and 3 (Arp2/3) complex, Formin proteins and Wiskott-Aldrich Homology 2 (WH2) based nucleators such as Spire proteins etc. These actin nucleators are all thought to organize actin monomers into a minimal stable actin nucleus to promote polymerization. Although much progress has been made regarding the nature of the different types of actin nuclei, the detailed pathways

for nucleus formation and the structure of the minimal stable filament element remain to be explored.

During bacterial infection, many pathogenic effectors disrupt or hijack the host cell actin cytoskeleton to increase their pathogenicity or motility. Among them, VopL is an effector protein recently identified which directly targets actin filament assembly by functioning as an actin nucleator. It has been shown that VopL potently promotes the assembly of actin filaments *in vitro* which grow at their barbed ends. The mechanism of actin filament nucleation by VopL, whether it acts in the same manner as its eukaryotic counterparts or adopts a distinct mechanism, has been unknown.

The second part of this thesis (chapter 4-6) is devoted to investigating the actin filament assembly mechanism by the bacterial effector VopL. In this chapter, I will first review eukaryotic actin nucleators and their proposed nucleation mechanisms. Then I will introduce VopL and other pathogenic effectors that hijack the host cell actin cytoskeleton during infection. In chapter 5, I will present structural and functional studies, which have led us to a model where VopL stabilizes the formation of an actin nucleus involving short-pitch interactions.

Actin molecule

All eukaryotes have the genes for actin (Richards and Cavalier-Smith 2005), and non-muscle cells contain a high concentration of non-polymerized actin, estimated to be ~ 100μM in some cells (Bray and Thomas 1976). Actin-monomer binding proteins such as profilin and thymosin suppress spontaneous nucleation. Actin is an ATPase and binds to adenosine nucleotide triphosphate (ATP) or adenosine diphosphate (ADP). Binding to ADP or ATP is a critical factor in controlling actin filament dynamics and interactions between actin and other binding proteins (Pollard, Blanchoin et al. 2000).

Actin is a highly conserved 375 residue protein with four subdomains surrounding a nucleotide binding cleft. The first crystal structure of actin was solved in complex with DNase I in 1990, which was a major breakthrough (Kabsch, Mannherz et al. 1990). Since then over 40 actin crystal and cocrystal structures have been solved. Due to the strong tendency to polymerize at high concentrations, it has been difficult to crystallize uncomplexed actin in the monomeric state. In 2001, Dominguez and colleagues successfully crystallized the first uncomplexed actin in the ADP bound state through covalent modification of residue Cys 374 in subdomain 4 with tetramethylrhodamine (TMR) (Otterbein, Graceffa et al. 2001) (Fig. 4-1A). Two years later, the same group solved a monomeric actin structure in the AMPPNP bound state using the same method (Graceffa and Dominguez 2003). Based on the comparison of the two crystal

structures, Dominguez and colleagues proposed that ATP hydrolysis promotes a series of changes originating from the nucleotide binding cleft, which ends up with a Loop-to-Helix structural transition in the DNase binding loop (D-loop) in subdomain 2 of actin. However, the cleft between subdomain 2 and 4 was closed in both structures, opposing to the controversial view held by many that ATP hydrolysis results in opening of the cleft. It was suspected that TMR modification might stabilize the inactive conformation of actin and make it not sensitive to nucleotide binding (Rould, Wan et al. 2006). Recently, Rould et al. solved the monomeric actin structure in both ADP and ATP state using actin with two mutations in subdomain 4 (A204E/P243K, called "AP" actin) (Rould, Wan et al. 2006). These two structures revealed that the D-loop was unstructured in both crystals, but there are nucleotide dependent structural differences in other regions of actin including residues 70-78 (H-loop) and residues 11-16 (S-loop). The various structures of monomeric actin suggest that the molecule has the potential to adopt different conformations, and in the context of actin filament, with actin-actin contacts, the structure of individual actin subunits may also differ from the unpolymerized actin monomer.

Monomeric actin (also called "G actin" for globular) polymerizes into actin filament (also called "F actin" for filamentous) in a polar fashion with a fast growing end (barbed end) and slower growing end (pointed end). ATP-actin polymerizes faster and dissociates more slowly than ADP-actin at filament barbed

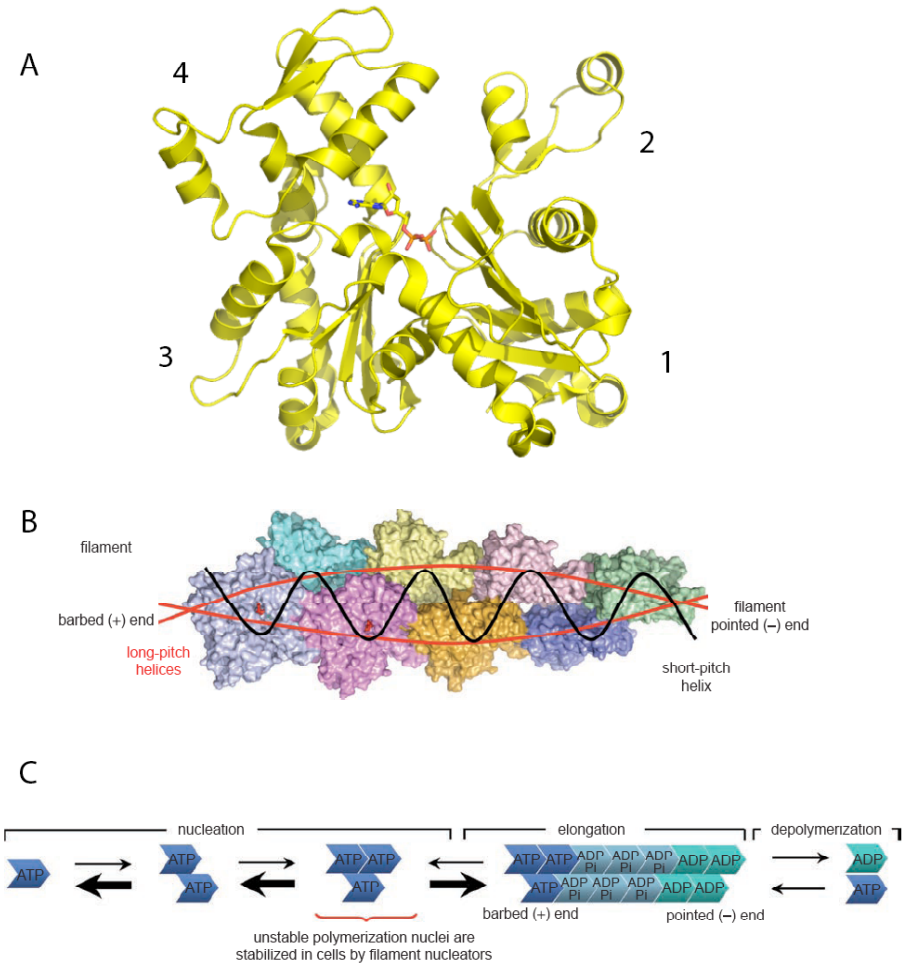


Fig. 4-1 Structure of monomeric actin and actin filament. (A) Structure of monomeric actin. Nucleotide binding pocket and subdomains 1-4 of actin are shown. (B) The structure of asymmetric actin filament. (C) Actin filament dynamics. Panels (B) and (C) are adapted from (Dominguez 2009).

end as well as at the pointed end (Pollard and Borisy 2003). Under physiological conditions, pure ATP-actin has a critical concentration of 0.1 μM at the barbed end and 0.6 μM at the pointed end (Pollard 1999). Once polymerized, ATP is hydrolyzed with a half time of 2 seconds (Pollard and Borisy 2003). Phosphate is dissociated from the filament slowly. As a result, the actin filament has three distinct regions: an ATP region near the barbed end, an intermediate ADP-Pi region, and an ADP region. It has been demonstrated that actin filaments with different nucleotide state have slightly different conformations (Belmont, Orlova et al. 1999).

Elongation of actin filaments is a rapid reaction. At the filament barbed end, elongation of ATP-actin is fast and diffusion limited (Drenckhahn 1986). At the filament pointed end, elongation is slower and not diffusion limited. With 10-100 μM of free actin, which is comparable to the physiological concentration of free actin in cells, the elongation rate of the filament barbed end has been measured to be as fast as roughly 100-1000 subunits per second, which equals 0.3-3 μm per second (Pollard, Blanchoin et al. 2000) In the cell cytoplasm, however, it is assumed that all of the unpolymerized actin monomers are bound to profilin or thymosin, which can lower the free actin concentration and alter the elongation rate (Pantaloni 1993; Kaiser, Vinson et al. 1999; Faraldo-Gomez and Roux 2007). Actin filaments turn over rapidly during multiple cellular processes including

pseudopod extension and centripetal flow (Wang 1985; Forscher and Smith 1988; Cassimeris, McNeill et al. 1990; Theriot and Mitchison 1991). In a continuously moving cell, actin filament assembly and disassembly must be balanced. The dissociation constants measured for ADP-actin is 7 s^{-1} at barbed end and 0.3 s^{-1} at pointed end. Actin-depolymerizing factor (ADF, also known as cofilin) is thought to bind to ADP-actin and promotes its dissociation from filament pointed end (Bamburg, Harris et al. 1980; Carlier, Laurent et al. 1997; Carlier and Pantaloni 1997; Blanchoin, Pollard et al. 2000). The mechanism of accelerated actin filament depolymerization by ADF/cofilin remains to be learned by additional work.

Under physiological salt conditions, purified actin polymerizes slowly in a concentration dependent manner. However, spontaneous nucleation of actin alone is highly inefficient and unfavorable (Fig. 4-1C). It involves the sequential formation of obligatory intermediates (actin dimer and trimer) that are very unstable and rapidly dissociate. Kinetic simulations from complete polymerization time courses have estimated the equilibrium constants to be: ~ 100,000 μM for actin dimers and ~ 10-100 μM for trimers (Pollard, Blanchoin et al. 2000; Sept and McCammon 2001). To overcome the energetic barrier of forming an actin nucleus, cells have developed three classes of factors to accelerate filament assembly: Arp2/3 complex, formin proteins and WASP

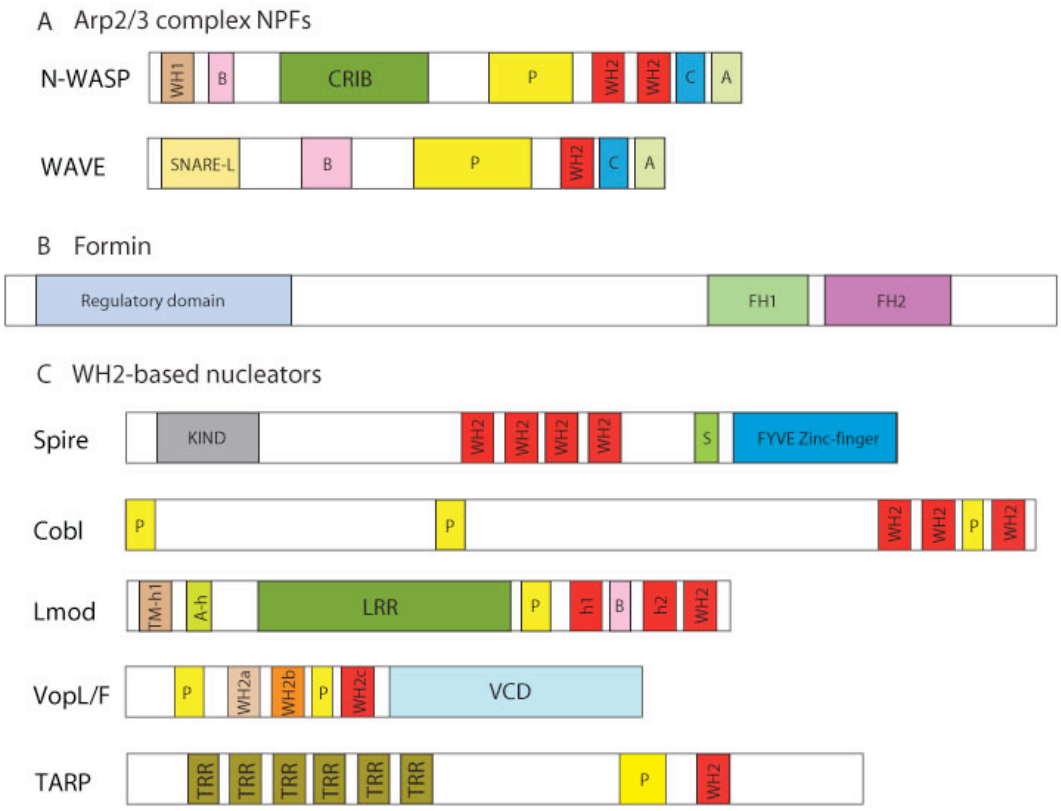


Fig. 4-2 Domain structures of actin nucleating factors. Abbreviations in the cartoon represent: CRIB: Cdc42 and Rac interactive binding domain, also known as GTPase binding domain (GBD), KIND: kinase non-catalytic C-lobe domain, SNARE-L: SNARE like motif, B: Basic, S: Spire like, LRR: Leucine rich region, TRR: Tyrosine rich repeat region.

Homology domain 2 (WH2)-based nucleators (Machesky, Atkinson et al. 1994; Pruyne, Evangelista et al. 2002; Quinlan, Heuser et al. 2005; Ahuja, Pinyol et al. 2007; Chereau, Boczkowska et al. 2008; Renault, Bugyi et al. 2008; Chesarone and Goode 2009). The domain structures of different types of actin nucleators are shown in Fig. 4-2. The proposed nucleation mechanisms are shown in fig. 4-3 and reviewed below.

Structure of the Actin Filament

In 1990, Holmes and colleagues determined the helical structure model of an actin filament by fitting the high resolution crystal structure of the actin monomer into a low resolution (8.4 Å) structure of actin filament determined from X-ray fiber diffraction (Holmes, Popp et al. 1990; Kabsch, Mannherz et al. 1990). In the model, actin filament can be seen as a single left handed short pitch (cross-filament) helix, each actin monomer is displaced by a rotation of 166° and a distance of 27.3 Å axially (Fig. 4-1B) (Holmes, Popp et al. 1990; Kabsch, Mannherz et al. 1990; Oda, Iwasa et al. 2009). The main contacts in the model are long pitch (along filament) interactions between subdomain 2 of one actin and subdomain 1 of another actin above, and between subdomain 4 of one actin and subdomain 3 of another actin above. Although this model requires the premise that there is no large conformational change in actin monomers, it has been

widely supported by a variety of mutagenesis and EM studies (Orlova and Egelman 1993; Bremer, Henn et al. 1994; Kim, Wriggers et al. 2000; Reisler and Egelman 2007). Recently, an improved F-actin model was reported by X-ray fiber diffraction analysis (Oda, Iwasa et al. 2009). In this model, contacts between subdomain 3/4 and subdomain 3/2 are completed by flattening of actin subunit, which enable the contacts between two protofilaments. Importantly, the contacts between subdomain 3 and subdomain 4 seem to contribute to the core of the filament. Oda et al. proposed that the rotations of subdomains 1 and 2 relative to subdomains 3 and 4 flatten the actin molecule in the structure. Following this study, Murakami et al. performed cryo-EM analysis on F-actin, and argued that a conserved proline-rich loop in subdomain 4 plays a critical role in coupling actin filament assembly with ATP hydrolysis (Murakami, Yasunaga et al. 2010).

Multiple studies have also revealed that actin filaments have considerable plasticity. It has been demonstrated that although the actin subunits have a fixed axial rise of 27.3 Å in the filament, the rotation angle between adjacent actin subunits can be variable (Egelman, Francis et al. 1982). In addition, EM studies have revealed that newly polymerized actin filaments appear to be more disordered, whereas older filaments rearrange into a more ordered and canonical helical structure (Steinmetz, Goldie et al. 1997; Schoenenberger, Bischler et al. 2002; Orlova, Shvetsov et al. 2004). ADF/cofilin binding to filament also causes the disordered structure to reappear (Galkin, Orlova et al. 2001). Single filament

fluorescence imaging studies further indicated that actin filaments exist in two populations with dramatically different shrinkage rates (Kueh, Brieher et al. 2008). Newly polymerized filaments shrink rapidly, whereas older filaments are more stable with slower shrinkage rate, likely due to a more ordered structure. Cofilin treatment accelerated filament shrinkage at both barbed end and pointed end. The intrinsic structural plasticity of the actin filament, in addition to ATP hydrolysis, thus are critical factors in determining actin dynamics. High resolution structures of actin filaments in different states will be important to understanding the dynamic properties of actin filaments with vastly different turnover rates in cells.

Actin Nucleating factors

Arp2/3 complex

The first identified actin nucleating factor is actin related protein 2/3 (Arp2/3) complex (Mullins, Heuser et al. 1998; Mullins, Kelleher et al. 1998). Arp2/3 complex is active at the leading edge of motile cells, and initiates the formation of filament branches on the sides of existing mother actin filaments (Pollard 2007). Branched filaments push cells forward to generate protrusions

which are important for cellular events such as cell migration. After nucleation, Arp2/3 complex remains attached to the mother filament and therefore is incorporated into the newly formed filament branch (Mullins, Heuser et al. 1998). The pointed end of the daughter filament nucleated by Arp2/3 complex is anchored to the mother filament, and the barbed end grows away at an average of 70° angle from the mother filament.

Arp2/3 complex consists of seven subunits: Arp2, Arp3, which are actin related proteins and structurally similar to actin monomer, ARPC1 (40 kD), ARPC2 (35 kD), ARPC3 (21 kD), ARPC4 (20 kD), and ARPC5 (16 kD). It has been proposed that during nucleation, Arp2 and Arp3 adopt a conformation that is similar to a short-pitch cross filament actin dimer, and generate a trimeric nucleus by binding to one actin monomer (Robinson, Turbedsky et al. 2001). The first crystal structure of Arp2/3 complex was solved in 2001 by Pollard and colleagues. In this structure Arp2 and Arp3 are separate from each other and subdomains 1 and 2 of Arp2 are disordered (Robinson, Turbedsky et al. 2001). It was claimed that this represented an inactive conformation of Arp2/3 complex, and a substantial conformational change will need to occur to bring the two subunits together during nucleation. Later, Nolen et al. crystallized Arp2/3 complex in several different nucleotide bound states (Nolen and Pollard 2007). These structures revealed a subtle stabilization of the Arp2 structure, but Arp2 and Arp3 were still separate. Thus it seems nucleotide binding is not sufficient to

activate Arp2/3 complex. Although there has been no high resolution structure of Arp2/3 complex in the active conformation, EM analysis performed by Egile et al. supported the idea that the Arp2 and Arp3 are the first two subunits in the daughter actin filament (Egile, Rouiller et al. 2005).

