Graduate Theses, Dissertations, and Problem Reports

2011

A New Member of the AFAP Family, AFAP1L1, Binds to Cortactin and Localizes to Invadosomes

Brandi Nicole Snyder
West Virginia University

Follow this and additional works at: https://researchrepository.wvu.edu/etd

A New Member of the AFAP Family, AFAP1L1, Binds to Cortactin and Localizes to Invadosomes

Brandi Nicole Snyder

Dissertation submitted to the School of Medicine at West Virginia University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Cancer Cell Biology

Peter Gannett, PhD, Chair
Jun Liu, PhD
Joan Olson, PhD
William Petros, PharmD
Jing Jie Yu, MD
Daniel C. Flynn, PhD, Mentor

Department of Cancer Cell Biology

Morgantown, West Virginia
2011

Keywords: AFAP1L1, AFAP1, XB130, podosome, invadopodia, invadosome, cortactin, dentate nucleus

ABSTRACT

A New Member of the AFAP Family, AFAP1L1,
Binds to Cortactin and Localizes to Invadosomes

Brandi Nicole Snyder

Cellular motility and invasion in normal cellular processes and disease states such as cancer are dependent upon the ability of a cell to efficiently interact with its microenvironment, rearrange its cytoskeleton and degrade tissue barriers for purposes of cell movement. The AFAP family of adaptor proteins, AFAP1, AFAP1L1 and AFAP1L2, integrates signals received from the microenvironment into coordinated cytoskeletal changes. While there have been many reports on the functions and binding partners of AFAP1 and AFAP1L2, this work aimed to determine the cellular location and function of newly discovered AFAP1L1. The overall amino acid and protein structures of AFAP family members were compared so as to determine similarities and differences as well as to propose an evolutionary link between all three family members. As AFAP1 and AFAP1L1 have been shown to be more closely related, studies focused on a detailed comparison of these two family members. AFAP1 and AFAP1L1 were shown to have similar cellular localization in the cell by associating with stress filaments and cortical actin and also showing localization to invadosomes. Immunohistochemistry demonstrated differential expression of AFAP1L1 in the brain, particularly surrounding the Purkinje neurons, granular cells and neurons of the dentate nucleus. Although AFAP1 is a well known cSrc binding partner and activator, AFAP1L1 was determined to be a binding partner for cortactin, possibly through the SH3 domain. As other AFAP family members have been shown to be increased in various cancers, AFAP1L1 expression levels are upregulated in a number of cancers, particularly neuroblastoma and glioblastoma. While the similar amino acid sequence and modular domain identifies AFAP1L1 as a previously undescribed member of the AFAP family, the ability of AFAP1L1 to interact with cortactin and localize to distinct areas of the brain implies that AFAP1L1 has unique functions separate from AFAP1 and AFAP1L2.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my mentor, Dr. Daniel C. Flynn for his unwavering support and constant enthusiasm throughout my graduate studies. His unending curiosity of and desire to advance the cancer field creates an environment in which I was able question and grow as a scientist. I am also grateful to Dr. Jess Cunnick and Dr. YoungJin Cho for their constant support and guidance throughout my graduate work.

I would like to express sincere gratitude to Dr. Yong Qian for his assistance and benevolence as I completed my doctoral work. I would also like to acknowledge the members of my dissertation committee, Dr. Peter Gannett, Dr. Jun Liu, Dr. Joan Olson, Dr. William Petros and Dr. Jing Jie Yu for their advice and suggestions during my development as a scientist. To all of the faculty and students in the Biomedical Sciences Graduate Program at West Virginia University, particularly the Cancer Cell Biology program, I am grateful for the honorable desire to see each student succeed in their endeavors.

Thank you to my fiance, William Talkington, for strengthening me and showing me unconditional love as I worked to fulfill my dreams. Thank you for being a friend to laugh with, a shoulder to cry on and an unwavering pillar of support.

Finally, I would like to thank my family. To my parents, William and Linda, and the rest of my support team, Jennifer, Rick, Kristen, Tyler, Preston, Vanessa, Aaron and Ethan, thank you for instilling in me a resolute work ethic and enduring spirit. Thank you for providing me with limitless love and a home where the door was always open. Thank you for never doubting that I would succeed, I could not have done this without you.