Arp2/3 complex is a highly inefficient actin nucleator by itself (Mullins, Heuser et al. 1998). Nucleation is activated by nucleation promoting factors (NPFs), of which Wiskott-Aldrich Syndrome Protein (WASP) family proteins are the best known and most extensively studied (Goley, Ohkawa et al. 2006; Pollard 2007). Classical WASP family proteins include WASP, N-WASP (Neuronal-WASP), WAVE (WASP-family VErprolin homolog) (Goley, Ohkawa et al. 2006), WASH (WASP/Scar homolog) (Linardopoulou, Parghi et al. 2007), WHAMM (WASP Homolog Associated with actin, Membrane and Microtubules) (Campellone, Webb et al. 2008) and JMY (Junction-Mediating regulatorY) (Zuchero, Coutts et al. 2009) proteins. NPFs are thought to recruit actin monomers and promote a conformational change in Arp2/3 complex (Pollard 2007). They are regulated themselves by a variety of signaling molecules particularly by Rho-family GTPases like Cdc42 and Rac (Bompard and Caron 2004) (Eden, Rohatgi et al. 2002; Goley, Ohkawa et al. 2006; Chen, Borek et al. 2010). WASP family proteins all contain a C-terminal VCA domain (Verprolin homology region also called “WH2” for Wiskott-Aldrich Homology 2, Central

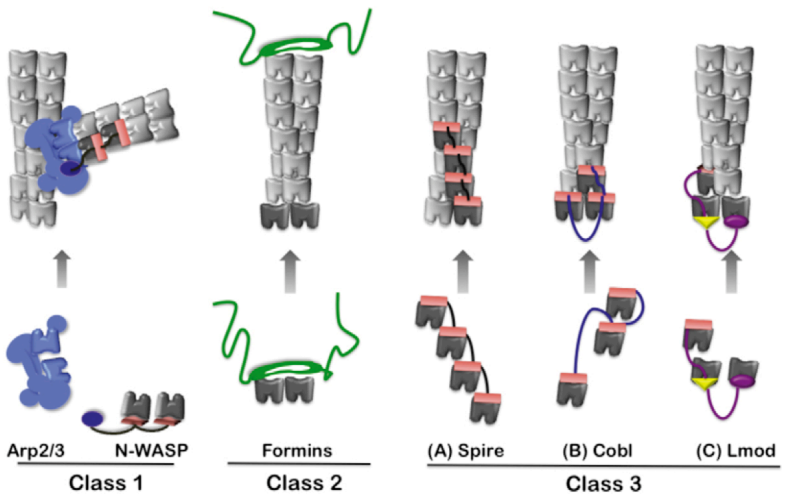


Fig. 4-3 Proposed mechanism of actin assembly factors (adapted from Chesarone, 2009 #129) Three classes of actin nucleators are shown. The actin subunits used by nucleators to seed polymerization are displayed in black and actin subunits polymerized from nuclei are displayed in gray. Class I actin nucleator: Arp2/3 complex (N-WASP uses its WH2 domain to recruit actin monomers and its acidic domain to bind to an actin-related protein subunit of Arp2/3. This structure stabilized by N-WASP may mimic an actin trimer). Class II: formin proteins (Spontaneously formed actin dimers and/or trimers are stabilized. Formins remain associated with the barbed end while permitting addition of actin subunits). Class III: WH2 based nucleators (An actin nucleus stabilized by lateral and/or longitudinal contacts). Note, in some respects N-WASP represents a specialized form of Class III nucleator, in which the third actin monomer-binding domain has been replaced with a domain that binds to actin-related proteins.

hydrophobic or connecting region, Acidic region), the shortest polypeptide necessary for activating nucleation with Arp2/3 complex. In the VCA, the V region recruits the first actin subunits in the new filament, and the C and A motifs interact with Arp2/3 complex and help to stabilize the active conformation. It is now believed that nucleotide binding, an actin loaded VCA motif and existing mother filament all contribute to the activation of Arp2/3 complex (Pollard 2007). However, the mechanism and interactions involved in the activation process remain to be investigated.

Formin proteins

In 2002, formin protein was identified by the Boone laboratory to be an actin nucleator that associated with the growing barbed end of the actin filament (Evangelista, Pruyne et al. 2002; Pruyne, Evangelista et al. 2002). Formin proteins are present in fungi, plants and animals (Pruyne, Evangelista et al. 2002). They nucleate unbranched actin filaments, which are often assembled into bundles, and are important for the formation of multiple subcellular structures, including stress fibers, contractile rings during cytokinesis, actin cables that direct polarized growth and mating (Chang, Drubin et al. 1997; Petersen, Nielsen et al. 1998; Feierbach and Chang 2001). The processive association of formin with the

filament barbed end protects the filament from capping proteins and allows for continuous filament growth (Kovar and Pollard 2004; Moseley, Sagot et al. 2004). It has been documented to date that fission yeast has three isoforms of formin proteins with distinct functions, and vertebrates have fifteen formin isoforms (Higgs 2005).

Formin proteins consist of an N terminal regulatory domain which binds to GTPases (not present in metazoan Formin proteins INF1, Delphilin, FHOD1, FMN1, but present in other metazoan and non-metazoan Formins), a proline rich formin homology 1 (FH1) domain that binds to profilin, a formin homology 2 (FH2) domain and a C terminal region (Pollard 2007). The filament nucleating activity of formin proteins resides in the FH2 domain. Deletion of the FH2 domain inhibits activity of formins in cells (Moseley, Sagot et al. 2004), and when purified and mixed with actin monomers, the isolated domain can nucleate actin filament assembly (Pruyne, Evangelista et al. 2002; Sagot, Rodal et al. 2002; Xu, Moseley et al. 2004). Crystal structures of FH2 domains have revealed that the domain forms a donut shaped head-to-tail dimer (Xu, Moseley et al. 2004) (Lu, Meng et al. 2007). It has been proposed that during nucleation, the FH2 domain stabilizes spontaneously formed actin dimer that acts as the seed for new filament. Although direct evidence is still lacking, this idea is supported by a crystal structure of FH2 dimer complexed with two actin molecules solved by Otomo et al. which shows one FH2 dimer bridging three actin subunits (Otomo, Tomchick

et al. 2005). After nucleation, unlike Arp2/3 complex, the FH2 domain processively moves with the growing barbed end of the filament (Kovar 2006; Vavylonis, Kovar et al. 2006). It was suggested that alternating contacts between the two halves of the FH2 dimer with the filament barbed end could account for this processive activity (Otomo, Tomchick et al. 2005). However, the precise mechanism of formin proteins' processive movement on the filament barbed end, in particular the conformation of formin proteins associated with filament ends instead of actin monomers, remains to be explored by EM or single molecule imaging techniques.

The crystal structure of mouse Formin protein mDia1 shows that its N terminus dimerizes and interacts with a segment distal to the FH2 domain (Otomo, Otomo et al. 2005). The molecule is therefore autoinhibited through intramolecular interactions. Rho-GTPases bind to the N terminus of formin proteins, displace the C terminus and activate its nucleating activity (Pollard 2007). The FH1 domain of formins appears to be a long and unstructured region, which could bind to profilin-actin monomers and deliver the actin to FH2 capped filament barbed end for elongation. This event was suggested to lead to the acceleration of filament elongation by up to five-fold (Kovar, Harris et al. 2006). The enhancement of elongation rate was shown to be dependent on the presence of profilin, and the number of profilin binding sites on FH1 domain (Romero, Le Clainche et al. 2004). Kinetic modeling studies explained how profilin and formin

FH1 and FH2 domains work together to accelerate filament assembly (Vavylonis, Kovar et al. 2006).

WH2 Based nucleators: Spire, Cobl, Lmod (Fig. 4-3) (Comparison in Table 4-1)

Since 2005, a new class of actin nucleators has been emerging, which include Spire, Cordon-Bleu (Cobl), and Leiomodin (Lmod). They all contain tandem repeats of WH2 domains for actin nucleation but appear to serve distinct cell functions. It has been recently understood that nucleation by these WH2 based nucleators involves the stabilization of an actin nucleus, which can be an actin trimer or tetramer (Dominguez 2009) by WH2 domains. A summarized comparison of WH2 based nucleators is shown in Table 4-1.

WH2 domain is a short actin binding motif with 17-23 amino acids that are poorly conserved (Dominguez 2007), making it hard to identify through sequence analysis. It is abundant among cytoskeletal proteins, has variable length and sequence, and has diverse functions. It is present in the actin sequestering protein thymosin, nucleation promoting factors like WASPs, nucleators like Spire and Cobl, the elongation factor VAsodilator-Stimulated Phosphoprotein (VASP), and scaffolding proteins like Missing In Metastasis protein (MIM). In 2005, Chereau et al. determined the crystal structure of WH2 domain bound actin, which

	Minimal Actin nucleus	Potency (The concentration used to get fast activation)	Location at nucleated filament	Binding partner or regulator	Other functions	Linker length dependence
Spire	Longitudinal tetramer	500 nM	Pointed end (Quinlan, Heuser et al. 2005) Barbed end (Bosch, Le et al. 2007)	Formin/ profilin	Sequester actin monomers Sever filament	No
Cobl	Short-pitch trimer	40 nM	Barbed end	Profilin?	Sequester actin monomers	Yes
Lmod	Short-pitch trimer	25 nM	Pointed end	Tmod	unclear	Unknown

Table 4-1 Comparison of WH2 based nucleators

revealed that the N terminal region of the WH2 domain forms an amphiphilic α helix and binds to a barbed face cleft between actin subdomains 1 and 3, with a hydrophobic surface facing to the actin cleft (Chereau, Kerff et al. 2005, see Fig. 4-4). This helix was shown to contribute most of the binding affinity between WH2 domain and actin. Following the helix is a linker region with a conserved "LKKT" motif that binds to a hydrophobic pocket of actin. Mutating the first L residue in LKKT motif to A or P of thymosin WH2 like domain decreased its affinity to actin substantially, and the relative binding activity between the mutant and wildtype was measured to be less than 6 percent in crosslinking experiment (Simenel, Van Troys et al. 2000). WH2 domains are divided into two groups: short WH2s and long WH2s. Short WH2 domains have only the helix and linker region, and long WH2 domains have 10 or more amino acids at the C terminus of LKKT motif (Dominguez 2007). The WH2 domain of thymosin is different from others in that it contains an extended region that binds to the nucleotide cleft of actin, and a second helix that binds between subdomain 2 and subdomain 4 at the pointed end of actin (Irobi, Aguda et al. 2004). It also has the highest affinity to actin.

During filament nucleation, Bosch et al. claimed that a steric clash will occur between WH2 domain and the D-loop of actin along the long pitch of the filament, therefore WH2 domain would have to dissociate for continuous growth (Bosch,

Le et al. 2007). In 2008, X-ray scattering studies performed by Rebowski et al. showed WH2 bound actin arrays along the long pitch of actin filaments, indicating the WH2 can remain bound to actin during filament growth without steric problems (Rebowski, Boczkowska et al. 2008). However, two years later the same group reported a crystal structure of actin dimer assembled by tandem WH2 domains, which showed differences between the short pitch actin dimer in this complex structure and in F-actin. The authors concluded that WH2 bound actin generated the steric clashes for filament contacts, and consequently the WH2 domain would have to dissociate from actin to allow filament growth (Rebowski, Namgoong et al. 2010). There are two potential explanations for the consistencies between these two studies: 1. The low-resolution quality of X-ray scattering experiments may cause mis-interpretations of the data; 2. F-actin can adopt multiple conformations, and WH2 domains can bind to F-actin only at some specific state.

Spire: Spire was the first identified WH2 based nucleator which could weakly nucleate new actin filaments *in vitro* (Quinlan, Heuser et al. 2005). It is present in Drosophila and mammals with preferential expression in nervous systems (Schumacher, Borawski et al. 2004). Spire is required for the establishment of oocyte polarity axes in Drosophila. A mutation in the drosophila *spire* gene

resulted in disruption of normal nervous system development (Manseau and Schupbach 1989). Spire is a multi-domain protein with four central WH2 motifs, and therefore can bind to four actin monomers. Quinlan et al. showed that Spire nucleated actin at a rate similar to formin proteins but slower than Arp2/3 complex, and remained associated with the filament pointed end (Quinlan, Heuser et al. 2005). The nucleation activity of Spire is dependent on the presence of multiple WH2 motifs, and the linker region between the third and fourth WH2 also appears to be critical (Quinlan, Heuser et al. 2005). EM studies demonstrated that Spire and actin could form a stable complex, with an elongated rod shape, and it has been proposed that Spire nucleates by organizing four actin subunits in a linear fashion to generate a filament seed (Quinlan, Heuser et al. 2005). This idea is supported by the fact that the linker region between the WH2s is too short to wrap around two cross filament actin subunits. However, although attempts have been made to determine the crystal structure of spire-actin complex, direct evidence for this model is still lacking (Ducka, Joel et al. 2010).

Spire was shown to work together with formin proteins and requires the presence of formin and profilin proteins for massive filament assembly during mid-oogenesis of drosophila (Manseau and Schupbach 1989; Quinlan and Kerkhoff 2008; Renault, Bugyi et al. 2008). It has been shown that Spire and the formin protein Cappuccino synergize to nucleate actin filaments both *in vivo* and *in vitro* (Bosch, Le et al. 2007) (Rosales-Nieves, Johndrow et al. 2006).

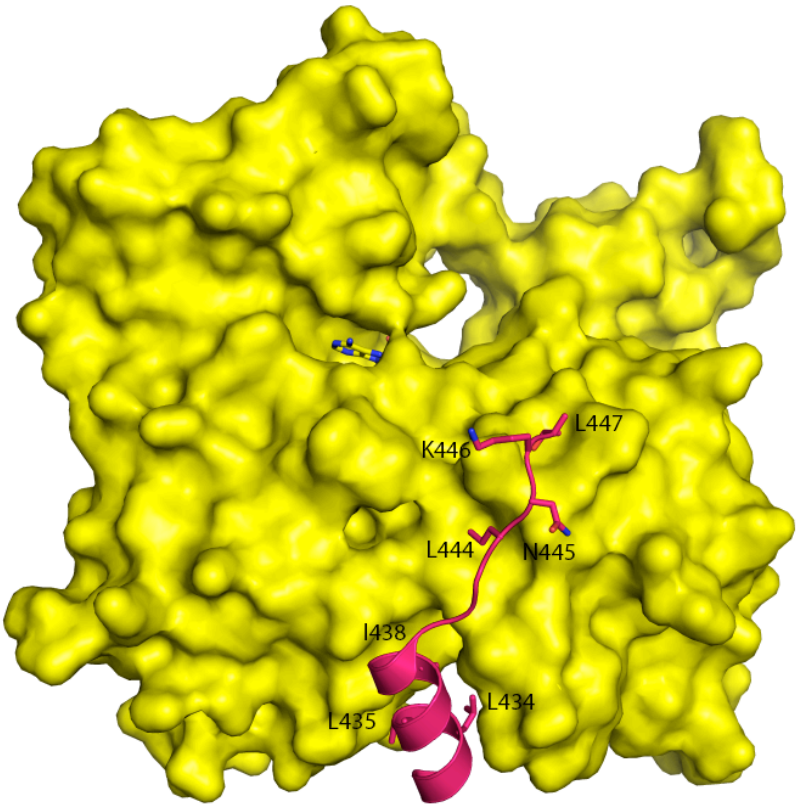


Fig. 4-4 Structure of WASP WH2 bound actin. WASP WH2 is shown in red, and actin is shown in yellow. Residues L444, N445, K446, L447 represent the "LKKT" motif.

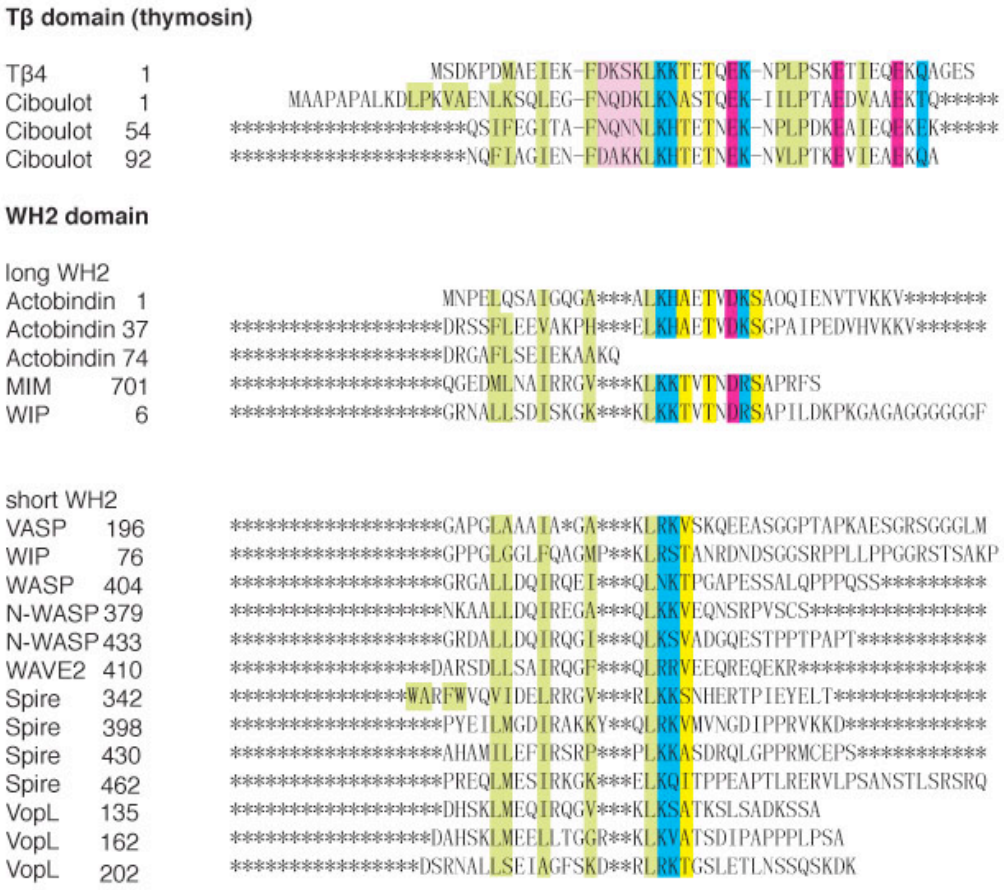


Fig. 4-5 Sequence alignment of WH2 domains

Quinlan et al. identified that Spire interacts with formin FH2 domain through the kinase non-catalytic C-lobe domain (KIND) at its N terminus (Quinlan, Hilgert et al. 2007). The mechanism of their cooperation for filament assembly is unclear. WASH protein, which is a NPF of Arp2/3 complex, was also indicated to synergize with Spire to control actin dynamics with mechanistic details remaining to be determined (Liu, Abreu-Blanco et al. 2009).

The functional role of Spire *in vivo* has nevertheless been under debate (Quinlan, Heuser et al. 2005; Bosch, Le et al. 2007). Bosch et al. observed that Spire slowed down filament barbed end depolymerization, and the length of actin filaments became shorter upon the addition of Spire. The authors claimed that Spire functions by severing and capping actin filaments instead of acting as an actin nucleator. In addition, they observed that Spire blocked filament barbed ends instead of pointed ends as seen by Quinlan et al. The mechanism of spire mediated actin dynamics appears to be far more complex than expected, and additional work is still required to resolve these discrepancies,

Cobl: Cobl is a vertebrate actin nucleator that controls neuronal morphology and development (Ahuja, Pinyol et al. 2007). It is enriched in brain and promotes the assembly of nonbundled, unbranched actin filaments. Cobl consists of three WH2 domains at its C-terminal region, with its nucleation activity dependent on all

three WH2 domains and an extended linker between the second and third WH2. The activity of Cobl in vitro was found to be comparable to the same concentration of N-WASP activated Arp2/3 complex and substantially higher than Spire proteins. Ahuja et al. proposed that the nucleation power of Cobl lies in the assembly of an actin trimer in a cross-filament orientation (Ahuja, Pinyol et al. 2007). This idea is supported by actin assembly assays that showed shortening the linker between the second and the third WH2 domain abolished Cobl activity. In addition, consistent with the model, the length of the linker instead of the sequence was found to be critical for Cobl-mediated nucleation. A crystal structure of Cobl complexed with actin will be particularly important to reveal the conformation of a minimal actin nucleus.

Interestingly, Ahuja et al. claimed that different from Spire, Cobl remains attached to filament barbed end since it protects the barbed end from depolymerization but not the pointed end. The locations of Cobl on the newly formed or growing filament remain to be validated in the future. Furthermore, the regulation of Cobl inside the cell is worth investigation.

Lmod: Lmod, identified recently by the Dominguez laboratory, is an actin nucleator expressed almost exclusively in muscle cells (Chereau, Boczkowska et al. 2008). Lmod knockdown cells lack organized sarcomeres, indicating Lmod

plays a role in sarcomere assembly and organization (Chereau, Boczkowska et al. 2008). Lmod is related to the filament pointed end capping protein tropomodulin (Tmod), with its N-terminal 340 residues sharing 40% identity with Tmod (Fischer and Fowler 2003; Kostyukova 2007). In this homologous region, there is an N-terminal unstructured region and a C-terminal leucine rich repeat, with two actin binding motifs in total (Krieger, Kostyukova et al. 2002). The important difference between the two proteins is Lmod has a C-terminal 150 residue extension that contains a poly-proline region, two helical motifs, and one WH2 domain. Lmod was shown to be a potent actin nucleator *in vitro* (Chereau, Boczkowska et al. 2008). The minimal fragment of Lmod which bears the nucleating activity contains both the leucine rich repeat and the C-terminal extension. The N terminal region was shown to recruit tropomyosin and increase the nucleation activity by three-fold. With three actin binding sites, Lmod has the potential to recruit three actin subunits. It was therefore proposed that Lmod nucleates by stabilizing the conformation of an actin trimer, also in a cross filament orientation like Cobl. Chereau et al. also suggested Lmod binds to filament pointed end by the homologous region with Tmod, and tropomyosin increases its nucleating activity by stabilizing it at the pointed end (Chereau, Boczkowska et al. 2008). The nucleation mechanism by Lmod remains to be validated by further stoichiometry and structural studies.

Taken together, the three types of WH2 based actin nucleators are utilized by cells to generate different kinds of actin structures and are individually important for specific cellular processes. Some nucleators may bear multiple functions other than promoting actin filament assembly. The actin nucleation mechanism by each nucleator appears to be similar in the sense of stabilizing the formation of a specific actin nucleus, which can be a long-pitch actin tetramer, or a short-pitch actin trimer. In addition, organizing actin nucleus in a short pitch orientation looks to boost the activity of the nucleator, i.e., Cobl and Lmod have much higher potency than Spire protein. This phenomenon can be explained by that after forming the long pitch actin dimer (the first energetic barrier), stabilizing a short pitch actin dimer helps to overcome the second significant energetic barrier to allow filament growth. There are still inconsistencies concerning the locations of these nucleators on an actin filament, and extra work will be needed to reconcile the different observations.

Pathogenic Actin nucleators: VopL, VopF, TARP

During bacterial infection, pathogens use a variety of mechanisms to manipulate the host cell cytoskeleton in order to prevent or induce phagocytosis or to increase their motility. Effector proteins take control of the actin cytoskeleton at different levels of its signaling pathway. For example, bacteria

Salmonella enters host cells by altering their actin cytoskeleton and inducing the formation of membrane ruffles. The effector protein "SopE" is secreted and inserted into the cytosol of the host cell. It acts as a strong guanine nucleotide exchange factor for Rac1 and Cdc42, two GTPases in the Rho family that regulate actin cytoskeleton (Wood, Rosqvist et al. 1996; Fu and Galan 1999). Another well known example is the ActA protein that is secreted from bacteria *Listeria monocytogenes*. It mimics the eukaryotic NPFs and hijacks Arp2/3 complex for actin polymerization (Cossart and Toledo-Arana 2008). Two other effector proteins: *Rickettsia* RickA also employ similar mechanisms as ActaA to control host cell actin cytoskeleton (Gouin, Welch et al. 2005; Rottner, Stradal et al. 2005).

It has recently been shown that bacterial effectors act by directly targeting nucleating actin filaments as well. Three such proteins have been identified to date, including Vibio Outer Protein L (VopL), Vibio Outer Protein F (VopF) and Translocated Actin-Recruiting Phosphoprotein (TARP) (Jewett, Fischer et al. 2006; Liverman, Cheng et al. 2007; Tam, Serruto et al. 2007; Tam, Suzuki et al. 2010). These three proteins were all shown to make use of multiple WH2 domains and have nucleating activity independent of the eukaryotic factors. Detailed studies done on each of these three proteins are reviewed below.

VopL: VopL is a pathogenic effector of the gram negative bacteria *Vibrio parahaemalyticus* (*V. para*). *V. para* is a rod shaped bacteria that is often associated with infections after consumption of raw seafood. It can cause gastro-intestinal diseases, diarrhea, and wound healing infections. It has been demonstrated that VopL is secreted from *V.para* through a type III secretion system (T3SS) (Liverman, Cheng et al. 2007). VopL consists of an N-terminal unstructured region (residues 1-114) including multiple proline residues, a central region with three WH2 domains (residues 115-245), and a C-terminal uncharacterized domain (residues 246-484, named "VCD" from now on). VopL was found to induce the formation of actin stress fibers during infection, and it potently promotes actin filament assembly in vitro, at a rate much faster than N-WASP activated Arp2/3 complex at similar concentrations (Liverman, Cheng et al. 2007). VopL mediated actin nucleation is dependent on the presence of WH2 domains, whose mutation result in activity loss both in vivo and in vitro. The filaments nucleated by VopL grow at their barbed end, since barbed end capping proteins were shown to inhibit their growth.

The nucleation mechanism of VopL accounting for its strong activity remains to be determined. Unlike the eukaryotic nucleators Cobl and Lmod, the linker region between the second and the third WH2 domain contains only ~ 19 residues, which is not long enough to wrap around two short pitch actin subunits. In addition, the function of VopL does not completely rely on three WH2 domains,

considering the fact that a dominant negative mutation on a single WH2 domain does not abolish VopL activity. These observations indicate VopL may nucleate actin filaments through a different mechanism from Cobl and Lmod. It would be therefore very interesting to understand the mechanistic basis of VopL mediataed filament assembly, both from a scientific and pharmalogical view.

VopF: Almost at the same time as VopL, VopF was identified to be secreted from bacteria *Vibrio chloerae* also through a T3SS system (Tam, Serruto et al. 2007). It was shown that VopF-mediated actin nucleation is important for intestinal colonization and VopF displays significant nucleating activity in vitro. The protein shares 55% amino acid sequence identity with VopL and contains three WH2 domains and VCD as well, suggesting the nucleation mechanism by VopF might be very similar to that by VopL.

Despite the great similarities between VopL and VopF, the two proteins have significantly different properties in several aspects. First, VopL promotes the formation of actin stress fibers whereas VopF induces aberrant actin-rich protrusions. How different actin structures are generated by them is an open question. Second, VopL localizes along stress fibers, in contrast with VopF that localizes to the tip of actin protrusions, suggesting VopF binds to the filament

barbed end. The locations of the two proteins at the filament are worth further investigation.

TARP: TARP is secreted from another gram-negative bacteria *Chlamydia trachomatis* through a T3SS system (Clifton, Fields et al. 2004). *Chlamydia trachomatis* is shown to be the leading cause of preventable blindness disease worldwide. It is also the most prevalent bacterial pathogen considered to cause sexually transmitted disease in the western world (Schachter 1999). TARP was found to be present in all pathogenic *Chlamydia* species to date and was implicated in the recruitment of actin to the site of bacteria internalization (Clifton, Fields et al. 2004; Clifton, Dooley et al. 2005). It was shown that TARP promotes actin assembly *in vitro* (Jewett, Fischer et al. 2006). Unlike VopL/VopF, TARP only contains one WH2 domain at its C-terminus. However, its nucleating activity was found to be dependent on its central poly-proline region, which seems to promote TARP oligomerization (Jewett, Fischer et al. 2006). Further studies are still needed to confirm the oligomerization state of full length TARP, and whether oligomerization is required for its function.

As described above, bacterial pathogens have evolved powerful nucleators that abuse WH2 domain mediated actin nucleation. With strong affinities and activities, these molecules can be very useful tools in another sense and allow us

to understand the molecular mechanism of actin assembly by repeats of WH2 domains. Understanding how pathogens interfere with host cells can also help with strategies to fight against bacterial infections. With these motivations in mind, I performed research on the pathogenic effector VopL as a model system, undertaking both structural and biochemical studies with detailed descriptions in chapter 5.

Chapter 5

Mechanism of Actin Filament Nucleation by the Bacterial Effector VopL[3]

Based on the published literature, VopL appears to be more potent than any other actin nucleator in promoting actin assembly. It could rapidly nucleate actin polymerization at a concentration as low as 5 nM. A unique feature of VopL/VopF protein is an uncharacterized domain at its C terminus (residue 246-484) that lacks homology to any known protein. We name this domain VopL C-terminal Domain (VCD). The VCD is predicted to be structured and mainly α helical based on secondary structure prediction analysis. More studies are obviously required in order to understand the role VCD is playing in nucleating actin. To elucidate the mechanism of actin nucleation by VopL, we are interested in what type of actin nucleus is formed during nucleation- is it an actin trimer or tetramer or something else? Moreover, we are most interested in understanding the nucleation problem structurally. Specifically, we want to know the conformation of VopL protein itself at first, and more importantly, we want to

[3] This chapter is mainly adapted from the work submitted to Nature Structure and Molecular Biology.

know the conformation of VopL bound to actin, which could tell us the specific interactions that stabilize an actin nucleus, and the interactions needed for the nascent actin filament to grow.

During the second half of my thesis work, I focused on determining the structure of VopL VCD alone and performing biochemical analyses to uncover the mechanism of actin filament nucleation by VopL. This work provided a structural basis for VopL mediated actin assembly, and led us to a model that is likely to be general among other WH2 based actin nucleators.

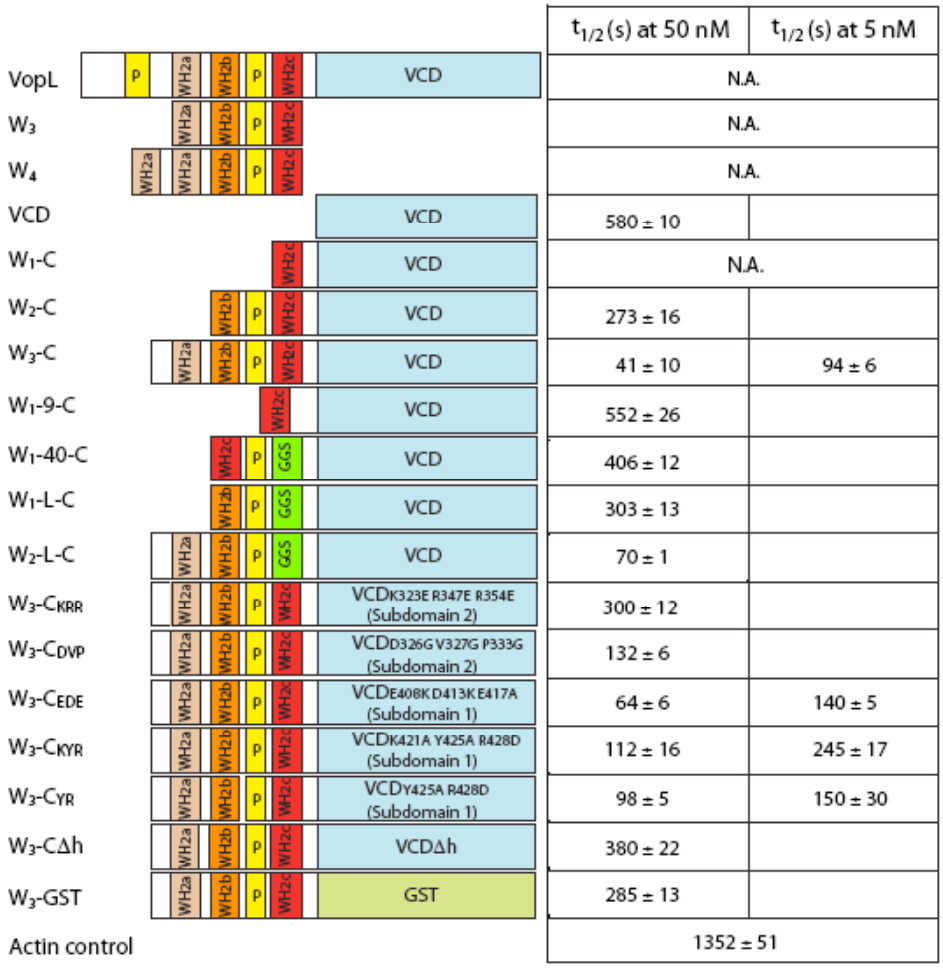
Results

WH2 motifs and VCD are both required for actin nucleation

VopL consists of an N terminal unstructured region that is likely a localization signal peptide for VopL (Liverman, Cheng et al. 2007), a central region containing two poly-proline regions (P) and three WH2 domains, and a C terminal VCD (Fig. 5-1). Sequence alignment shows that each WH2 domain in VopL contains both the N-terminal hydrophobic residues and C-terminal basic sequence that are important for binding to actin monomers in eukaryotic WH2 motifs (Chereau, Kerff et al. 2005). To test whether these three WH2 motifs all bind to actin with similar affinities, I used isothermal titration calorimetry (ITC)

method to measure the binding affinity of each motif (named WH2a, WH2b and WH2c from N- to C-terminus, Fig. 5-1) for monomeric actin (Fig. 5-2). The dissociation equilibrium constants (K_D) range from 5 to 15 nM, explaining the nM potency of VopL in actin assembly assays (Liverman, Cheng et al. 2007), and indicating little difference between the ability of the motifs to bind actin. Note that these assays were performed in low salt buffer to prevent actin polymerization (see Materials and Methods), and the affinities could be somewhat different in the presence of higher KCl and $MgCl_2$ concentrations.

Dominant negative mutations in all three WH2 domains results in abolishment of VopL nucleating activity, indicating its function highly relies on the presence of WH2 domain (Liverman, Cheng et al. 2007). Linear arrays of just WH2 motifs in other nucleation factors such as Spire and Cobl have substantial activity in vitro. I tested the activity of a similar VopL fragment containing only the three WH2 motifs without the VCD (construct W_3). Unlike Spire or Cobl, it was completely inactive in pyrene-actin assembly assays (Fig. 5-3, half times of actin assembly ($t_{1/2}$) are tabulated in Fig. 5-1). High concentrations of W_3 inhibited assembly, likely by sequestration of actin monomers (the analysis of actin sequestration by W_3 is shown in Fig. 5-4). By contrast, the VCD alone retained some actin assembly activity, albeit with ~100-fold lower potency than W_3-C, which contains the three WH2 motifs and the VCD (Figs. 5-1, 5-5; 500 nM VCD has activity similar to 5 nM W_3-C (not shown)). These results indicate that



	$t_{1/2}$ (s) at 50 nM	$t_{1/2}$ (s) at 5 nM
VopL	N.A.	
W_3	N.A.	
W_4	N.A.	
VCD	580 ± 10	
W_1-C	N.A.	
W_2-C	273 ± 16	
W_3-C	41 ± 10	94 ± 6
W_1-9-C	552 ± 26	
W_1-40-C	406 ± 12	
W_1-L-C	303 ± 13	
W_2-L-C	70 ± 1	
W_3-C_{KRR}	300 ± 12	
W_3-C_{DVP}	132 ± 6	
W_3-C_{EDE}	64 ± 6	140 ± 5
W_3-C_{KYR}	112 ± 16	245 ± 17
W_3-C_{YR}	98 ± 5	150 ± 30
W_3-CΔh	380 ± 22	
W_3-GST	285 ± 13	
Actin control	1352 ± 51	

Fig. 5-1 Domain structure of VopL constructs used and activities of VopL proteins represented by half times. P: poly proline region. WH2: WASP Homology 2 domain. VCD: VopL C-terminal domain. GGS: flexible linker region consisting of mainly G and S residues. Half times are measured from pyrene actin polymerization assays. Half-times are represented by an average ± standard error calculated from three measurements.

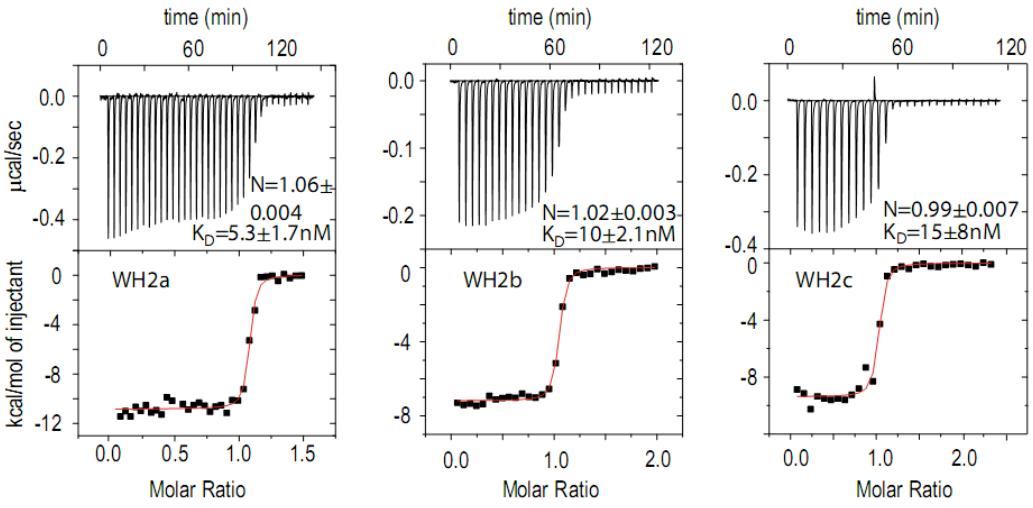


Fig. 5-2 Isothermal titration calorimetry analysis of the binding of the individual VopL WH2 domains to actin in G-buffer (see methods and materials). Upper and lower panel show raw and integrated heats of injection. Red lines in lower panels represent fits of the data to a single site binding model.

VCD in VopL is critical for nucleating actin. To further investigate the number of WH2 motifs required for nucleation, I compared the activity of constructs with zero, one, two or three WH2 motifs (Figs. 5-1, 5-5 constructs VCD, W_1-C, W_2-C, W_3-C, respectively). In this series, activity decreased in the following order: W_3-C > W_2-C >> VCD > W_1-C. Thus, proteins containing three or two WH2 motifs had appreciably higher activity than those with fewer. A single WH2 motif appears to inhibit activity relative to the bare VCD (see below). Although these activities could in principle reflect differences in rates of actin filament elongation (e.g. by binding of the WH2 motifs to the filament barbed end (Co, Wong et al. 2007), single filament elongation rates determined by Dominguez and colleagues are unchanged by VopL (Suk et al., submitted). Thus, the differences in activity likely reflect differences in nucleation efficiency. Together, these data demonstrate that actin binding by the WH2 motifs alone is insufficient to promote nucleation, and both the WH2 motifs and VCD play important roles in VopL function.

The VopL VCD is a U-shaped dimer stabilized by a coiled-coil

The W_3-C and VCD proteins form high affinity dimers in solution, based on multi-angle laser light scattering (MALLS) analyses (Fig 5-6). Efforts were then made to crystallize VCD protein alone. I found that VCD protein could be

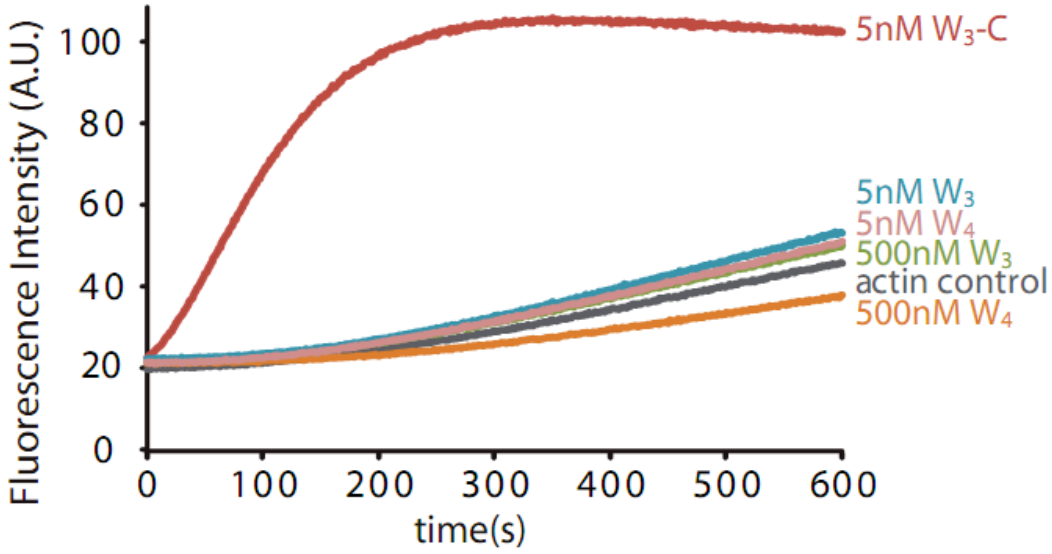


Fig. 5-3 Both the WH2 motifs and C-terminal domain contribute to actin nucleation by VopL. Actin assembly assays performed with 4 μM actin and the indicated concentrations of VopL constructs W3-C, W3 or W4.