TABLE OF CONTENTS

LIST OF FIGURES

Unique Methods

Appendix A

GLOSSARY

+	basic residue
A	alanine
ABD	actin binding domain
ABD2	actin binding domain 2
AFAP	Actin Filament Associated Protein
AFAP1	Actin Filament Associated Protein 1/AFAP-110
AFAP1L1	Actin Filament Associated Protein 1-like 1
AFAP1L2	Actin Filament Associated Protein 1-like 2/XB130
AP-1	activating-protein 1
Arp2/3	Actin related protein 2/3
BSA	bovine serum albumin
C	celsius
C-	carboxy terminal
CCDC6	coiled-coil domain containing 6
CEF	chicken embryo fibroblasts
Crk	avian sarcoma virus CT10 oncogene homolog
CTTN	cortactin
D	aspartic acid
Da	daltons
DTT	dithiolthreitol
E	glutamic acid
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
EMT	epithelial to mesenchymal transition
ERK	extracellular signal regulated kinase
EST	expressed sequence tag
F	phenylalanine
F-	filamentous
FAK	focal adhesion kinase
G	glycine
GAP	GTPase activating protein
GEFs	guanine nucleotide exchange factors
GFP	green fluorescent protein
GST	glutathione S-transferase
GTPase	guanosine triphosphatases
HEK	human embryonic kidney
I	isoleucine

IL-8	interluekin-8
K	lysine
kDa	kilodalton
L	leucine
LB	luria broth
LPA	lysophosphatidic acid
LSB	Laemmli sample buffer
Lzip	leucine zipper
M	methionine
MAGUK	membrane-associated guanylate kinase
MEF	mouse embryo fibroblasts`
MMPs	matix metalloproteases
M_r	relative mobility
N	asparagine
N-	amino terminal
NMDA	*N*-methyl *D*-aspartate
OD	optical density
P	proline
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDBu	phorbol 12,13-dibutyrate
PH	pleckstrin homology
PI3K	phosphatidylinositol-3 kinase
PKC	protein Kinase C
PMSF	phenylmethanesulfonylfluoride
PSD-95	post synaptic density protein 95/SAP90
PTC	papillary thyroid carcinoma
Q	glutamine
RET/PTC	rearranged in transformation/papillary thyroid carcinoma
RISC	RNA-induced silencing complex
RNAi	RNA interference
S	serine
SD	substrate domain
SDS	sodium dodecyl sulfate
SFK	Src Family Kinases
SH2	Src homology 2
SH3	Src homology 3
shRNA	short hairpin RNA
siRNA	short interfering RNA
SNP	single nucleotide polymorphism

SPTA1	spectrin alpha chain erythroytic
SPTAN1	spectrin alpha chain non-erythrocytic
SRE	serum response element
TBST	tris buffered saline tween 20%
TPC1	thyroid papillary carcinoma
TRITC	tetramethylrodamine-isothiocyanate
V	valine
X	any amino acid
Y	tyrosine
Ψ	aliphatic residue

CHAPTER 1

Introduction and Review of Literature

I. Introduction and Significance

Cells have the ability to take instructional cues from their environment and convert them into a complex and intricate cascade of information directed by signaling proteins so as to perform a large array of cellular processes. At the heart of these signaling complexes are adaptor proteins, proteins with no intrinsic enzymatic activity themselves but which do have the ability to link proteins together by virtue of their protein binding domains. These protein binding domains, identified through consensus amino acid sequence, allow a single adaptor protein to interact with a variety of other proteins so as to create a sophisticated signaling complex (Pawson and Nash, 2003). The functions of these signaling complexes are diverse and include apoptosis, actin dynamics and cell cycle regulation, among many others (Duncan et al., 2010; Qian et al., 2000; Roberts et al., 2002). The Actin Filament-Associated Protein (AFAP) family of adaptor proteins is involved in signaling cascades that regulate cell shape, motility and cytoskeletal integrity and are associated with cSrc kinase activity, protein kinase C alpha (PKCa) activation and activation of the phosphatidylinositol-3 kinase (PI3K) pathway (Baisden et al., 2001b; Lodyga et al., 2009; Xu et al., 2007). As the discovery of multiple AFAP family members is a relatively new area of research, this literature review will serve as an in depth comparison of the AFAP family structure and function.

II. Actin Filaments and Cell Morphology

It is well known that many cells have the ability to migrate during processes such as embryonic development, wound closure, and inflammation (Ho et al., 2008; Le Clainche and Carlier, 2008; Parsons et al., 2010; Rathinam and Alahari, 2010). It is

through rearrangement of the actin cytoskeleton that cells are able to effectively respond to mechanical forces, extracellular matrix proteins or external cues (Parsons et al., 2010). The ability of a cell to remodel the cytoskeleton is important in normal cellular processes such as embryogenesis, immune cell motility and adhesion; however, cancer cells can also acquire the ability to undergo abnormal cell migration (Le Clainche and Carlier, 2008; Weed and Parsons, 2001). Mediated through a number of proteins such as the Rho family of small guanosine triphosphatases (GTPases), cells move towards a signal by first extending the cell body and creating a variety of adhesions to their substratum.