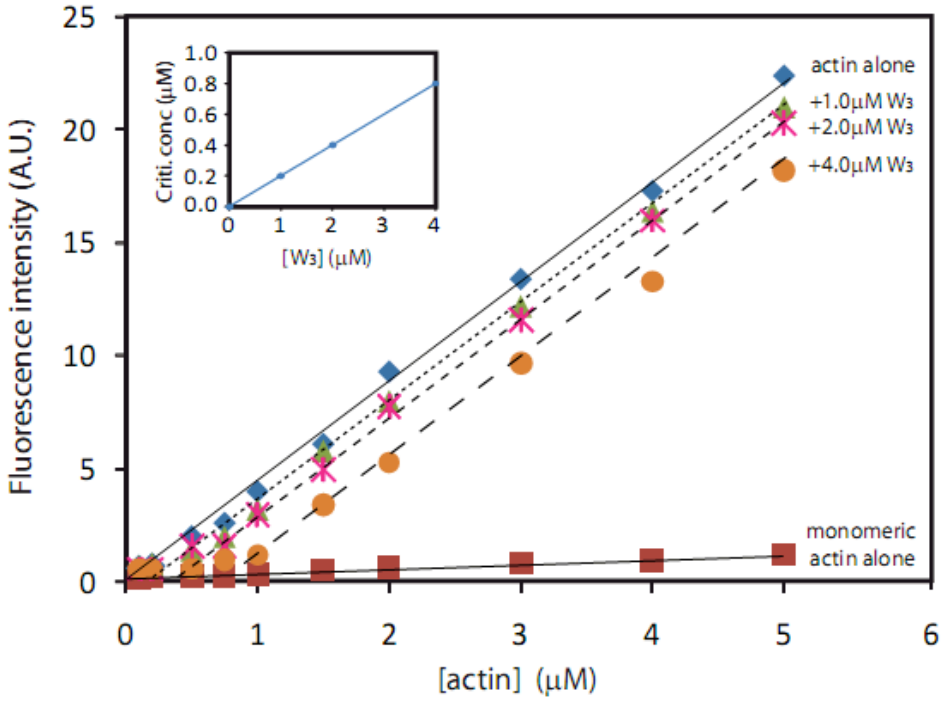


Fig. 5-4 Critical concentration of actin measured to be 0.02 μM, 0.21 μM, 0.40 μM, 0.77 μM in the presence of 0 μM, 1 μM, 2 μM or 4 μM VopL W_3 peptide. Blue diamonds and red squares represent actin in standard polymerization buffer (KMEI buffer) or G buffer. Green triangles, magenta stars and orange circles represent actin in the presence of 1 μM, 2 μM or 4 μM VopL W_3 peptide respectively. Inset demonstrates the linear relationship (slope = 0.19, R^2=0.99) between the critical concentration of actin and VopL W_3 concentration.

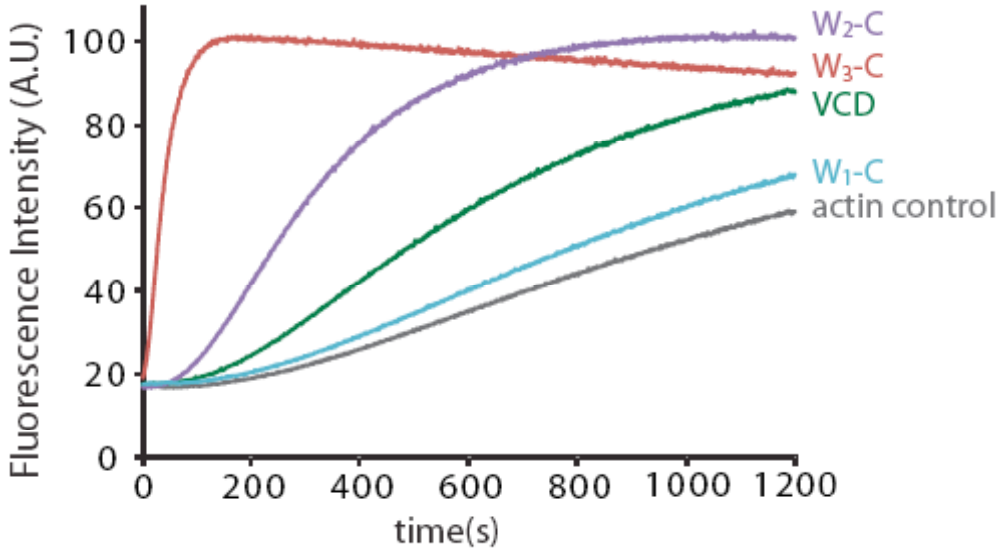


Fig. 5-5 Actin assembly assays contained 4 μM actin and 50 nM W_3-C, W_2-C, W_1-C or VCD.

crystallized at room temperature when polyethylene glycol (PEG 1500 to PEG 8000) was used as precipitant. The quality of protein crystal was the best and diffracted to 3.0 Å when PEG 1500 was precipitant and pH of the drop was in the range between 6.5 to 8.0. Under this condition, VCD was crystallized in space group $P2_12_12_1$ with two molecules in the asymmetric unit. The diffraction resolution was further improved to 2.36 Å resolution by crystal dehydration using PEG20,000. The structure of VCD was determined through molecular replacement using seleno-methionine derivative VCD structure model. The data collection and structure refinement statistics are listed in Table 5-1. Both molecules in the asymmetric unit are well-defined in the electron density map, except for residues 323-338. The VCD is mainly α-helical except for two anti-parallel β strands formed by residues 355-360 and 364-369 (Fig. 5-7). Each monomer is composed of three structural regions: subdomain 1, subdomain 2 and C-terminal helix. Subdomain 1 consists of residues 244-278 and 395-456, and forms an elongated cluster of five α-helices. Subdomain 2 is essentially an insert in subdomain 1 that consists of residues 283-384. These form a bundle of three helices capped distally to subdomain 1 by the two-stranded β-sheet, and with an additional perpendicular helix proximal to subdomain 1. Finally, a C-terminal helix (residues 462-484) emerges perpendicularly to the end of subdomain 1 opposite to subdomain 2.

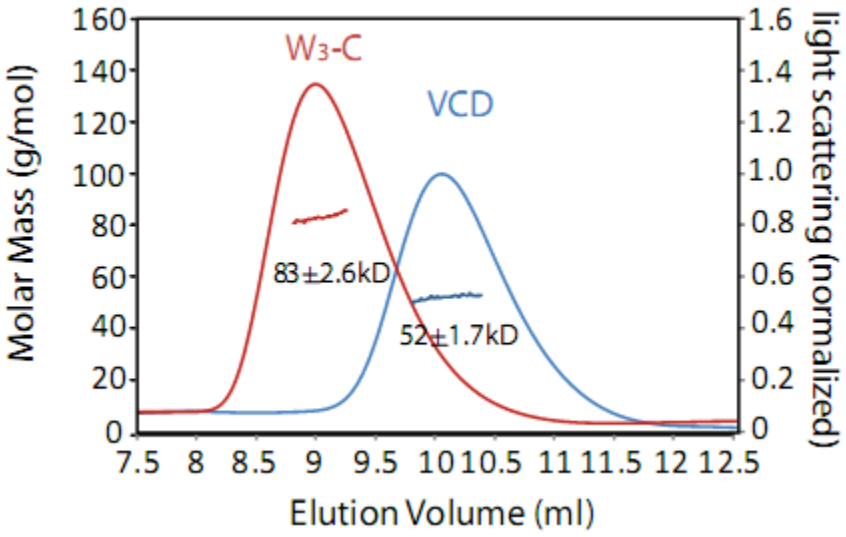


Fig. 5-6 Size exclusion chromatography—multi-angle laser light scattering (MALLS) analysis of VopL constructs W_3-C (red) and VCD (blue). Normalized scattered light in the chromatographic elution (right y-axis) is superimposed on the molecular weight distribution (left y-axis). The molecular weights of monomeric W3-C and VCD are 41 kD and 26 kD respectively.

	Selenomethionine-VopL			Native VopL
Data collection				
Space group	$P2_12_12_1$			$P2_12_12_1$
Cell dimensions				
a, b, c (Å)	56.19 88.64 101.36			55.15 89.81 102.98
	Peak	*Inflection*	*Remote*	
Wavelength	0.9794	0.9795	0.9717	0.9792
Resolution (Å)	50.0-3.10	50.0-3.17	50.0-3.21	50.0-2.30
Rmerge	0.074(0.713)	0.075(0.788)	0.068(0.665)	0.034(0.618)
I / σ(I)	64.2(5.1)	40.5(3.8)	43.7(4.5)	26.4(1.4)
Completeness (%)	99.8(100)	99.9(100)	99.9(100)	99.1(99.0)
Redundancy	23.5(24.2)	11.8(11.9)	11.7(12.1)	2.8(2.5)
Refinement				
Resolution (Å)				29-2.30
No. reflections				23040
Rwork / Rfree				0.221(0.276)/0.275(0.334)
No. atoms				
Protein				3510
Ligand/ion				9
Water				36
B-factors				
Protein				88.92
Ligand/ion				115.4
Water				64.9
R.m.s deviations				
Bond lengths (Å)				0.003
Bond angles (°)				0.564
Missing residues				Chain A: 293-295, 323-338; Chain B: 325-337

Table 5-1. Data collection, phasing and refinement statistics of VopL VCD

*One seleno-methionine and one native crystal were used for data collection. *Values in parentheses are for highest-resolution shell.

There are two potential dimers in the asymmetric unit of the VCD crystals (Figs. 5-7a and 5-8). One is formed by coiled-coil interactions between the C-terminal helices of the two molecules, and by contacts between the adjacent ends of subdomain 1 (Fig. 5-7A). The second is formed by contacts between subdomain 1 from one monomer and subdomain 2 from another monomer (Fig. 5-8). The former dimer has an appreciably larger (1065 Å2 versus 780 Å2 buried) and more hydrophobic intermolecular interface, and is predicted by PDBePISA (Protein Interfaces, Surfaces and Assemblies)(Krissinel and Henrick 2007) to be the more probable of the two possibilities (complexation significance score of 0.2 versus 0.0). Moreover, a VopL protein lacking the C-terminal helix (VCDΔh, residues 247-460) is monomeric in solution by MALLS analysis (Fig. 5-9). Thus, both the structure and mutagenesis data support the functional relevance of the coiled-coil mediated dimer of Figure 5-7A, and we base our further analyses on it.

The VopL dimer forms a shallow U shape, with a wide base formed by end to end association of subdomain 1, and two shorter arms formed by subdomain 2 of each monomer (Fig. 5-7A). The C-terminal coiled coil extends perpendicularly from the center of the base. The N-terminus of the structure is at the subdomain 1-subdomain 2 elbow; the WH2 element would emerge from the VCD at this location. As discussed below, a 24 residue linker connects the first ordered residue in the structure with the C-terminus of WH2c. At the intermolecular

interface a hydrophobic pocket formed by residues T422, A426, L465 of one molecule partially surrounds the sidechain of Y462 in the neighboring molecule (Fig. 5-7B). The sidechains of K421 and Q418 hydrogen bond to those of E417 and T422 from the other chain. The C-terminal coiled coil is mediated by packing of the L465, V469, L472, L476, L479 and L483 sidechains of the two monomers (Fig. 5-7C). The R468 sidechain also hydrogen bonds to the E473 sidechain across this interface.

Subdomain 1 is very similar in structure for the two molecules in the asymmetric unit, with all-atom root mean square deviation (r.m.s.d.) of 0.36 Å. The same is true of subdomain 2, with rmsd of 0.85 Å. However, the organization of the two domains differs between the two molecules due to rigid body movement about the interdomain hinge (residues 279-283 and 384-395, Fig. 5-10). Thus, in solution, VopL likely samples a range of conformations involving fluctuations in the relative orientation of subdomains 1 and 2. This notion is consistent with small angle x-ray scattering data by Dominguez and colleagues that suggest an average co-linear organization of the subdomains in solution (Suk et al., submitted).

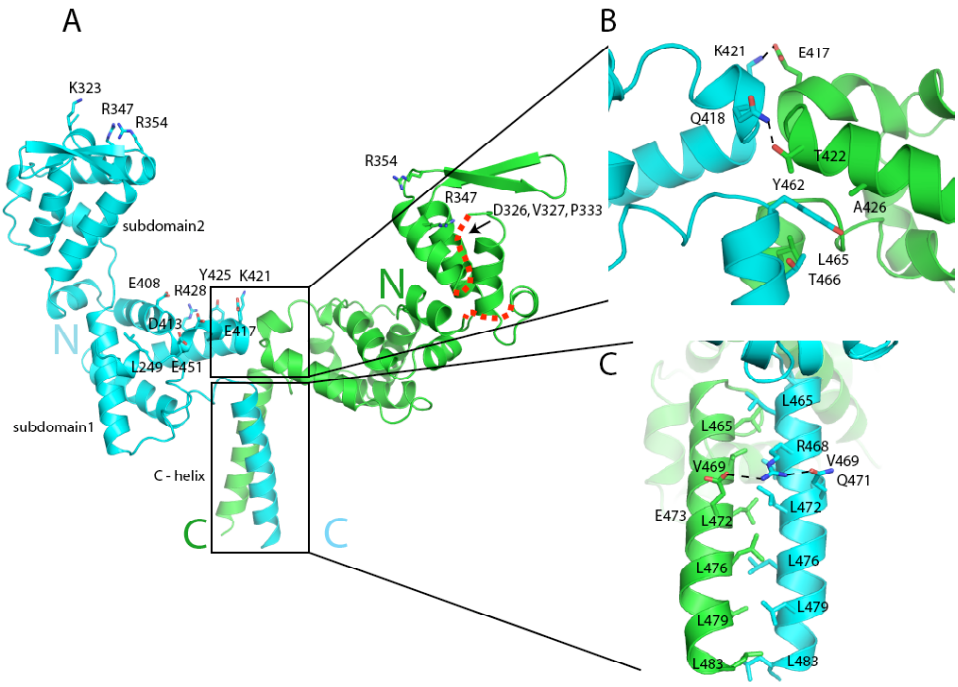


Fig. 5-7 Structure of the VopL VCD dimer. (A) Ribbon diagram of the structure; monomers are colored blue and green. The N- and C-termini of each monomer are indicated. Red dashed lines indicate regions not observed in the electron density map for the green monomer (analogous regions not shown in the blue monomer). Boxed region are enlarged in (B) and (C), which show detailed views of dimer interface. Black dashed lines indicate hydrogen bonds. In all panels sidechains of residues discussed in the text are shown as sticks.

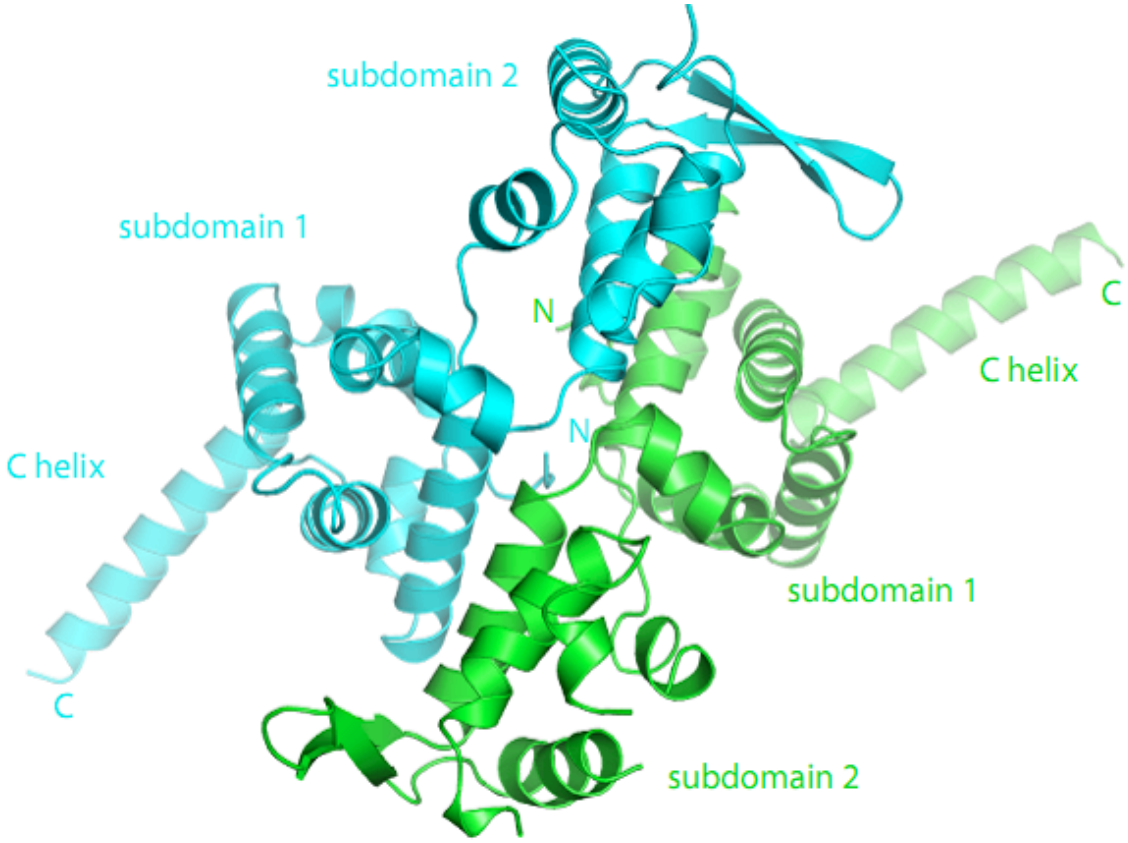


Fig. 5-8 Alternative VopL VCD dimer from that shown in Figure 5-7. The N terminus, C terminus and subdomains of each monomer are denoted.

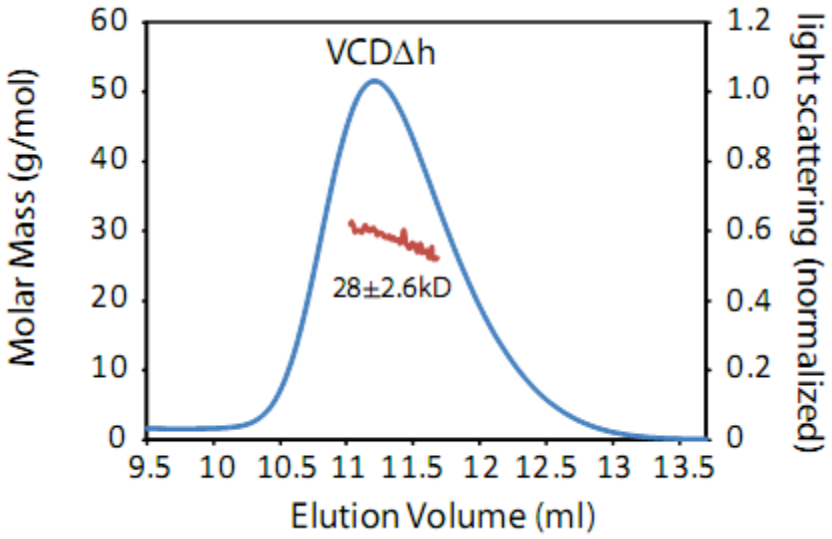


Fig. 5-9 MALLS analysis of VopL VCDΔh. Data plotted as in Fig. 5-6. The molecular weight of monomeric VCDΔh is 24 kD. MALLS analysis indicates VCDΔh is in monomeric form.

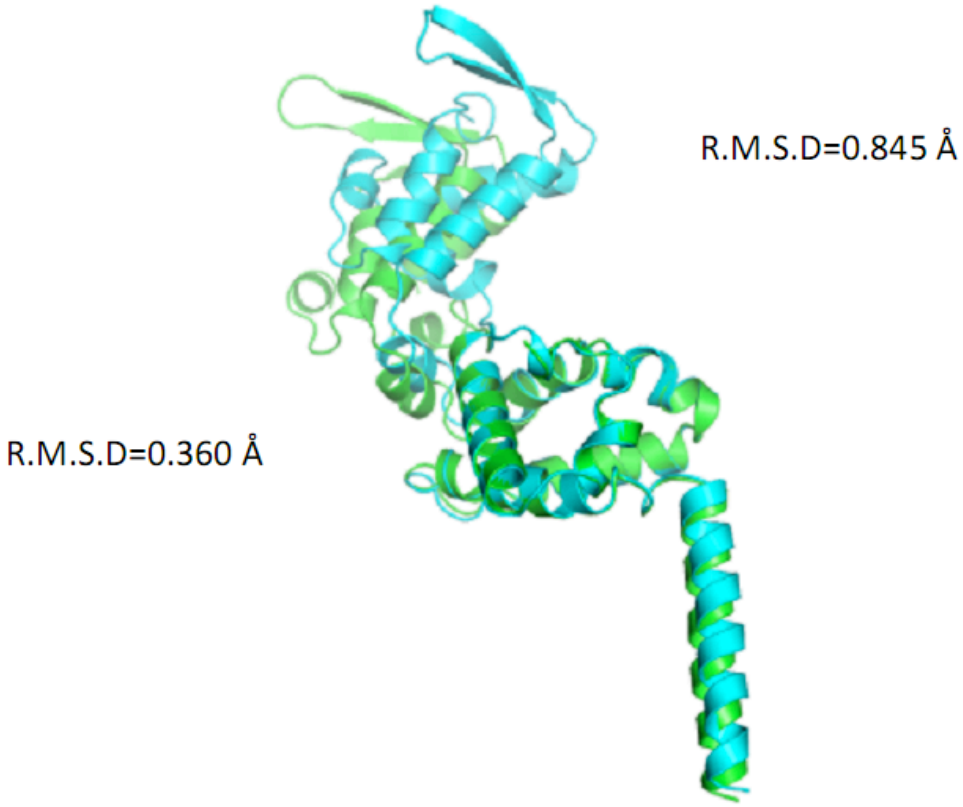


Fig. 5-10 Structure alignment of the two VopL VCD molecules in the dimer. Molecules were superimposed based on subdomain 1. Root mean squared deviations (R.M.S.D.) of all atoms of the individual subdomains 1 and 2 are indicated.

Dimerization and actin contacts of the VCD are important for activity

Compared to the inactive W_3 protein, the highly active W3-C protein is both dimeric (Fig. 5-6) and contains additional potential contact sites to actin. To better dissect these two contributions to activity I performed three different experiments: artificial dimerization of WH2 proteins, mutagenic monomerization of W_3-C, and increasing the number of WH2 motifs in a linear sequence. In the first experiment, I replaced the C terminal domain of WH2-containing constructs with glutathione S-transferase (GST), which forms high affinity dimers (Padrick, Cheng et al. 2008). At a concentration of 50 nM, the engineered W_3-GST protein exhibits substantial nucleation activity, which is comparable to that of W_2-C (Fig. 5-11). Both W_2-GST and W_1-GST are inactive at this concentration. Similar behaviors are observed in N-terminal GST fusions. Thus, dimerization alone can greatly increase the activity of W_3, which is completely inactive as a monomer even at high concentrations. Nevertheless, W_3-GST and W_2-GST remain appreciably less active than their native counterparts dimerized by the VCD. In the second experiments I found that while the activity of the monomeric W_3-CΔh is substantially below that of the dimeric W_3-C counterpart, the former protein is still more active than W3 alone (Fig. 5-12A). This residual activity does not appear to result from residual dimerization, as a W_3-CΔh protein that also

contains K421E and T422A mutations (W3-CΔh_{KT}) designed to weaken dimerization contacts between subdomains 1 is similarly active (Fig. 5-12B). Finally, I generated a W_4 construct by adding an additional WH2 motif to the N terminus of W_3. As shown in Figure 5-3, W_4 proved to be completely inactive. Together, the data suggest that a specific spatial organization of the WH2 motifs, rather than merely their total number, is likely important for function. Moreover, while dimerization of the WH2 element is necessary for high activity, the VCD appears to contribute in ways beyond mere dimerization, perhaps through an ability to bind and orient actin molecules to favor nucleation (see below). Thus, the data imply a specific organization of actin in the nucleus, generated by the WH2 elements juxtaposed by the VCD.

To test the hypothesis that the VCD might contact actin, I mutated residues at the distal end of subdomain 2 and at the base of the VCD "U" in subdomain 1. Residues K323, R347 and R354 form a conserved basic patch at the distal end of subdomain 2 (Fig. 5-7A). Mutations of these three residues to glutamic acid in W_3-C (construct W_3-C_{KRR}) or the VCD alone (construct VCD_{KRR}) substantially decreased activity (Figs. 5-13). I also mutated to glycine the three residues D326, V327 and P333 in the loop Y322-G338 near the end of subdomain 2 (constructs W_3-C_{DVP} and VCD_{DVP}, Fig. 5-7A), which is not observed in the electron density and is likely flexible in solution. These mutations also produced a small decrease

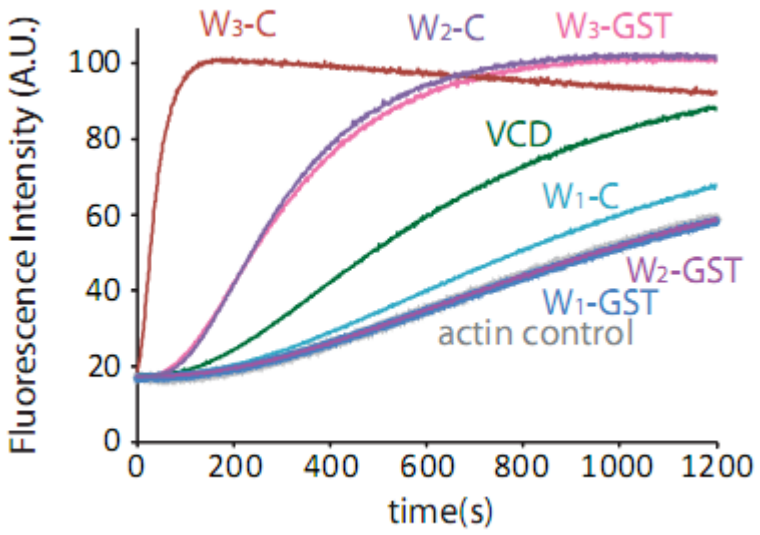


Fig. 5-11 Actin polymerization assays performed with 4μM actin and 50nM W_3-C, W_2-C , W_1-C, C, W_3-GST, W_2-GST or W_1-GST.

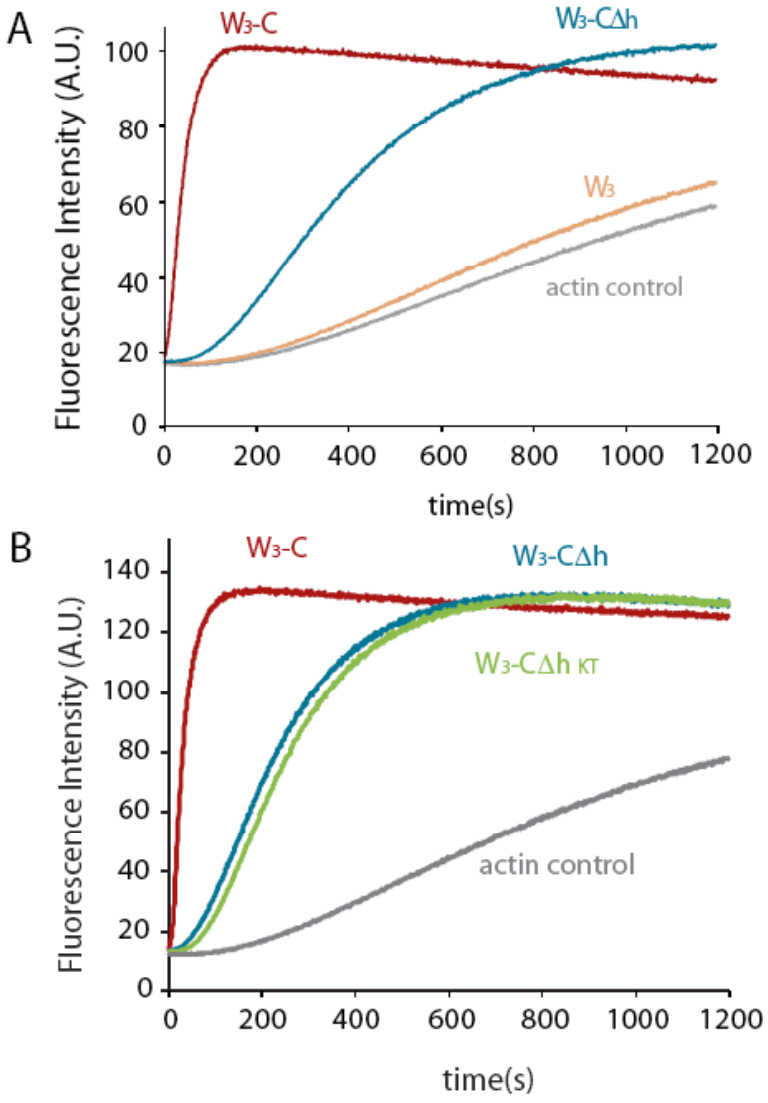


Fig. 5-12 VCD dimerization is important for VopL nucleating activity. Assays were performed with 4 μM actin and 50 nM indicated VopL proteins. (A) Comparison of W_3-C with the VCD monomeric mutant, W_3-CΔh and W_3. (B) Comparison of W_3-C with W_3-CΔh and W3-CΔh_{KT}.

in activity (Figs. 5-13). Thus, the distal end of subdomain 2 appears to contact actin during nucleation.

The surface of subdomain 1 that faces into the VCD "U" consists of three distinct patches: a basic patch containing residues K421, Y425, and R428; an acidic patch containing E408, D413, and E417; and a mixed patch containing L249 and E251 (Fig. 5-14A). Mutation of the first patch (KYR to AAD (W_3-C_{KYR}) or to KAD (W_3-C_{YR})), or second patch (EDE to KKA (W_3-C_{EDE})) produced small but reproducible decreases in activity (Fig. 5-14B). Mutation of the third patch had no effect on activity (data not shown). Thus, subdomain 1 may also contact actin during nucleation, although these interactions do not appear as important for activity as those to subdomain 2.

Multi-angle laser light scattering analyses indicate that the wild type and mutant proteins are largely (and similarly) dimeric at concentrations as low as 500 nM (Fig. 5-15A). In addition, the activities of several VopL proteins are proportional to concentration between 5 nM and 50 nM, suggesting the proteins are mostly dimeric even at low concentrations (Fig. 5-15B). These data support the argument that the mutations perturb contacts to actin rather than causing loss of dimerization. Together, my biochemical analyses suggest that VCD likely plays a direct role in organizing actin monomers in the VopL nucleus, in addition to dimerizing and orienting the WH2 elements.

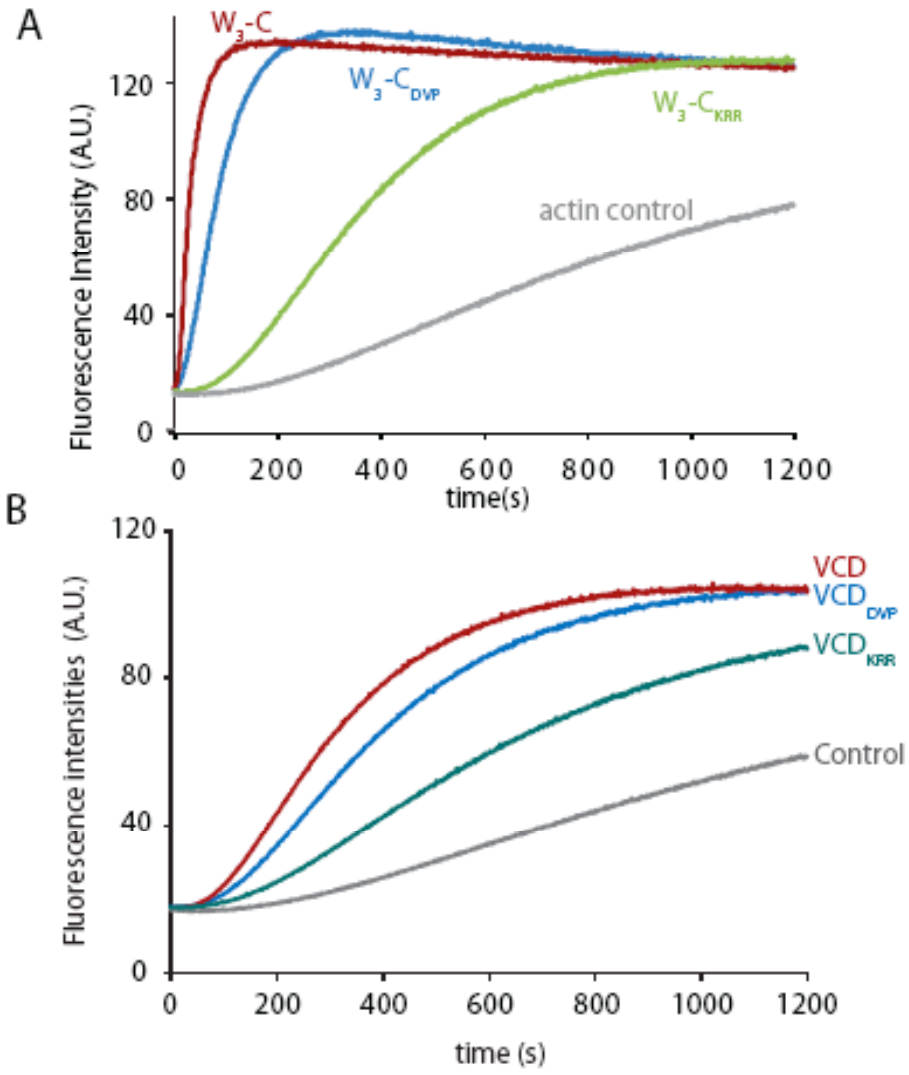


Fig. 5-13 The basic arm of VCD subdomain 2 contributes to VopL nucleating activity. Figure (A) shows the comparison of W_3-C with the VCD subdomain 2 mutants, W_3-C_{KRR} and W_3-C_{DVP}. Figure (B) shows the comparison of VCD with the VCD subdomain 2 mutants, VCD_{KRR} and VCD_{DVP}. Assays were performed with 4 μM actin and 50 nM indicated VopL proteins.

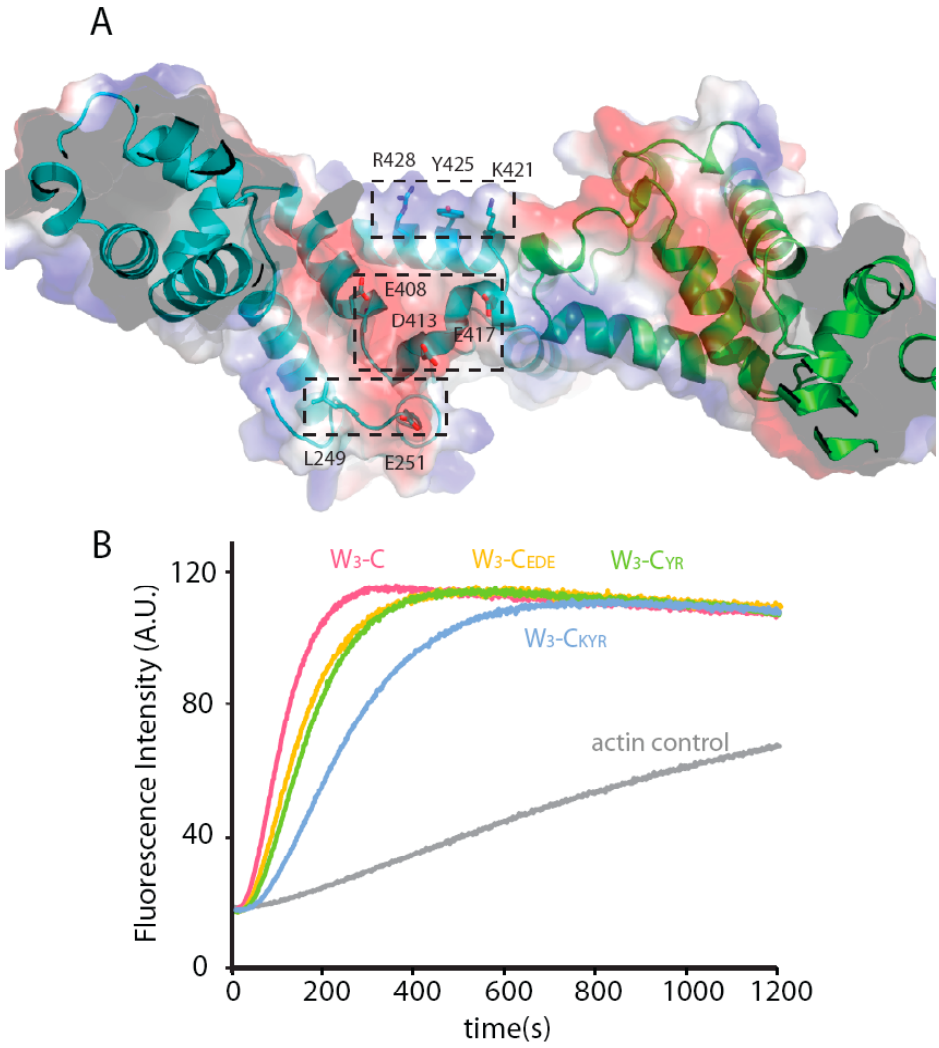


Fig. 5-14 VCD subdomain 1 may also contact actin and contribute to nucleating activity. (A)The surface charge of VopL subdomain 1 from a top view. Blue color represents positive charged and red color represents negative charged. The boxed regions show the three distinct patches. The residues' side chains are shown as sticks. (B) Comparison of W_3-C with the VCD subdomain 1 mutants, W_3-C_{KYR}, W_3-C_{YR} and W_3-C_{EDE}.

VopL activity is sensitive to the length of linker between WH2c and VCD

It is somewhat paradoxical that W1-C has lower activity than both W2-C and VCD (Fig. 5-5). That is, adding one WH2 motif to the VCD is inhibitory but adding two WH2 motifs is stimulatory. We hypothesized that this effect might arise because the natural 24-residue linker between the C-terminus of WH2c (T223, of the basic sequence LRKT) and the beginning of the VCD (R247) is too short, causing sub-optimal orientation of the initial bound actin monomer(s).

To test this idea, I inserted 9- and 40-residue linkers into W_1-C, to give a total of 33 and 64 residues, respectively, between the WH2c motif and VCD (W_1-9-C and W_1-40-C, Fig. 5-1). The linker in W_1-40-C is identical in length to that between WH2b and the C-terminus of W_2-C. As shown in Figure 5-16, increasing linker length in W_1-C progressively increases activity. In related experiments, when the WH2c motif in W_2-C or W_3-C is replaced by an equal length Gly-Gly-Ser linker (to give constructs W_1-L-C and W_2-L-C), activity decreases only modestly. Thus, a smaller number of WH2 motifs (e.g. W_3-C versus W_2-L-C) can be largely compensated for by increasing the spacing of those motifs from the VCD. These results are consistent with the idea that the natural linker between WH2c and VCD is too short (in one or both molecules in the dimer), likely preventing optimal organization of actin monomers for nucleation activity.

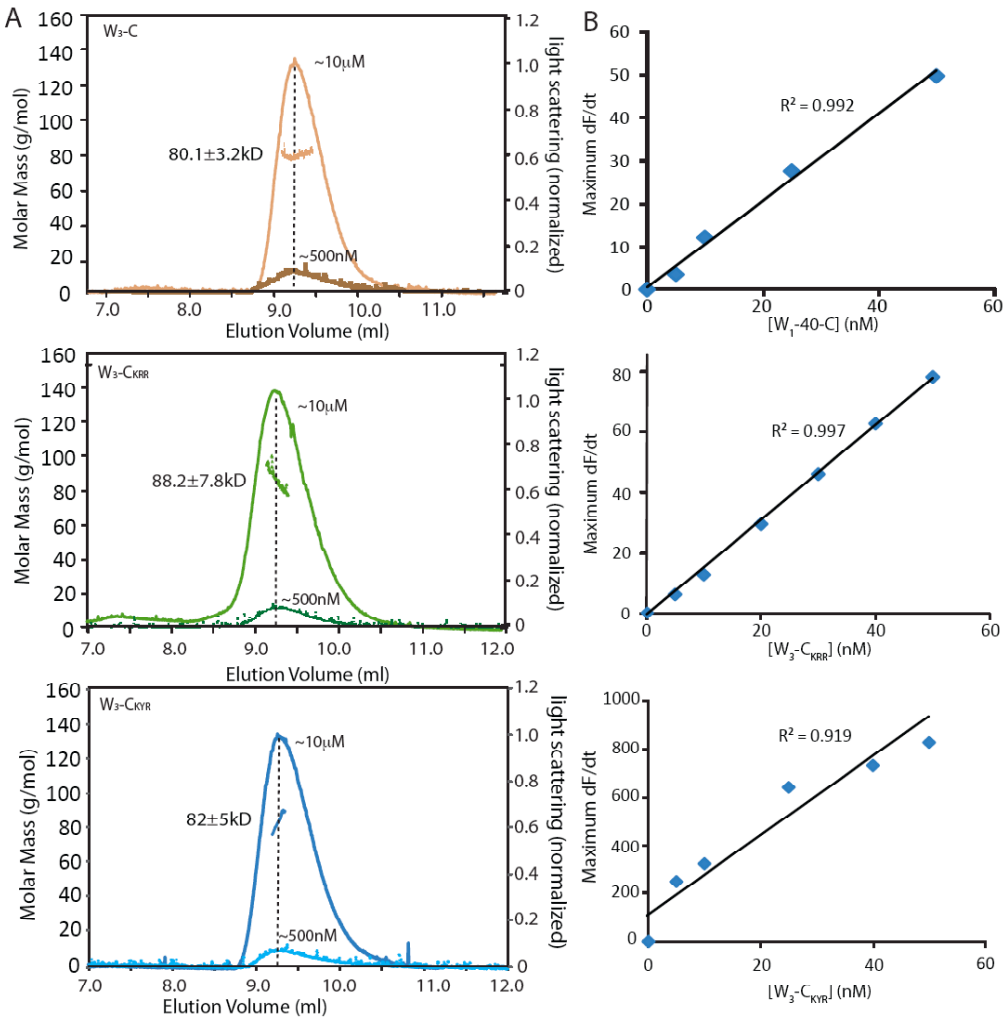


Fig. 5-15 Stability of the VopL dimer. (A). MALLS analysis of W3-C, W3-C_{KRR} and W3-C_{KYR} proteins at concentrations of 10 μM and ~500 nM. The same elution position at the two concentrations indicates the same oligomeric state of the protein. (B). Maximum actin polymerization rates after subtracting the rate of actin alone (represented by the maximum fluorescence changes over time: dF/dt) produced by W_1-40-C, W_3-C_{KRR}

and W_3-C_{KYR} proteins are plotted against VopL concentrations (5-50 nM). Lines show linear fits to the data with errors shown as R square. The linear relationship between activity and VopL concentration suggests the VCD dimer does not appreciably dissociate at concentrations as low as 5 nM.