Rho Family Members

Rho family members Rho, Rac and Cdc42 are GTPases that are regulators of cell movement. Small GTPases are activated to GTP bound forms through guanine nucleotide exchange factors (GEFs) which exchange bound GDP for GTP and are inactivated to GDP bound forms through GTPase activating proteins (GAPs) which hydrolyze GTP. Activation of different Rho family members can lead to the formation of different actin structures within the cell. The activation of Rho is known to induce the formation of stress fibers, contractile bundles of anti-parallel cross-linked actin that have the ability to contract due to association with non-muscle myosin (Parsons et al., 2010). Rac activation is known to induce the formation of lamellipodia, a thin cytoplasmic protrusion at the leading edge of the cell that is 1-5μm in width. Lamellipodia contain branched actin structures and are vital to the formation of cellular adhesions while Cdc42 activation results in the formation of actin-rich finger-like protrusions which extend outside of the lamellipodium called filopodia (Small et al., 2002). Filopodia play multiple roles in the

cell by directing cell migration through sensing of the microenvironment, forming sheets of cells for both embryonic development and wound healing through cell adhesion and forming the precursors of dendritic spines in neurons (Mattila and Lappalainen, 2008). It is through activation of various Rho family GTPases and their subsequent signaling pathways that a cell can transform extra- and intracellular cues into a coordinated cell movement.

Focal Adhesions

As a cell begins the process of migration, it begins by forming transient contacts with the substratum called focal contacts (Zamir and Geiger, 2001). These immature nascent adhesions of bundled actin have the ability to mature into larger and longer lasting focal adhesions; a complex array of proteins along the ventral plasma membrane rich in cytoskeletal proteins such as paxillin and talin and protein kinases such as Src, focal adhesion kinase (FAK) and protein kinase C (PKC), as well as many others that link the actin cytoskeleton to the substratum (Zamir and Geiger, 2001). Key players in focal adhesions are integrins, transmembrane proteins whose cytoplasmic tails bind to the array of proteins in the focal adhesion and whose extracellular domains bind to proteins in the extracellular matrix (ECM) such as fibronectin and collagen (Parsons et al., 2010). The fast growing filamentous actin lamellipodium at the protruding edge of the cell undergoes actin polymerization that pushes the lamellipodium forward, forming immature focal contacts with the extracellular matrix. These early contacts mature into integrin-containing focal adhesions through mechanical tension applied to the contact by myosin-II mediated contractile forces from the cytoskeleton. The application of force at a

maturing contact establishes new binding sites for adhesion-associated proteins through conformational changes in proteins involved in the immature contact. An additional result of myosin-II activation is cross-linking and bundling of actin filaments, combined with myosin-II mediated contractility results in the maturation of focal adhesions at the leading edge of the cell which signals for actin filament polymerization and protrusion of the leading edge while this protrusive force causes actin filament depolymerization and disassembly of the focal adhesions at the rear of the cell (Parsons et al., 2010; Ridley et al., 2003; Small and Resch, 2005). The forward movement of the cell is mediated by these contractile forces between the cellular adhesions and stress fibers. Three types of stress fibers have been described: ventral, dorsal and transverse arcs (Pellegrin and Mellor, 2007). Ventral stress fibers extend between focal adhesions at the leading edge of the cell to the central cell body (Small et al., 1998). Contraction of these ventral stress fibers causes a pulling towards the newly formed lamellipodium and away from the trailing edge of the cell. Dorsal stress fibers are attached at one end to a focal adhesion and rise into the dorsal surface of the cell where they interact with a meshwork of actin. Although they do not have the ability to contract themselves, dorsal stress fibers may function to transmit the contraction of transverse arcs which are not tethered to focal adhesions and form behind the lamellipodium (Pellegrin and Mellor, 2007).