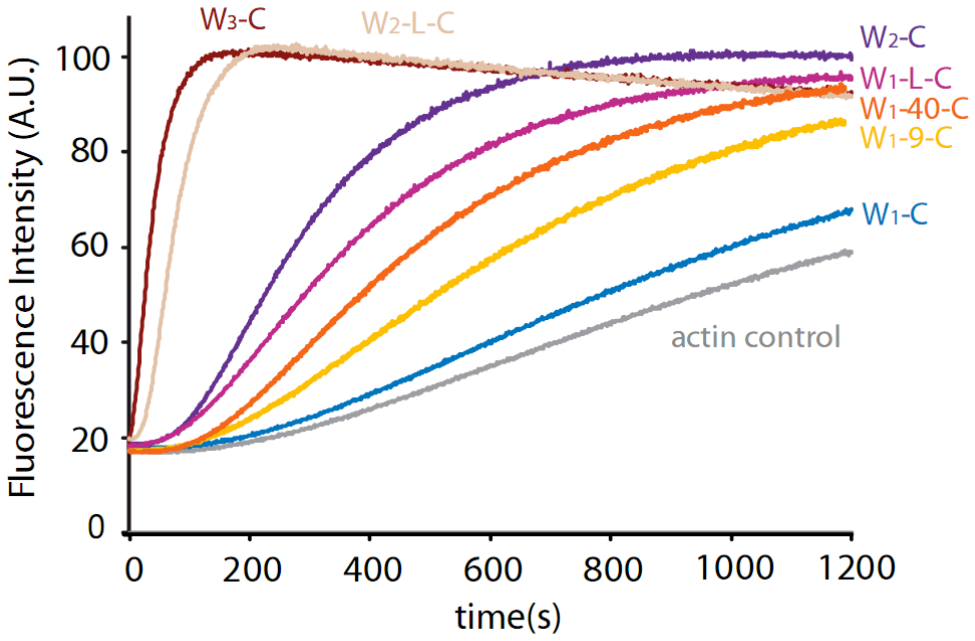


Fig. 5-16 Increasing the linker between WH2c and the VCD increases actin assembly activity. Assays contained of each 50 nM VopL protein, W_3-C, W_2-C, W_1-C, W_2-L-C or W_1-L-C.

Discussion

Our data demonstrate that multiple WH2 motifs and the VCD of VopL are both required for high potency actin filament assembly. Dimerization, either by the VCD or artificially by GST, is necessary for the WH2 motifs to nucleate. The VCD itself also appears to contact actin, since it alone can nucleate weakly, and mutagenesis of conserved residues in subdomains 1 and 2 decreases activity. These key data, along with the geometry of the VCD structure, suggest a preliminary model for the actin nucleus assembled by VopL. A central concept in this model is that the WH2 motifs efficiently recruit actin monomers, while the VCD provides organization that leads to a productive nucleus.

The requirement for dimerization of the WH2 element suggests that in the nucleus, at least one actin monomer is contributed/recruited by each chain. An individual multi-WH2 protein should be able to assemble a single strand of the nascent paired actin filament (i.e. stabilize long-pitch contacts between actin monomers). But such activity is evidently insufficient for nucleation by VopL. This observation suggests that in VopL dimers, the paired WH2 elements act by stabilizing lateral (short-pitch) actin-actin contacts. Thus, the nucleus should minimally contain a short-pitch actin dimer.

How might such a structure interact with the VCD? One possibility is suggested by geometric considerations. In a short pitch actin dimer, the two subunits are axially displaced on average by 27.5 Å (Holmes, Popp et al. 1990). In the VCD, the distal end of subdomain 2 is ~30 Å from the platform created by the two subdomains 1. Thus, the axial displacement of an actin dimer could be accommodated by only small changes in the structural organization of the VCD. In this configuration the penultimate actin monomer would contact the distal end of subdomain 2 and the terminal monomer would contact subdomain 1 (Fig. 5-17). Such an arrangement would be consistent with several observations. First, my mutagenesis data (Fig. 5-13, 5-14) suggest that residues at both the distal end of subdomain 2 and the VCD base formed by subdomain 1 contribute to nucleation activity, probably through contacts to actin. Second, in order for WH2c to bind this penultimate actin, the WH2c-VCD linker would need to extend from the N-terminus of the VCD, around subdomain 2 to its tip, a distance of ~50 Å, and then approximately half the length of an actin monomer to the N-terminus of the WH2 motif(Chereau, Kerff et al. 2005), ~28 Å. While these distances are obviously very approximate, they would be consistent with the idea that the natural 24 residue linker (~80 Å fully extended) may be poorly able to position a WH2c-bound actin on the VCD, explaining the low activity of W1-C, and the progressive increase in activity as the linker is lengthened (Fig. 5-16). In native VopL, the penultimate monomer may be recruited by WH2b. Third, SAXS

analyses from Dominguez and colleagues on the W_1-C: actin complex suggest that an ordered actin contacts the subdomain 1 platform (Suk et al., submitted). This terminal monomer could be recruited by the WH2c of the opposing VCD. Finally, the short-pitch dimer bound to the VCD would be further stabilized by additional actin monomers recruited by additional WH2 motifs, consistent with the activity series W_3-C > W_2-C > W_1-L-C (Fig. 5-16). While this model remains speculative, it makes testable predictions that will be explored in future work.

The stabilization of short-pitch contacts in the nascent actin filament appears to be a common feature of WH2-based nucleation factors, both prokaryotic and eukaryotic. In Cobl and Lmod, the stoichiometry of actin binding and linker length dependencies of activity have led to models in which WH2 and other motifs stabilize a short-pitch actin trimer (Ahuja, Pinyol et al. 2007; Chereau, Boczkowska et al. 2008). In Spire, a tandem array of WH2 motifs was shown by electron microscopy to organize a linear actin structure, which was proposed to nucleate through serving as a single, long-pitch strand of an actin filament (Quinlan, Heuser et al. 2005). However, Spire alone is a relatively weak nucleation factor, whose activity is significantly enhanced through dimerization mediated by the formin protein, Cappuccino (Quinlan, Hilgert et al. 2007). Similarly, the actin nucleation factor TARP from the pathogen *Chlamydia trachomatis* requires oligomerization of its WH2 motif by an adjacent poly-

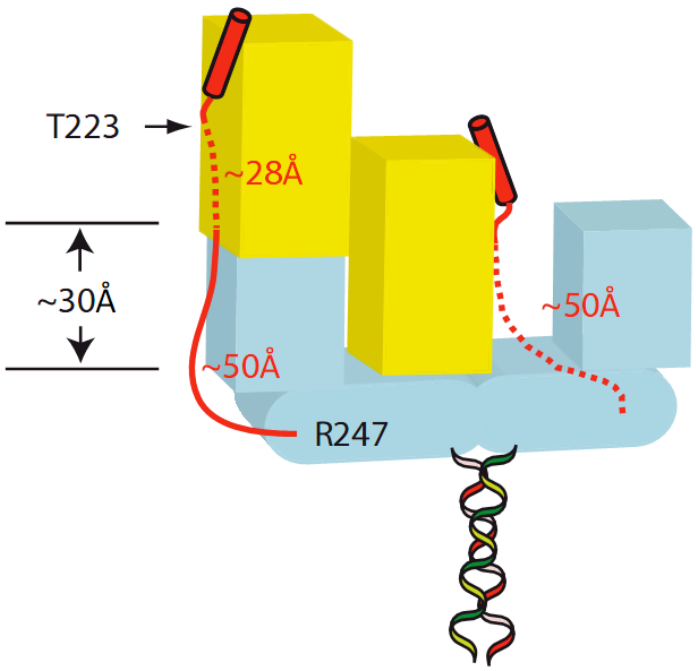


Fig. 5-17 Model for a minimal actin nucleus assembled by VopL. VCD is indicated by blue blocks with ribbons representing the coiled coil. Initial actin monomers are indicated by yellow boxes. Actin-bound WH2c motifs are shown as red cylinders. Red lines indicate linkers between WH2c and VCD. Additional actin monomers assembled by WH2a and WH2b would bind to the upper surfaces of the WH2c-bound actins, making additional short- and long-pitch contacts, further stabilizing the assembly. Actins are implicitly shown with barbed ends directed away from the VCD, but my current data do not speak to the orientation of actin monomers in the VopL nucleus.

proline region for activity(Jewett, Fischer et al. 2006). In both Spire and TARP, higher potency likely arises, as in VopL, from the ability of dimers to organize lateral contacts between actin monomers in the nascent filament. These experimental findings are all consistent with computational analyses of actin nucleation, which suggest that a stabilized short-pitch dimer will recruit a third monomer with much greater affinity than will a long-pitch dimer (Sept and McCammon 2001). Thus, short-pitch dimers should act as more effective nuclei.

Together, our findings provide an initial mechanistic model for the potent actin nucleation activity of VopL. Yet many questions remain to be answered. What is the conformation of the VopL-bound actin nucleus, and what contacts to both WH2 motifs and the VCD stabilize it? Does VopL remain bound to filaments after nucleation? If so, does it bind filament ends or sides; does it remain static like the Arp2/3 complex or process like formins? If not, what triggers its release from the nascent filament? Why do VopL and VopF, which presumably nucleate through very similar mechanisms, produce different actin structures in cells? My work here provides an initial framework to address these questions and others in the future.

Methods and Materials

Molecular Biology and Protein Purification.

We cloned VopL fragments (W_3-C: 115-484, W_2-C: 155-484, W_1-C: 194-484, VCD: 247-484 and WH2 peptides: WH2a: 129-157, WH2b: 158-187, WH2c: 199-226, W_3: 115-246) into pGEX vector by PCR and verified them by DNA sequencing. Proteins were expressed in *Escherichia coli* BL21(DE3)^{T1R} cells grown in LB media and induced with 1 mM IPTG at optical density ($O.D._{600}$) of 0.8. Proteins were purified with successive glutathione-sepharose affinity chromatography, tobacco etch virus (TEV) protease cleavage, followed by strong anion exchange and gel filtration chromatographies. GST fusions of VopL WH2 fragments (W_3-GST, W_2-GST and W_1-GST respectively) were expressed from pET11a and purified using glutathione-sepharose affinity, strong anion exchange and gel filtration chromatographies.

Pyrene Fluorescence Based Actin Polymerization Assay

Actin was prepared from rabbit muscle and kept at 4 °C with consistent dialysis against G-buffer. The buffer should be replaced periodically. Our experience is that actin of good quality can be extracted from muscle acetone powder that has been stored at -80 °C for up to 4 months. Pyrene labeled actin

was obtained by mixing actin and pyrene iodoacetamide (Molecular Probes) in 10-fold excess for overnight and subsequent purification. Concentration of pyrene and actin can be calculated by the following equations.

$$[\text{Pyrene } (\mu M)] = OD_{344}/2.2*10^4(M^{-1})$$

$$[\text{Actin}(\mu M)] = (OD_{290} - (OD_{344} *0.127)) /2.66*10^4(M^{-1})$$

All actin polymerization assays contained 4 μM actin (5% pyrene labeled) and 5 nM-500 nM VopL proteins in KMEI-G Buffer (10 mM Imidazole (pH 7.0), 50 mM KCl, 1 mM $MgCl_2$, and 1 mM EGTA), and were performed at 20 °C. All protein samples that are added in very small volumes are kept on ice until the start of a reaction. Assays were performed using a Jobin Yvon Horiba Fluorimeter (www.jyhoriba.com). I monitored the progress of actin polymerization through the increase of pyrene fluorescence intensity. Pyrene-actin fluorescence (λ_{ex} = 365 nm, λ_{em} = 407 nm) is measured every second until it plateaus. The half-time ($t_{1/2}$) of actin assembly was measured as the time for 50% polymerization. Errors were calculated as the standard deviation from three independent measurements.

Protein Crystallization and X-ray Diffraction Data Collection.

VCD crystals (50 × 150 × 200 μm) were grown at 20°C using the hanging-drop vapor-diffusion method. Drops contained 1μl of 10-15 mg/ml protein

solution (20 mM Tris, pH 7.5, 100 mM NaCl, 1 mM DTT) plus 1 μl of reservoir solution (100 mM Tris pH 7.5, 20% polyethylene glycol (PEG) 1500) and were equilibrated against 200 μl of reservoir solution. Crystals of selenomethionine-labeled protein were obtained under similar conditions with 1% PEG 3350 as the additive. Crystals were cryoprotected by addition of 30% (v/v) ethylene glycol to reservoir solution. Native crystals were soaked in a cryosolution that also contained 5% PEG20,000 for 30 minutes prior to flash cooling. Native and anomalous dispersion data were collected at the Advanced Photon Source 19ID beamline at 100 K. Native VCD crystallized in space group $P2_12_12_1$ with unit cell parameters of a = 56 Å, b =88 Å, c = 101 Å, containing two molecules in each asymmetric unit with ~50% solvent. Crystals diffracted X-rays to a minimum Bragg spacing of 2.30 Å. Data indexing and scaling were carried out using the HKL3000 suite of programs (Minor, Cymborowski et al. 2006). Data collection statistics are in Table 5-1.

Phase Determination and Structure Refinement

We used multi-wavelength anomalous diffraction data to a resolution of 3.1-3.21 Å from a selenomethionine VopL crystal to obtain experimental phases. The peak, reflection and remote wavelengths were 0.9794 Å, 0.9795 Å and 0.9717 Å, respectively. Ten of ten expected selenium sites were located using the program

SHELXD (Schneider and Sheldrick 2002). Phases were refined with the program MLPHARE (Otwinowski 2001) and further improved to a figure-of-merit of 0.90 by density modification using the program DM (CCP4 (Collaborative Computational Project 1996). An initial structure model was built using the program Buccaneer (Cowtan 2006) and further improved by manual model building using the program Coot (Emsley and Cowtan 2004). I collected a native data set at the same temperature and beamline, at a wavelength of 0.9792 Å. An electron density map of native VopL was obtained through molecular replacement using the selenomethionine VopL as a search model using the program Phaser (McCoy, Grosse-Kunstleve et al. 2007). The native structure model was again initially built using Buccaneer and improved with manual model building. Structure refinement with native data was carried out using the program PHENIX (Adams, Grosse-Kunstleve et al. 2002), consisting of refinements of individual sites, individual B-factors and TLS parameters. The current model contains two VopL chains in one asymmetric unit. Statistics of refinement and structural quality are listed in Table 5-1. In the refined structure, 97.5%, 2.5% and 0.0% of residues are in the favored, allowed and disallowed regions of the Ramachandran plot, respectively, according to the Molprobity structural validation tools (Davis, Leaver-Fay et al. 2007).

Multi-Angle Laser Light Scattering (MALLS)

We applied 500 μl of 40 μM VopL proteins at a flow rate of 0.5 ml/min to either a Superdex 200 10/300 or a Superdex 75 10/300 column (Amersham Biosciences) equilibrated in buffer containing 20 mM Tris pH 8.0, 100 mM NaCl and 2 mM DTT. The chromatography column was combined in line with a three angle light scattering detector (miniDAWN TREOS) coupled to a refractive index detector (Optilab rEX, Wyatt Technology). Data analysis and molecular weight calculations were carried out using the program ASTRA V (Wyatt technology). For Figure 5-15A, VopL proteins were first applied to the column and the estimation of protein concentration at the elution peak was 10 μM after taking into account the effect of dilution. Then the same sample was diluted by 20 times and applied to the column to check the elution position at roughly 500 nM.

Isothermal Titration Calorimetry (ITC)

All ITC experiments were performed using a VP-ITC MicroCalorimeter at 20°C. Prior to each experiment, WH2 peptides and actin were dialyzed extensively in the same beaker against 2 liters of G-Buffer for at least 24 hours. G-Buffer contains 2 mM Tris, pH 8.0, 0.1 mM $CaCl_2$, 0.5 mM DTT, 1 mM NaN_3 and 0.2 mM ATP. In each experiment, 100 μM of WH2 peptide was titrated into

10 μM actin in G-Buffer. Data were analyzed and fitted to a single site binding model using the program Origins 7, Microcal for ITC.

Chapter 6

Concluding Remarks II

WH2 motif-based actin nucleators are a group of newly identified players in actin polymerization. To date, three such eukaryotic proteins have been discovered: Spire, Cobl and Lmod (Quinlan, Heuser et al. 2005; Ahuja, Pinyol et al. 2007; Chereau, Boczkowska et al. 2008). They all utilize WH2 motifs to recruit and nucleate actin. However, the employed nucleation mechanisms look to be quite different, and the function these molecules exhibit inside the cell also appear to be distinct. On the other hand, many bacterial pathogens hijack the host cell actin cytoskeleton during infection (Gouin, Welch et al. 2005; Bhavsar, Guttman et al. 2007; Cossart and Toledo-Arana 2008). *Vibrio* effector VopL directly targets actin assembly to take control of actin dynamics. It makes use of WH2 motifs and has high potency in promoting actin filament nucleation. The molecular mechanism underlying this strong activity was yet to be uncovered. The studies performed here are not only additions to our knowledge of WH2 based nucleators, but also provide an initial framework for future studies.

Based on our analysis of VopL, we are now able to compare the different aspects of eukaryotic and prokaryotic WH2 based nucleators. We can tell that the potency of the nucleator is highly related to the type of actin nucleus formed. Through oligomerization of WH2 motifs, VopL can recruit up to six actin monomers and potentially stabilize a short pitch actin hexamer for filament seeding. Stabilization of a short pitch actin dimer through oligomerization may not be a unique feature of VopL. Spire protein, for example, gains great enhancement in activity through binding to the dimeric formin protein. The nucleating activity of TARP also relies on the oligomerization of the N-terminal proline-rich region. These observations all support the idea that a short pitch dimer is more efficient in promoting actin filament growth.

A unique feature of the VopL VCD is the potential contacts between VCD and actin that seem to stabilize the nucleus formation. My mutagenesis results suggest that a basic patch on subdomain 2 of VCD may be the key binding site for this interaction. A crystal structure of VCD complexed with actin will address this problem nicely. However, the difficulty lies in the low affinity between VCD and monomeric actin. Adding WH2 motifs at the N terminus may help to stabilize the VopL-actin complex, but in the meantime it can introduce more flexibility issues. Attempts should be made to optimize the linker length between WH2 and VCD to form the minimal and most stable complex for crystallization.

There is still debate regarding the positions of WH2 based nucleator at growing filament. Based on my preliminary single filament fluorescence imaging results (thanks to the studies performed by our collaborators in Jeff Gelles' Laboratory), we can see distinct events where VopL binds to either the filament barbed end, pointed end or side. However, it seems in most events, VopL binds to the pointed end of filaments during the nucleation process. My current hypothesis is that the VopL VCD contacts the pointed end of actin monomers since WH2 domain binds to the barbed end of actin and it is at the N-terminus of VCD. The WH2 motifs deliver four actin monomers to form the minimal nucleus. VopL dissociates from the nucleus quickly after its formation, and subsequently WH2 motifs bind to either the filament side or barbed end independently of the VCD (it has been shown that the WH2 motif can bind to filament barbed end, see (Co, Wong et al. 2007). This hypothesis is yet to be tested by more structural and imaging studies.

The multiple binding positions of WH2 based nucleators to the filament are associated with the multiple potential functions of the proteins. Spire protein was shown to sequester actin monomers in a 1:4 complex, and it inhibits filament barbed end growth from profilin-actin (Bosch, Le et al. 2007). It was therefore proposed that Spire may sequester actin monomers or sever actin filaments *in vivo*. It will be interesting to investigate whether other nucleators also possess multifunctionality like Spire. Notably, one difference between Spire and other

nucleators is that it is a much weaker nucleation promoting factor. Thus it is likely that Spire has a higher tendency to fulfill its other functions. The observation that high concentrations (above 40 nM) of Cobl leads to a decline in the nucleation rate implies it may also sequester actin monomers (Ahuja, Pinyol et al. 2007). In my studies, high concentration (up to 500 nM) of VopL did not result in lower nucleation rate, but the presence of high concentration of W_3 motif would increase the critical concentration of actin. These results suggest that VopL WH2 motif alone may sequester actin monomers, but that full length VopL mainly functions as an actin nucleator. In addition, the filament side binding feature of VopL suggests that it may bundle and crosslink actin filaments together. This idea is consistent with the observation that VopL promoted the formation of stress fibers in host cell during infection (Liverman, Cheng et al. 2007).

The poly-proline regions of VopL, which albeit do not seem to affect its nucleating activity *in vitro*, may play an important role in recruiting profilin sequestered actin monomers and delivering actin to WH2 motifs. Therefore, the N-terminal fragment may be essential for VopL-mediated actin nucleation *in vivo*. Cell based studies will be required to address the importance of this region.

The discovery of WH2 based nucleators has opened a new door for understanding actin dynamics. Although notable progress has been made, many questions remain to be addressed including the nature of the complex structure of

actin bound to WH2 based nucleator, the mechanism of generating different actin structures by these nucleators, and the crosstalk between WH2 based nucleators and other signaling molecules in regulating actin cytoskeleton rearrangements. Related research is ongoing and is likely to dominate the actin field for the coming years.

References:

Abe, K., I. P. Whitehead, et al. (1999). "Involvement of NH(2)-terminal sequences in the negative regulation of Vav signaling and transforming activity." J Biol Chem **274**(43): 30410-30418.

Adams, J. M., H. Houston, et al. (1992). "The hematopoietically expressed vav proto-oncogene shares homology with the dbl GDP-GTP exchange factor, the bcr gene and a yeast gene (CDC24) involved in cytoskeletal organization." Oncogene **7**(4): 611-618.

Adams, P. D., R. W. Grosse-Kunstleve, et al. (2002). "PHENIX: building new software for automated crystallographic structure determination." Acta Crystallogr D Biol Crystallogr **58**(Pt 11): 1948-1954.

Aghazadeh, B., W. E. Lowry, et al. (2000). "Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation." Cell **102**(5): 625-633.

Ahuja, R., R. Pinyol, et al. (2007). "Cordon-bleu is an actin nucleation factor and controls neuronal morphology." Cell **131**(2): 337-350.

Alai, M., A. L. Mui, et al. (1992). "Steel factor stimulates the tyrosine phosphorylation of the proto-oncogene product, p95vav, in human hemopoietic cells." J Biol Chem **267**(25): 18021-18025.

Alexandropoulos, K. and D. Baltimore (1996). "Coordinate activation of c-Src by SH3- and SH2-binding sites on a novel p130Cas-related protein, Sin." Genes Dev **10**(11): 1341-1355.

Amarasinghe, G. K. and M. K. Rosen (2005). "Acidic region tyrosines provide access points for allosteric activation of the autoinhibited Vav1 Dbl homology domain." Biochemistry **44**(46): 15257-15268.

Bamburg, J. R., H. E. Harris, et al. (1980). "Partial purification and characterization of an actin depolymerizing factor from brain." FEBS Lett **121**(1): 178-182.

Belmont, L. D., A. Orlova, et al. (1999). "A change in actin conformation associated with filament instability after Pi release." Proc Natl Acad Sci U S A **96**(1): 29-34.

Bhavsar, A. P., J. A. Guttman, et al. (2007). "Manipulation of host-cell pathways by bacterial pathogens." Nature **449**(7164): 827-834.

Blanchoin, L., T. D. Pollard, et al. (2000). "Interactions of ADF/cofilin, Arp2/3 complex, capping protein and profilin in remodeling of branched actin filament networks." Curr Biol **10**(20): 1273-1282.

Bompard, G. and E. Caron (2004). "Regulation of WASP/WAVE proteins: making a long story short." J Cell Biol **166**(7): 957-962.

Booden, M. A., S. L. Campbell, et al. (2002). "Critical but distinct roles for the pleckstrin homology and cysteine-rich domains as positive modulators of Vav2 signaling and transformation." Mol Cell Biol **22**(8): 2487-2497.

Bosch, M., K. H. Le, et al. (2007). "Analysis of the function of Spire in actin assembly and its synergy with formin and profilin." Mol Cell **28**(4): 555-568.

Braiman, A., M. Barda-Saad, et al. (2006). "Recruitment and activation of PLCgamma1 in T cells: a new insight into old domains." EMBO J **25**(4): 774-784.

Bray, D. and C. Thomas (1976). "Unpolymerized actin in fibroblasts and brain." J Mol Biol **105**(4): 527-544.

Bremer, A., C. Henn, et al. (1994). "Towards atomic interpretation of F-actin filament three-dimensional reconstructions." J Mol Biol **242**(5): 683-700.

Briggs, S. D. and T. E. Smithgall (1999). "SH2-kinase linker mutations release Hck tyrosine kinase and transforming activities in Rat-2 fibroblasts." J Biol Chem **274**(37): 26579-26583.

Brown, N. R., M. E. Noble, et al. (1999). "Effects of phosphorylation of threonine 160 on cyclin-dependent kinase 2 structure and activity." J Biol Chem **274**(13): 8746-8756.

Buck, M., W. Xu, et al. (2001). "Global disruption of the WASP autoinhibited structure on Cdc42 binding. Ligand displacement as a novel method for monitoring amide hydrogen exchange." Biochemistry **40**(47): 14115-14122.

Bustelo, X. R. (2001). "Vav proteins, adaptors and cell signaling." Oncogene **20**(44): 6372-6381.

Bustelo, X. R. and M. Barbacid (1992). "Tyrosine phosphorylation of the vav proto-oncogene product in activated B cells." Science **256**(5060): 1196-1199.

Bustelo, X. R., J. A. Ledbetter, et al. (1992). "Product of vav proto-oncogene defines a new class of tyrosine protein kinase substrates." Nature **356**(6364): 68-71.

Campellone, K. G., N. J. Webb, et al. (2008). "WHAMM is an Arp2/3 complex activator that binds microtubules and functions in ER to Golgi transport." Cell **134**(1): 148-161.

Carlier, M. F., V. Laurent, et al. (1997). "Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: implication in actin-based motility." J Cell Biol **136**(6): 1307-1322.

Carlier, M. F. and D. Pantaloni (1997). "Control of actin dynamics in cell motility." J Mol Biol **269**(4): 459-467.

Cassimeris, L., H. McNeill, et al. (1990). "Chemoattractant-stimulated polymorphonuclear leukocytes contain two populations of actin filaments that differ in their spatial distributions and relative stabilities." J Cell Biol **110**(4): 1067-1075.

CCP4 (Collaborative Computational Project, N. (1996). "The CCP4 suite: programs for protein crystallography." Acta Crystallogr. D Biol. Crystallogr(50): 760-763.

Cen, H., A. G. Papageorge, et al. (1992). "Isolation of multiple mouse cDNAs with coding homology to Saccharomyces cerevisiae CDC25: identification of a region related to Bcr, Vav, Dbl and CDC24." EMBO J **11**(11): 4007-4015.

Chang, F., D. Drubin, et al. (1997). "cdc12p, a protein required for cytokinesis in fission yeast, is a component of the cell division ring and interacts with profilin." J Cell Biol **137**(1): 169-182.

Chen, Z., D. Borek, et al. (2010). "Structure and control of the actin regulatory WAVE complex." Nature **468**(7323): 533-538.

Chereau, D., M. Boczkowska, et al. (2008). "Leiomodin is an actin filament nucleator in muscle cells." Science **320**(5873): 239-243.

Chereau, D., F. Kerff, et al. (2005). "Actin-bound structures of Wiskott-Aldrich syndrome protein (WASP)-homology domain 2 and the implications for filament assembly." Proc Natl Acad Sci U S A **102**(46): 16644-16649.

Chesarone, M. A. and B. L. Goode (2009). "Actin nucleation and elongation factors: mechanisms and interplay." Curr Opin Cell Biol **21**(1): 28-37.

Chrencik, J. E., A. Brooun, et al. (2008). "Structural basis of guanine nucleotide exchange mediated by the T-cell essential Vav1." J Mol Biol **380**(5): 828-843.

Clifton, D. R., C. A. Dooley, et al. (2005). "Tyrosine phosphorylation of the chlamydial effector protein Tarp is species specific and not required for recruitment of actin." Infect Immun **73**(7): 3860-3868.

Clifton, D. R., K. A. Fields, et al. (2004). "A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin." Proc Natl Acad Sci U S A **101**(27): 10166-10171.

Co, C., D. T. Wong, et al. (2007). "Mechanism of actin network attachment to moving membranes: barbed end capture by N-WASP WH2 domains." Cell **128**(5): 901-913.

Colon-Gonzalez, F. and M. G. Kazanietz (2006). "C1 domains exposed: from diacylglycerol binding to protein-protein interactions." Biochim Biophys Acta **1761**(8): 827-837.

Coppola, J., S. Bryant, et al. (1991). "Mechanism of activation of the vav protooncogene." Cell Growth Differ **2**(2): 95-105.

Cossart, P. and A. Toledo-Arana (2008). "Listeria monocytogenes, a unique model in infection biology: an overview." Microbes Infect **10**(9): 1041-1050.

Cowtan, K. (2006). "The Buccaneer software for automated model building. 1. Tracing protein chains." Acta Crystallogr D Biol Crystallogr **62**(Pt 9): 1002-1011.

Crespo, P., K. E. Schuebel, et al. (1997). "Phosphotyrosine-dependent activation of Rac-1 GDP/GTP exchange by the vav proto-oncogene product." Nature **385**(6612): 169-172.

Das, B., X. Shu, et al. (2000). "Control of intramolecular interactions between the pleckstrin homology and Dbl homology domains of Vav and Sos1 regulates Rac binding." J Biol Chem **275**(20): 15074-15081.

Davis, I. W., A. Leaver-Fay, et al. (2007). "MolProbity: all-atom contacts and structure validation for proteins and nucleic acids." Nucleic Acids Res **35**(Web Server issue): W375-383.

Delano, W. L. (2002). "The PyMOL Molecular Graphics System DeLano Scientific." San Carlos, CA, USA.

Dominguez, R. (2007). "The beta-thymosin/WH2 fold: multifunctionality and structure." Ann N Y Acad Sci **1112**: 86-94.

Dominguez, R. (2009). "Actin filament nucleation and elongation factors--structure-function relationships." Crit Rev Biochem Mol Biol **44**(6): 351-366.

Drenckhahn, D., Pollard, T.D. (1986). "Elongation of actin filaments is a diffusion-limited reaction at the barbed end and is accelerated by inert macromolecules." J. Biol. Chem. **261**: 12754-12758.

Ducka, A. M., P. Joel, et al. (2010). "Structures of actin-bound Wiskott-Aldrich syndrome protein homology 2 (WH2) domains of Spire and the implication for filament nucleation." Proc Natl Acad Sci U S A **107**(26): 11757-11762.

Eden, S., R. Rohatgi, et al. (2002). "Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck." Nature **418**(6899): 790-793.

Egelman, E. H., N. Francis, et al. (1982). "F-actin is a helix with a random variable twist." Nature **298**(5870): 131-135.

Egile, C., I. Rouiller, et al. (2005). "Mechanism of filament nucleation and branch stability revealed by the structure of the Arp2/3 complex at actin branch junctions." PLoS Biol **3**(11): e383.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr **60**(Pt 12 Pt 1): 2126-2132.

Evangelista, M., D. Pruyne, et al. (2002). "Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast." Nat Cell Biol **4**(3): 260-269.

Evans, G. A., O. M. Howard, et al. (1993). "Interleukin-2 induces tyrosine phosphorylation of the vav proto-oncogene product in human T cells: lack of requirement for the tyrosine kinase lck." Biochem J **294 (Pt 2)**: 339-342.

Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallogr D Biol Crystallogr **62**(Pt 1): 72-82.

Fackler, O. T., W. Luo, et al. (1999). "Activation of Vav by Nef induces cytoskeletal rearrangements and downstream effector functions." Mol Cell **3**(6): 729-739.

Faraldo-Gomez, J. D. and B. Roux (2007). "On the importance of a funneled energy landscape for the assembly and regulation of multidomain Src tyrosine kinases." Proc Natl Acad Sci U S A **104**(34): 13643-13648.

Feierbach, B. and F. Chang (2001). "Roles of the fission yeast formin for3p in cell polarity, actin cable formation and symmetric cell division." Curr Biol **11**(21): 1656-1665.

Fernandez-Zapico, M. E., N. C. Gonzalez-Paz, et al. (2005). "Ectopic expression of VAV1 reveals an unexpected role in pancreatic cancer tumorigenesis." Cancer Cell **7**(1): 39-49.

Fernandez, C. and G. Wider (2003). "TROSY in NMR studies of the structure and function of large biological macromolecules." Curr Opin Struct Biol **13**(5): 570-580.

Fischer, R. S. and V. M. Fowler (2003). "Tropomodulins: life at the slow end." Trends Cell Biol **13**(11): 593-601.

Forscher, P. and S. J. Smith (1988). "Actions of cytochalasins on the organization of actin filaments and microtubules in a neuronal growth cone." J Cell Biol **107**(4): 1505-1516.

Fu, Y. and J. E. Galan (1999). "A salmonella protein antagonizes Rac-1 and Cdc42 to mediate host-cell recovery after bacterial invasion." Nature **401**(6750): 293-297.

Galkin, V. E., A. Orlova, et al. (2001). "Actin depolymerizing factor stabilizes an existing state of F-actin and can change the tilt of F-actin subunits." J Cell Biol **153**(1): 75-86.

Galland, F., S. Katzav, et al. (1992). "The products of the mcf-2 and vav proto-oncogenes and of the yeast gene cdc-24 share sequence similarities." Oncogene **7**(3): 585-587.

Goley, E. D., T. Ohkawa, et al. (2006). "Dynamic nuclear actin assembly by Arp2/3 complex and a baculovirus WASP-like protein." Science **314**(5798): 464-467.

Gouin, E., M. D. Welch, et al. (2005). "Actin-based motility of intracellular pathogens." Curr Opin Microbiol **8**(1): 35-45.

Graceffa, P. and R. Dominguez (2003). "Crystal structure of monomeric actin in the ATP state. Structural basis of nucleotide-dependent actin dynamics." J Biol Chem **278**(36): 34172-34180.

Groysman, M., M. Nagano, et al. (1998). "Mutagenic analysis of Vav reveals that an intact SH3 domain is required for transformation." Oncogene **17**(12): 1597-1606.

Groysman, M., C. S. Russek, et al. (2000). "Vav, a GDP/GTP nucleotide exchange factor, interacts with GDIs, proteins that inhibit GDP/GTP dissociation." FEBS Lett **467**(1): 75-80.

Gulbins, E., K. M. Coggeshall, et al. (1993). "Tyrosine kinase-stimulated guanine nucleotide exchange activity of Vav in T cell activation." Science **260**(5109): 822-825.

Han, J., B. Das, et al. (1997). "Lck regulates Vav activation of members of the Rho family of GTPases." Mol Cell Biol **17**(3): 1346-1353.

Han, J., K. Luby-Phelps, et al. (1998). "Role of substrates and products of PI 3-kinase in regulating activation of Rac-related guanosine triphosphatases by Vav." Science **279**(5350): 558-560.

Hantschel, O., B. Nagar, et al. (2003). "A myristoyl/phosphotyrosine switch regulates c-Abl." Cell **112**(6): 845-857.

Hart, M. J., A. Eva, et al. (1991). "Catalysis of guanine nucleotide exchange on the CDC42Hs protein by the dbl oncogene product." Nature **354**(6351): 311-314.

Henzler-Wildman, K. and D. Kern (2007). "Dynamic personalities of proteins." Nature **450**(7172): 964-972.

Heo, J., R. Thapar, et al. (2005). "Recognition and activation of Rho GTPases by Vav1 and Vav2 guanine nucleotide exchange factors." Biochemistry **44**(17): 6573-6585.
Higgs, H. N. (2005). "Formin proteins: a domain-based approach." Trends Biochem Sci **30**(6): 342-353.
Hoffman, G. R. and R. A. Cerione (2002). "Signaling to the Rho GTPases: networking with the DH domain." FEBS Lett **513**(1): 85-91.
Holmes, K. C., D. Popp, et al. (1990). "Atomic model of the actin filament." Nature **347**(6288): 44-49.
Hornstein, I., A. Alcover, et al. (2004). "Vav proteins, masters of the world of cytoskeleton organization." Cell Signal **16**(1): 1-11.
Huse, M. and J. Kuriyan (2002). "The conformational plasticity of protein kinases." Cell **109**(3): 275-282.
Irobi, E., A. H. Aguda, et al. (2004). "Structural basis of actin sequestration by thymosin-beta4: implications for WH2 proteins." EMBO J **23**(18): 3599-3608.
Jewett, T. J., E. R. Fischer, et al. (2006). "Chlamydial TARP is a bacterial nucleator of actin." Proc Natl Acad Sci U S A **103**(42): 15599-15604.
Kabsch, W., H. G. Mannherz, et al. (1990). "Atomic structure of the actin:DNase I complex." Nature **347**(6288): 37-44.
Kaiser, D. A., V. K. Vinson, et al. (1999). "Profilin is predominantly associated with monomeric actin in Acanthamoeba." J Cell Sci **112 (Pt 21)**: 3779-3790.
Katzav, S. (2009). "Vav1: a hematopoietic signal transduction molecule involved in human malignancies." Int J Biochem Cell Biol **41**(6): 1245-1248.
Katzav, S., J. L. Cleveland, et al. (1991). "Loss of the amino-terminal helix-loop-helix domain of the vav proto-oncogene activates its transforming potential." Mol Cell Biol **11**(4): 1912-1920.
Katzav, S., D. Martin-Zanca, et al. (1989). "vav, a novel human oncogene derived from a locus ubiquitously expressed in hematopoietic cells." EMBO J **8**(8): 2283-2290.
Khosravi-Far, R., M. Chrzanowska-Wodnicka, et al. (1994). "Dbl and Vav mediate transformation via mitogen-activated protein kinase pathways that are distinct from those activated by oncogenic Ras." Mol Cell Biol **14**(10): 6848-6857.
Kim, A. S., L. T. Kakalis, et al. (2000). "Autoinhibition and activation mechanisms of the Wiskott-Aldrich syndrome protein." Nature **404**(6774): 151-158.
Kim, E., W. Wriggers, et al. (2000). "Cross-linking constraints on F-actin structure." J Mol Biol **299**(2): 421-429.
Korzhnev, D. M., K. Kloiber, et al. (2004). "Multiple-quantum relaxation dispersion NMR spectroscopy probing millisecond time-scale dynamics in proteins: theory and application." J Am Chem Soc **126**(23): 7320-7329.
Kostyukova, A. S. (2007). "Leiomodin/tropomyosin interactions are isoform specific." Arch Biochem Biophys **465**(1): 227-230.
Kovar, D. R. (2006). "Molecular details of formin-mediated actin assembly." Curr Opin Cell Biol **18**(1): 11-17.