Invadosomes

In addition to focal adhesions, some highly motile cells such as macrophages, osteoclasts, dendritic cells and also smooth muscle cells exhibit podosomes, dynamic cellular contacts with the substratum that consist of an F-actin core surrounded by a ring

of proteins involved in cytoskeletal organization such as actin-related protein (Arp)2/3, cortactin, integrins, cSrc and many others. Found on the ventral membrane of cells, podosomes have a half-life longer than focal contacts ranging from approximately 2 to 12 minutes and form a structure approximately 0.5-1μm in diameter with a depth of 0.2-0.4μm (Linder and Kopp, 2005). Although focal adhesions and podosomes share some of their integral proteins such as protein kinases, RhoGPTases and integrins, podosomes have the ability to secrete matrix metalloproteases (MMPs) which focal adhesions do not (Linder and Aepfelbacher, 2003). The ability of podosomes to secrete MMPs allows them to remodel the surrounding extracellular matrix for purposes of motility and crossing of tissue barriers (Linder and Kopp, 2005). Similar to podosomes are invadopodia, podosome-like structures found in transformed cells (cancer cells) which are longer lived (minutes to hours) and actively degrade matrix, which is required for tumor invasion (Yamaguchi et al., 2006). Invadopodia are larger than podosomes, reaching diameters of up to 8μm and extending projections deep into the extracellular matrix (Linder and Kopp, 2005). A well-known feature of transformed cells is epithelial to mesenchymal transition (EMT) which involves the detachment of tumor cells from their place of origin, invasion through their surroundings, intravasation into and extravasation from the blood and lymphatic vessels and colonization of a secondary metastatic site in which this temporary cell phenotype is reversed (Yilmaz and Christofori, 2009). Naturally, loss of cellular contacts and rearrangement of the actin cytoskeleton are dynamic activities in the EMT process. Invadopodia play a role in EMT by degrading the extracellular matrix and providing a pathway for invasion of malignant cells.

Collectively termed invadosomes for their ability to degrade components of the extracellular matrix, podosomes and invadopodia share some common characteristics while there are also identifiable differences between the two (Linder, 2009). It is hypothesized that podosomes may be a precursor to invadopodia, and, as such, both contain similar cellular location and machinery such as their dot-like shape and localization to the ventral membrane as well as their association with actin and other focal complex regulators (Linder, 2009; Saltel et al., 2010). While many podosomes may form in a single cell and have a lifetime of a few minutes, fewer and longer-lived invadopodia are usually found in transformed cells (Linder, 2009). Podosomes are commonly characterized as adhesive structures that may play a role in migration and ECM remodeling while also serving a sensory function for guidance (Gimona et al., 2008; Saltel et al., 2010). Invadopodia have an increased ability to degrade the ECM and are a hallmark of highly invasive cancers as it is by virtue of this matrix degradation that highly motile cancer cells are able to metastasize throughout the body. The origin of and an *in vivo* role for invadosomes is currently a subject of exploration (Gimona et al., 2008; Linder, 2009).

cSrc and the Actin Cytoskeleton

cSrc is a well-known and well-studied non-receptor protein tyrosine kinase whose activation affects cellular transformation through the tyrosine phosphorylation of cellular proteins thus resulting in a transformed phenotype in which the cells are rounded, lose their integrin-based substrate attachment, become more motile and can degrade and invade through cellular matrix (Frame et al., 2002; Yeatman, 2004). The cSrc structure

consists of a kinase domain, a Src homology 2 (SH2) domain, a Src homology 3 (SH3) domain and a myristoylation site that is necessary for membrane localization (Yeatman, 2004). cSrc remains in an intramolecularly autoinhibited conformation by phosphorylation of a tyrosine residue at Y527 in chicken (Y530 in human) that interacts with its' own SH2 domain. An additional interaction of the SH3 domain with the kinase domain acts as a second level of intramolecular regulation to insure the proper functioning of cSrc (Yeatman, 2004). As activation of cSrc is a hallmark of cancer and results in a transformed cellular phenotype, this tight regulation of activation is necessary for the proper functioning of the cell. Phosphatases that dephosphorylate the regulatory tyrosine in the C-terminus are also highly regulated as dephosphorylation of phospho-Y527 (Y530 in humans) allows cSrc to open into an active conformation and autophosphorylate itself and cellular substrates. While inactive cSrc resides around the perinuclear region, active cSrc indirectly associates with actin and can move to the cell periphery where the myristoylation site is necessary for association with the plasma membrane. From here, cSrc has the ability to phosphorylate a variety of proteins. FAK, for example, is a non-receptor tyrosine kinase that associates with integrins and is activated when phosphorylated by cSrc which results in a loss of and turnover of focal adhesions. This turnover results in the regulated process of assembly and disassembly of cell-matrix adhesions and allows the cell to increase its motility. Also a result of cSrc activation, cells develop the ability to invade through matrix by the activation of MMPs (Yeatman, 2004).

Classical Protein Kinase C and the Cytoskeleton

The Protein Kinase C family are serine/threonine kinases divided into classical (PKC a,bI,bII (splice variants from a single gene) and g), novel (PKC d,e,h and q) and atypical (PKC i/l and x) designations which each contain a catalytic and regulatory domain (Larsson, 2006). The classical PKCs have a necessary activation site that can bind phosphatidylserine, diacylglycerol and phorbol esters, a calcium binding region, an ATP binding region and a catalytic domain (Brandt et al., 2002; Martiny-Baron and Fabbro, 2007). While PKCb and PKCg are known to have roles in inflammation, angiogenesis and neuronal tissues, the PKCa isoform is a general promoter of cell migration, proliferation and apoptosis (Larsson, 2006; Martiny-Baron and Fabbro, 2007). In A7r5 smooth muscle cells, classical PKCs, PKCa in particular, were shown to be responsible for phorbol ester-induced cytoskeletal remodeling of stress fibers into podosomes (Hai et al., 2002). Activation of PKCs by phorbol ester in these cells resulted in a decrease in RhoA activity and subsequent dissolution of stress fibers by virtue of the upregulation of p190RhoGAP (Brandt et al., 2002). P190RhoGAP is activated by phosphorylation, possibly by cSrc. Indeed, cSrc activity is also increased upon stimulation of A7r5 cells with phorbol ester. It is hypothesized that activation of classical PKC family members, PKCa in particular in A7r5 cells, can result in the activation of cSrc which in turn phosphorylates and activates p190RhoGAP which results in the downregulation of Rho, the dissolution of stress fibers and the formation of podosome-like structures (Brandt et al., 2002; Fincham et al., 1999). In addition, the activation of cSrc can also have a profound effect on the phenotype and motility of the cell.

In addition to its role in cell shape, motility and, in the case of cancer cells, invasion, the actin cytoskeleton also plays a role in cell signaling, transport and cell division, among others. It is therefore through highly conserved and concerted pathways involving the actin cytoskeleton that a cell is able to perform all of its necessary processes.

III. The AFAP Family: A Historical Perspective

The first AFAP family member discovered was described as a tyrosine phosphorylated protein of approximately 110 kilodaltons (kDa) which could co-immunoprecipitate with active cSrc and was referred to as pp110 for phosphoprotein of 110kDa (Kanner et al., 1991; Reynolds et al., 1989). Although cells expressing active cSrc were known to have a transformed phenotype demonstrated by a loss of focal adhesions and bundled actin filaments, as well as an increase in motility and invasion (Frame et al., 2002), it was unknown through which proteins these signals were relayed. While overexpression of cSrc is not enough to induce such a transformed phenotype, a mutated form of activated avian Src, cSrc^{527F}, contains a mutation in the C-terminus regulatory tyrosine which disrupts internal regulation by changing this tyrosine to a phenylalanine. In the cSrc^{527F} mutant, cSrc is not able to be held in its intramolecular inhibition state and is thus constitutively activated. To determine proteins involved in cSrc cellular transformation, cSrc^{527F} was transfected into chicken embryo fibroblasts (CEF), immunoprecipitated using the cSrc-specific antibody EC10 and immune complexes were analyzed through the use of phosphotyrosine antibodies leading to the discovery of two tyrosine phosphorylated proteins, pp130 (later known as pp130Cas) and

pp110, that interacted directly with active cSrc (Reynolds et al., 1989). Expression of deletion variants of cSrc in CEF cells in which amino acids 155-157 of the SH2 domain and amino acids 92-95 were removed confirmed the phosphorylated proteins involvement in binding and activation of the cSrc molecule. Monoclonal antibodies were raised against pp110 (Kanner et al., 1990). Antibody 4C3 against pp110 was used to screen cDNA λgt11 peptide and λgt10 cDNA expression libraries derived from chicken embryo brains and chick embryo fibroblasts, thus resulting in the identification of various overlapping cDNA sequences which were analyzed so as to construct the overall cDNA coding sequence of pp110 (Flynn et al., 1993). A rabbit polyclonal antibody created using this cDNA, F1, identified pp110's association with stress fibers and cortical actin in chick embryo cells, thus pp110 was named Actin Filament Associated Protein of 110kDa, hereinafter known as AFAP1 (Flynn et al., 1993). The human homologue of chicken AFAP1 is also associated with cortical actin and stress fibers. Human AFAP1 is found at chromosomal location 4p16.1 and contains 16 exons which encode 730 amino acids with an approximate molecular mass of 80,725 daltons (Da) as identified by the Ensembl and UniProt databases. A second alternatively spliced isoform of AFAP1, AFAP-120, contains an additional 86 amino acids encoded by a novel exon near the carboxy terminus and has been shown to have a regulated expression pattern during the development of the mouse brain. AFAP-120 is thought to direct signaling through the protein tyrosine kinases cSrc and Fyn during neural development (Clump et al., 2003; Flynn et al., 1995).