Kovar, D. R., E. S. Harris, et al. (2006). "Control of the assembly of ATP- and ADP-actin by formins and profilin." Cell **124**(2): 423-435.
Kovar, D. R. and T. D. Pollard (2004). "Insertional assembly of actin filament barbed ends in association with formins produces piconewton forces." Proc Natl Acad Sci U S A **101**(41): 14725-14730.
Krieger, I., A. Kostyukova, et al. (2002). "Crystal structure of the C-terminal half of tropomodulin and structural basis of actin filament pointed-end capping." Biophys J **83**(5): 2716-2725.
Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." J Mol Biol **372**(3): 774-797.
Kueh, H. Y., W. M. Brieher, et al. (2008). "Dynamic stabilization of actin filaments." Proc Natl Acad Sci U S A **105**(43): 16531-16536.
Lazer, G., L. Pe'er, et al. "Tyrosine residues at the carboxyl terminus of Vav1 play an important role in regulation of its biological activity." J Biol Chem **285**(30): 23075-23085.
Levinson, D. F. (2006). "The genetics of depression: a review." Biol Psychiatry **60**(2): 84-92.
Li, P., I. R. Martins, et al. (2008). "Internal dynamics control activation and activity of the autoinhibited Vav DH domain." Nat Struct Mol Biol **15**(6): 613-618.
Lietha, D., X. Cai, et al. (2007). "Structural basis for the autoinhibition of focal adhesion kinase." Cell **129**(6): 1177-1187.
Lim, W. A. (2010). "Designing customized cell signalling circuits." Nat Rev Mol Cell Biol **11**(6): 393-403.
Linardopoulou, E. V., S. S. Parghi, et al. (2007). "Human subtelomeric WASH genes encode a new subclass of the WASP family." PLoS Genet **3**(12): e237.
Liu, R., M. T. Abreu-Blanco, et al. (2009). "Wash functions downstream of Rho and links linear and branched actin nucleation factors." Development **136**(16): 2849-2860.
Liverman, A. D., H. C. Cheng, et al. (2007). "Arp2/3-independent assembly of actin by Vibrio type III effector VopL." Proc Natl Acad Sci U S A **104**(43): 17117-17122.
Llorca, O., E. Arias-Palomo, et al. (2005). "Global conformational rearrangements during the activation of the GDP/GTP exchange factor Vav3." EMBO J **24**(7): 1330-1340.
Lopez-Lago, M., H. Lee, et al. (2000). "Tyrosine phosphorylation mediates both activation and downmodulation of the biological activity of Vav." Mol Cell Biol **20**(5): 1678-1691.
Lu, J., W. Meng, et al. (2007). "Structure of the FH2 domain of Daam1: implications for formin regulation of actin assembly." J Mol Biol **369**(5): 1258-1269.
Machesky, L. M., S. J. Atkinson, et al. (1994). "Purification of a cortical complex containing two unconventional actins from Acanthamoeba by affinity chromatography on profilin-agarose." J Cell Biol **127**(1): 107-115.
Mackay, D. J. and A. Hall (1998). "Rho GTPases." J Biol Chem **273**(33): 20685-20688.

Manseau, L. J. and T. Schupbach (1989). "cappuccino and spire: two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the Drosophila embryo." Genes Dev **3**(9): 1437-1452.

Margolis, B., P. Hu, et al. (1992). "Tyrosine phosphorylation of vav proto-oncogene product containing SH2 domain and transcription factor motifs." Nature **356**(6364): 71-74.

Martinerie, C., L. A. Cannizzaro, et al. (1990). "The human VAV proto-oncogene maps to chromosome region 19p12----19p13.2." Hum Genet **86**(1): 65-68.

Masterson, L. R., A. Mascioni, et al. (2008). "Allosteric cooperativity in protein kinase A." Proc Natl Acad Sci U S A **105**(2): 506-511.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2007). "Phaser crystallographic software." J Appl Crystallogr **40**(Pt 4): 658-674.

Miletic, A. V., D. B. Graham, et al. (2007). "Vav proteins control MyD88-dependent oxidative burst." Blood **109**(8): 3360-3368.

Minor, W., M. Cymborowski, et al. (2006). "HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes." Acta Crystallogr D Biol Crystallogr **62**(Pt 8): 859-866.

Moarefi, I., M. LaFevre-Bernt, et al. (1997). "Activation of the Src-family tyrosine kinase Hck by SH3 domain displacement." Nature **385**(6617): 650-653.

Morris, R. J., P. H. Zwart, et al. (2004). "Breaking good resolutions with ARP/wARP." J Synchrotron Radiat **11**(Pt 1): 56-59.

Moseley, J. B., I. Sagot, et al. (2004). "A conserved mechanism for Bni1- and mDia1-induced actin assembly and dual regulation of Bni1 by Bud6 and profilin." Mol Biol Cell **15**(2): 896-907.

Mosteller, R., J. Han, et al. (2000). "Biochemical analysis of regulation of Vav, a guanine-nucleotide exchange factor for Rho family of GTPases." Methods Enzymol **325**: 38-51.

Movilla, N. and X. R. Bustelo (1999). "Biological and regulatory properties of Vav-3, a new member of the Vav family of oncoproteins." Mol Cell Biol **19**(11): 7870-7885.

Mullins, R. D., J. A. Heuser, et al. (1998). "The interaction of Arp2/3 complex with actin: nucleation, high affinity pointed end capping, and formation of branching networks of filaments." Proc Natl Acad Sci U S A **95**(11): 6181-6186.

Mullins, R. D., J. F. Kelleher, et al. (1998). "Arp2/3 complex from Acanthamoeba binds profilin and cross-links actin filaments." Mol Biol Cell **9**(4): 841-852.

Murakami, K., T. Yasunaga, et al. (2010). "Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release." Cell **143**(2): 275-287.

Nolen, B. J. and T. D. Pollard (2007). "Insights into the influence of nucleotides on actin family proteins from seven structures of Arp2/3 complex." Mol Cell **26**(3): 449-457.

Nolz, J. C., T. S. Gomez, et al. (2005). "The Ezh2 methyltransferase complex: actin up in the cytosol." Trends Cell Biol **15**(10): 514-517.

Oda, T., M. Iwasa, et al. (2009). "The nature of the globular- to fibrous-actin transition." Nature **457**(7228): 441-445.

Orlova, A. and E. H. Egelman (1993). "A conformational change in the actin subunit can change the flexibility of the actin filament." J Mol Biol **232**(2): 334-341.

Orlova, A., A. Shvetsov, et al. (2004). "Actin-destabilizing factors disrupt filaments by means of a time reversal of polymerization." Proc Natl Acad Sci U S A **101**(51): 17664-17668.

Otomo, T., C. Otomo, et al. (2005). "Structural basis of Rho GTPase-mediated activation of the formin mDia1." Mol Cell **18**(3): 273-281.

Otomo, T., D. R. Tomchick, et al. (2005). "Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain." Nature **433**(7025): 488-494.

Otterbein, L. R., P. Graceffa, et al. (2001). "The crystal structure of uncomplexed actin in the ADP state." Science **293**(5530): 708-711.

Otwinowski, Z., Ed. (2001). Isomorphous Replacement and Anomalous Scattering Proceedings of CCP4 Study Weekend Warrington, UK.

Paccani, S. R., M. Boncristiano, et al. (2005). "Defective Vav expression and impaired F-actin reorganization in a subset of patients with common variable immunodeficiency characterized by T-cell defects." Blood **106**(2): 626-634.

Padrick, S. B., H. C. Cheng, et al. (2008). "Hierarchical regulation of WASP/WAVE proteins." Mol Cell **32**(3): 426-438.

Palmby, T. R., K. Abe, et al. (2002). "Critical role of the pleckstrin homology and cysteine-rich domains in Vav signaling and transforming activity." J Biol Chem **277**(42): 39350-39359.

Pantaloni, D., Carlier, M.F. (1993). "How profilin promotes actin filament assembly in the presence of thymosin b4." Cell **75**: 1007-1014.

Pawson, T. and M. Kofler (2009). "Kinome signaling through regulated protein-protein interactions in normal and cancer cells." Curr Opin Cell Biol **21**(2): 147-153.

Petersen, J., O. Nielsen, et al. (1998). "FH3, a domain found in formins, targets the fission yeast formin Fus1 to the projection tip during conjugation." J Cell Biol **141**(5): 1217-1228.

Pluskey, S., T. J. Wandless, et al. (1995). "Potent stimulation of SH-PTP2 phosphatase activity by simultaneous occupancy of both SH2 domains." J Biol Chem **270**(7): 2897-2900.

Pollard, T. D., Ed. (1999). Introduction to actin and actin-binding proteins. Guidebook to the cytoskeletcal and motor proteins.

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

Pollard, T. D., L. Blanchoin, et al. (2000). "Molecular mechanisms controlling actin filament dynamics in nonmuscle cells." Annu Rev Biophys Biomol Struct **29**: 545-576.

Pollard, T. D. and G. G. Borisy (2003). "Cellular motility driven by assembly and disassembly of actin filaments." Cell **112**(4): 453-465.

Prehoda, K. E., J. A. Scott, et al. (2000). "Integration of multiple signals through cooperative regulation of the N-WASP-Arp2/3 complex." Science **290**(5492): 801-806.

Pruyne, D., M. Evangelista, et al. (2002). "Role of formins in actin assembly: nucleation and barbed-end association." Science **297**(5581): 612-615.

Pufall, M. A. and B. J. Graves (2002). "Autoinhibitory domains: modular effectors of cellular regulation." Annu Rev Cell Dev Biol **18**: 421-462.

Quinlan, M. E., J. E. Heuser, et al. (2005). "Drosophila Spire is an actin nucleation factor." Nature **433**(7024): 382-388.

Quinlan, M. E., S. Hilgert, et al. (2007). "Regulatory interactions between two actin nucleators, Spire and Cappuccino." J Cell Biol **179**(1): 117-128.

Quinlan, M. E. and E. Kerkhoff (2008). "Actin nucleation: bacteria get in-Spired." Nat Cell Biol **10**(1): 13-15.

Raberger, J., N. Boucheron, et al. (2008). "Impaired T-cell development in the absence of Vav1 and Itk." Eur J Immunol **38**(12): 3530-3542.

Ramos-Morales, F., B. J. Druker, et al. (1994). "Vav binds to several SH2/SH3 containing proteins in activated lymphocytes." Oncogene **9**(7): 1917-1923.

Rapley, J., V. L. Tybulewicz, et al. (2008). "Crucial structural role for the PH and C1 domains of the Vav1 exchange factor." EMBO Rep **9**(7): 655-661.

Rebowski, G., M. Boczkowska, et al. (2008). "X-ray scattering study of actin polymerization nuclei assembled by tandem W domains." Proc Natl Acad Sci U S A **105**(31): 10785-10790.

Rebowski, G., S. Namgoong, et al. (2010). "Structure of a longitudinal actin dimer assembled by tandem w domains: implications for actin filament nucleation." J Mol Biol **403**(1): 11-23.

Reisler, E. and E. H. Egelman (2007). "Actin structure and function: what we still do not understand." J Biol Chem **282**(50): 36133-36137.

Renault, L., B. Bugyi, et al. (2008). "Spire and Cordon-bleu: multifunctional regulators of actin dynamics." Trends Cell Biol **18**(10): 494-504.

Richards, T. A. and T. Cavalier-Smith (2005). "Myosin domain evolution and the primary divergence of eukaryotes." Nature **436**(7054): 1113-1118.

Rittling, S. R. and A. F. Chambers (2004). "Role of osteopontin in tumour progression." Br J Cancer **90**(10): 1877-1881.

Robinson, R. C., K. Turbedsky, et al. (2001). "Crystal structure of Arp2/3 complex." Science **294**(5547): 1679-1684.

Romero, F. and S. Fischer (1996). "Structure and function of vav." Cell Signal **8**(8): 545-553.

Romero, S., C. Le Clainche, et al. (2004). "Formin is a processive motor that requires profilin to accelerate actin assembly and associated ATP hydrolysis." Cell **119**(3): 419-429.

Rosales-Nieves, A. E., J. E. Johndrow, et al. (2006). "Coordination of microtubule and microfilament dynamics by Drosophila Rho1, Spire and Cappuccino." Nat Cell Biol **8**(4): 367-376.

Rottner, K., T. E. Stradal, et al. (2005). "Bacteria-host-cell interactions at the plasma membrane: stories on actin cytoskeleton subversion." Dev Cell **9**(1): 3-17.

Rould, M. A., Q. Wan, et al. (2006). "Crystal structures of expressed non-polymerizable monomeric actin in the ADP and ATP states." J Biol Chem **281**(42): 31909-31919.

Sagot, I., A. A. Rodal, et al. (2002). "An actin nucleation mechanism mediated by Bni1 and profilin." Nat Cell Biol **4**(8): 626-631.

Sauzeau, V., M. Jerkic, et al. (2007). "Loss of Vav2 proto-oncogene causes tachycardia and cardiovascular disease in mice." Mol Biol Cell **18**(3): 943-952.

Sauzeau, V., M. A. Sevilla, et al. (2006). "Vav3 proto-oncogene deficiency leads to sympathetic hyperactivity and cardiovascular dysfunction." Nat Med **12**(7): 841-845.

Schachter, J., Ed. (1999). in Chlamydia; Intracellular Biology, Pathogenesis, and Immunity, Am Soc Microbiol Press, Washington, DC.

Schapira, V., G. Lazer, et al. (2006). "Osteopontin is an oncogenic Vav1- but not wild-type Vav1-responsive gene: implications for fibroblast transformation." Cancer Res **66**(12): 6183-6191.

Schindler, T., W. Bornmann, et al. (2000). "Structural mechanism for STI-571 inhibition of abelson tyrosine kinase." Science **289**(5486): 1938-1942.

Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." Acta Crystallogr D Biol Crystallogr **58**(Pt 10 Pt 2): 1772-1779.

Schoenenberger, C. A., N. Bischler, et al. (2002). "Actin's propensity for dynamic filament patterning." FEBS Lett **529**(1): 27-33.

Schuebel, K. E., X. R. Bustelo, et al. (1996). "Isolation and characterization of murine vav2, a member of the vav family of proto-oncogenes." Oncogene **13**(2): 363-371.

Schuebel, K. E., N. Movilla, et al. (1998). "Phosphorylation-dependent and constitutive activation of Rho proteins by wild-type and oncogenic Vav-2." EMBO J **17**(22): 6608-6621.

Schumacher, N., J. M. Borawski, et al. (2004). "Overlapping expression pattern of the actin organizers Spir-1 and formin-2 in the developing mouse nervous system and the adult brain." Gene Expr Patterns **4**(3): 249-255.

Sept, D. and J. A. McCammon (2001). "Thermodynamics and kinetics of actin filament nucleation." Biophys J **81**(2): 667-674.

Simenel, C., M. Van Troys, et al. (2000). "Structural requirements for thymosin beta4 in its contact with actin. An NMR-analysis of thymosin beta4 mutants in solution and correlation with their biological activity." Eur J Biochem **267**(12): 3530-3538.

Sondermann, H., S. M. Soisson, et al. (2004). "Structural analysis of autoinhibition in the Ras activator Son of sevenless." Cell **119**(3): 393-405.

Songyang, Z. and L. C. Cantley (2004). "ZIP codes for delivering SH2 domains." Cell **116**(2 Suppl): S41-43, 42 p following S48.

Songyang, Z., S. E. Shoelson, et al. (1993). "SH2 domains recognize specific phosphopeptide sequences." Cell **72**(5): 767-778.

Steinmetz, M. O., K. N. Goldie, et al. (1997). "A correlative analysis of actin filament assembly, structure, and dynamics." J Cell Biol **138**(3): 559-574.

Tam, V. C., D. Serruto, et al. (2007). "A type III secretion system in Vibrio cholerae translocates a formin/spire hybrid-like actin nucleator to promote intestinal colonization." Cell Host Microbe **1**(2): 95-107.

Tam, V. C., M. Suzuki, et al. (2010). "Functional Analysis of VopF Activity Required for Colonization in Vibrio cholerae." MBio **1**(5).

Tamas, P., Z. Solti, et al. (2003). "Mechanism of epidermal growth factor regulation of Vav2, a guanine nucleotide exchange factor for Rac." J Biol Chem **278**(7): 5163-5171.

Theriot, J. A. and T. J. Mitchison (1991). "Actin microfilament dynamics in locomoting cells." Nature **352**(6331): 126-131.

Tugarinov, V., P. M. Hwang, et al. (2003). "Cross-correlated relaxation enhanced 1H[bond]13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes." J Am Chem Soc **125**(34): 10420-10428.

Tybulewicz, V. L. (2005). "Vav-family proteins in T-cell signalling." Curr Opin Immunol **17**(3): 267-274.

Tybulewicz, V. L., L. Ardouin, et al. (2003). "Vav1: a key signal transducer downstream of the TCR." Immunol Rev **192**: 42-52.

Vajpai, N., A. Strauss, et al. (2008). "Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib." J Biol Chem **283**(26): 18292-18302.

Van Aelst, L. and C. D'Souza-Schorey (1997). "Rho GTPases and signaling networks." Genes Dev **11**(18): 2295-2322.

Vavylonis, D., D. R. Kovar, et al. (2006). "Model of formin-associated actin filament elongation." Mol Cell **21**(4): 455-466.

Vicente-Manzanares, M., A. Cruz-Adalia, et al. (2005). "Control of lymphocyte shape and the chemotactic response by the GTP exchange factor Vav." Blood **105**(8): 3026-3034.

Vogtherr, M., K. Saxena, et al. (2006). "NMR characterization of kinase p38 dynamics in free and ligand-bound forms." Angew Chem Int Ed Engl **45**(6): 993-997.

Wang, Y. L. (1985). "Exchange of actin subunits at the leading edge of living fibroblasts: possible role of treadmilling." J Cell Biol **101**(2): 597-602.

Wilson, R., R. Ainscough, et al. (1994). "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans." Nature **368**(6466): 32-38.

Wood, M. W., R. Rosqvist, et al. (1996). "SopE, a secreted protein of Salmonella dublin, is translocated into the target eukaryotic cell via a sip-dependent mechanism and promotes bacterial entry." Mol Microbiol **22**(2): 327-338.

Worthylake, D. K., K. L. Rossman, et al. (2000). "Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1." Nature **408**(6813): 682-688.

Xu, W., A. Doshi, et al. (1999). "Crystal structures of c-Src reveal features of its autoinhibitory mechanism." Mol Cell **3**(5): 629-638.

Xu, W., S. C. Harrison, et al. (1997). "Three-dimensional structure of the tyrosine kinase c-Src." Nature **385**(6617): 595-602.

Xu, Y., J. B. Moseley, et al. (2004). "Crystal structures of a Formin Homology-2 domain reveal a tethered dimer architecture." Cell **116**(5): 711-723.

Yabana, N. and M. Shibuya (2002). "Adaptor protein APS binds the NH2-terminal autoinhibitory domain of guanine nucleotide exchange factor Vav3 and augments its activity." Oncogene **21**(50): 7720-7729.

Yohe, M. E., K. Rossman, et al. (2008). "Role of the C-terminal SH3 domain and N-terminal tyrosine phosphorylation in regulation of Tim and related Dbl-family proteins." Biochemistry **47**(26): 6827-6839.

Yohe, M. E., K. L. Rossman, et al. (2007). "Auto-inhibition of the Dbl family protein Tim by an N-terminal helical motif." J Biol Chem **282**(18): 13813-13823.

Young, M. A., S. Gonfloni, et al. (2001). "Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation." Cell **105**(1): 115-126.

Yu, B., I. R. Martins, et al. (2010). "Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1." Cell **140**(2): 246-256.

Zeng, L., P. Sachdev, et al. (2000). "Vav3 mediates receptor protein tyrosine kinase signaling, regulates GTPase activity, modulates cell morphology, and induces cell transformation." Mol Cell Biol **20**(24): 9212-9224.

Zheng, L., A. Sjolander, et al. (1996). "Antibody-induced engagement of beta 2 integrins on adherent human neutrophils triggers activation of p21ras through tyrosine phosphorylation of the protooncogene product Vav." Proc Natl Acad Sci U S A **93**(16): 8431-8436.

Zheng, Y., D. Zangrilli, et al. (1996). "The pleckstrin homology domain mediates transformation by oncogenic dbl through specific intracellular targeting." J Biol Chem **271**(32): 19017-19020.

Zhou, Z., J. Yin, et al. (2007). "The calponin homology domain of Vav1 associates with calmodulin and is prerequisite to T cell antigen receptor-induced calcium release in Jurkat T lymphocytes." J Biol Chem **282**(32): 23737-23744.

Zuchero, J. B., A. S. Coutts, et al. (2009). "p53-cofactor JMY is a multifunctional actin nucleation factor." Nat Cell Biol **11**(4): 451-459.

Zugaza, J. L., M. A. Lopez-Lago, et al. (2002). "Structural determinants for the biological activity of Vav proteins." J Biol Chem **277**(47): 45377-45392.