Work by Han et al. identified that AFAP1 and cSrc were in complex after mechanical stretch-induced cytoskeletal deformation and that this was dependent upon the integrity of the N-terminal SH3 binding motif of AFAP1. Once bound, cSrc had the

ability to phosphorylate AFAP1 on multiple tyrosine residues to enhance the interaction (Guappone and Flynn, 1997; Guappone et al., 1998; Han et al., 2004; Qian et al., 1998). Further studies on the role of AFAP1 in mechanical stretch-induced cSrc activation by Xu et al. identified several human expressed sequence tag (EST) clones that shared similarities with AFAP1, including a transcript containing a partial C-terminal open reading frame (Xu et al., 2007). Analysis of this transcript by reverse transcription and polymerase chain reaction amplification resulted in the identification of a novel protein of 818 amino acids with a predicted mass of 91,300 Da. An antibody created against this protein revealed a relative mobility (M_r) of 130kDa, thus this protein was titled XB130 which will hereinafter be referred to as AFAP1L2 for Actin Filament Associated Protein 1-like 2 (Xu et al., 2007). AFAP1L2 is predicted to be a paralogue of AFAP1 by the Ensembl database and contains 19 exons found on chromosome 10q25.3.

A third paralogue of AFAP1, identified as Actin Filament Associated Protein 1-like 1 or AFAP1L1, is also predicted in the Ensembl database although very little is known aside from general sequence knowledge. AFAP1L1 is found at chromosomal position 5q33.1 and contains 19 exons which encode 768 amino acids with an approximate mass of 86,432 Da. A second isoform of AFAP1L1 encodes 725 amino acids and is lacking the 43 amino acids encoded by exon 18.

IV: The AFAP Family: A Genomic Study

While AFAP family members are found on separate chromosomes, there may be an evolutionary link between their locations. First proposed in 1970, the hypothesis that two rounds of whole genome duplication occurred in the evolution of the vertebrate

genome from invertebrate is known as the 2R hypothesis (Ohno, 1970). The 2R

hypothesis suggests that the presence of paralogous genes in a species that are located on

different chromosomes is due to regional duplication during evolution (Ohno, 1970). The

theory originated with an early observation that gene families which have one member in

invertebrates can have up to four members in vertebrates (Dehal and Boore, 2005).

Although this "4:1" rule does not hold up in many vertebrates, this can be attributed to

the many deletions, translocations and other genetic anomalies that can occur throughout

evolution (Dehal and Boore, 2005; Ohno, 1970). Various groups have still shown that in

some gene families, one family member exists in invertebrates while multiple members

exist in the vertebrate lineage and can be mapped to predictable sites within

chromosomes. Of particular interest to the AFAP family is the work of Pebusque et al.

Through the study of various genes on human chromosome 8 (PA family, ANK family

and FGFR family among others), it was found that paralogous regions exist between

human chromosomes 4p16, 5q33-35, 8p12-21 and 10q24-26 (Pebusque et al., 1998).

AFAP family members are found in these regions: AFAP1 at 4p16.1, AFAP1L1 at

5q33.1 and AFAP1L2 at 10q25.3. The 2R hypothesis would predict that only one AFAP

family member would be found in invertebrates while the evolution of the vertebrate

lineage would produce multiple family members. To study the AFAP family, the defining

characteristic which establishes these three proteins as a family, the amino acid sequence

of the AFAP1 pleckstrin homology 1 (PH1) domain, was used to determine paralogous

genes in the Ensenbl database.

Urochordates, also known as tunicates or sea squirts, are thought to be the

invertebrate most closely related to vertebrates (Delsuc et al., 2006; Kasahara, 2007).

Indeed, *Ciona intestinalis*, the vase tunicate, has a 493 amino acid PH domain containing protein predicted to be an orthologue of the AFAP family by the Ensembl database which shares 13% overall identity (40% similarity) with AFAP1, 13% overall identity (37% similarity) with AFAP1L1 and 14% overall identity (35% similarity) with AFAP1L2 (Figure 1A). Similarity is particularly strong between all three AFAP family member PH1 domains and the *Ciona intestinalis* PH domain containing protein with 29% identity (68% similarity) with the AFAP1 PH1 domain, 21% identity (59% similarity) with the AFAP1L1 PH1 domain and 27% identity (63% similarity) with the AFAP1L2 PH1 domain (Figure 1B). *Ciona intestinalis*, while predicted by Ensembl to have an orthologous relationship with all three AFAP family members in vertebrate lineages, is not predicted to have any paralogous genes in its genome. Flajnik et al. predict that the first round of whole genome duplication took place after the split of jawless vertebrates and protochordates. It is then predicted that a second round of whole genome duplication took place upon the split of cartilaginous fish from all other jawed vertebrates (Flajnik and Kasahara, 2001). This would be consistent with the finding of one orthologous PH domain containing protein in urochordates, the closest invertebrate ancestor to vertebrates, and the finding of all three paralogous AFAP family members in higher vertebrates. Taking into account similarity with the PH domain containing protein in *Ciona intestinalis*, one could predict, upon evolution of the vertebrate genome, that AFAP1 was the first AFAP family member. It can then be predicted that AFAP1L2 was the second AFAP family member to arise. AFAP1L1 shares more overall similarity with *Ciona intestinalis* than AFAP1L2, but the core factor that makes AFAP family members a family, the PH domains, is more conserved between AFAP1L2 and *Ciona intestinalis*.

A possible reason for AFAP1L1 to have higher overall similarity, though lower similarity in the PH1 domain, is the fact that AFAP1L1 is more closely related to AFAP1 than to AFAP1L2. It could be predicted that after the appearance of AFAP1L2 from AFAP1, AFAP1L1 then appeared after a second round of whole genome duplication, being duplicated from AFAP1 and not AFAP1L2. This is supported by both cladistic analysis of the AFAP family members, showing a closer relationship between AFAP1L1 and AFAP1 than AFAP1L1 and AFAP1L2 (Figure 2A) and also an Ensembl generated phylogram using all Ensembl AFAP family member sequence data (Figure 2B) (Larkin et al., 2007). *Ciona intestinalis*, which is predicted to be the closest invertebrate available in the ENSEMBL database that is related to the AFAP family, is indeed shown to have come from an ancestor earlier than any vertebrate AFAP family member. Red nodes in Figure 2B indicate predicted duplication events while blue nodes indicate predicted speciation events. A predicted duplication event occurred in a vertebrate ancestor that lead to two arms, one from which AFAP1L2 arose and another from which AFAP1 arose. A second predicted duplication event occurred in the arm of AFAP1 which effectively split that arm again into two, one from which the AFAP1 lineage arose and a second from which the AFAP1L1 lineage arose. This supports our hypothesis that AFAP1L1 and AFAP1 are more closely related and came from a more common ancestor than AFAP1L2. Interestingly, a third duplication event is predicted to have occurred in the AFAP1L1 lineage of Clupeocephala, or bony fish (Kasahara, 2007). In support of this hypothesis, AFAP1L1 has two predominant isoforms in the Ensembl database that are found in the bony fish *Tetraodon nigroviridis, Takifugu rubripes, Gasterosteus aculeatus, Oryzias latipes,* and *Danio rerio* while AFAP1 and AFAP1L2 only contain one isoform

in these species (Figure 2B). Though there are multiple AFAP family members in the vertebrate lineage, the closest predicted invertebrate ancestor has only one homologous PH domain containing protein and shares closest homology with AFAP1.

We hypothesize that AFAP1 arose in vertebrates from the invertebrate line and two rounds of whole genome duplication gave rise to the similar family members, AFAP1L1 and AFAP1L2. A third round of whole genome duplication may have occurred in the Clupeocephala lineage of AFAP1L1, leading to a possible second isoform of AFAP1L1 in bony fish. Across species, AFAP family members are well conserved. A detailed phylogram of AFAP1 (Figure 3) shows the sequential relationship between various vertebrates. Both AFAP1 and AFAP1L1 share 99% identity with their nearest ancestor sequence, chimpanzee, with only a one amino acid change while AFAP1L2 has five amino acid changes between chimpanzee and human sequence. In a simpler description of the amino acid sequences across human, chimpanzee, mouse, lizard and zebrafish, AFAP1 sequences show a high level of identity and similarity throughout the entire AFAP1 sequence (Figure 4). While AFAP1 shares the most identity across species, a similar analysis of AFAP1L1 (Figure 5) and AFAP1L2 (Figure 6) also shows a high level of consensus amino acid sequence across species in these family members. This consensus amino acid sequence in the AFAP family encodes for multiple protein binding domains which establish the AFAP proteins as a family of adaptor proteins.

V: The AFAP Family: Protein Structure

Protein binding domains and their complementary protein binding motifs play a key role in regulating cellular processes. While numerous protein binding domains have

been described, each binding domain can interact with multiple ligands thus creating an even larger array of signaling possibilities. A single protein binding domain may be prevalent throughout various proteins, as is characterized by the many phosphotyrosine-binding SH2 and poly-proline binding SH3 domains that have been identified (Pawson and Nash, 2003). These protein binding domains, in conjunction with numerous others, allow proteins to create signaling complexes that are not only specific with regard to associated proteins but also to the spaciotemporal patterns of cellular processes.

In its earliest studies, AFAP1 was found to be associated with actin filaments via an actin binding domain in the carboxy terminus and was composed of both amino-terminal and carboxy-terminal SH2 binding motifs and two juxtaposed SH3 binding motifs of which the N-terminal motif is required for cSrc binding (Guappone and Flynn, 1997; Guappone et al., 1998; Qian et al., 1998). Two PH domains are found within the interior of AFAP1. The PH1 domain has been shown to be involved in binding of the serine/threonine kinase PKCα and in the intra-molecular regulation of AFAP1 by binding to a downstream leucine zipper (Qian et al., 2004). Both PH domains are involved in the stabilization of AFAP1 multimers (Clump et al., 2010). The amino acid sequence between the PH domains (substrate domain, SD) is rich in serine and threonine residues. Phosphorylation of AFAP1 in this region by PKCα plays a role in podosome formation (Dorfleutner et al., 2008; Gatesman et al., 2004). The leucine zipper, an alpha-helical structure containing a heptad repeat of leucine residues found amino-terminal to the actin binding domain, is necessary for both an inhibitory intramolecular interaction with the AFAP1 PH1 domain and also to multimerize with other AFAP1 molecules (Kouzarides and Ziff, 1988). AFAP1 has the ability to cross-link actin filaments by binding via its

actin binding domain and then interacting with other AFAP1 molecules via the leucine zipper motif and possibly other internal sequences.

A similarity search of AFAP1 in the Ensembl database identified AFAP1L1 and AFAP1L2 which share 71% and 64% overall similarity with AFAP1 respectively (Xu et al., 2007) (Figure 7). Of interest between these proteins was both similarity in the modular domain structure (Figure 8A) and conservation of amino acid sequence in the PH domains (Figure 8B). While the approximately 250 known PH domain containing proteins have a highly conserved structure in these domains, a b-barrel with four b strands on one side and three b strands on the other connected by three variable loops connecting b1/b2, b3/b4 and b6/b7, they do not have conserved amino acid sequences (DiNitto and Lambright, 2006). The PH domains of AFAP1, AFAP1L1 and AFAP1L2 have both similar structure and similar sequence, with 44% identity (80% similarity) between the PH1 domains and 40% identity (74% similarity) between the PH2 domains (Figure 7). It is these similar sequences and overall domain structure that designate AFAP1, AFAP1L1 and AFAP1L2 as related proteins. Consistent with AFAP1, AFAP1L1 and AFAP1L2 also contain at least one SH3 binding motif in their N-terminus. An SH3 domain is a region of approximately sixty amino acids known to bind to proline-rich regions with a core PXXP motif (Ren et al., 1993). AFAP1 contains two juxtaposed SH3 binding motifs, PPQMPLPEIP and PPDSGPPPLP, and binds cSrc using the N-terminal PEIP sequence (Guappone and Flynn, 1997). AFAP1L2 is also known to be a cSrc binding partner by virtue of its SH3 binding motif, PDLPPPKMIP (Xu et al., 2007). While AFAP1L1 contains an SH3 binding motif, DLPPPLPNKP, this sequence is not consistent with a consensus cSrc binding motif and instead resembles that of a consensus

binding motif for cortactin, which preferentially binds to a +PPYPXKPXWL motif where + is a basic residue, Y is an aliphatic residue, and X is any amino acid (Sparks et al., 1996). While consensus amino acid sequence designates these as SH3 binding motifs, it is the specificity of the surrounding amino acids that can direct AFAP family members to interact with different SH3 domain containing proteins.

Also consistent with AFAP1, AFAP1L1 and AFAP1L2 have potential sites for phosphorylation as predicted by Scansite (Figure 9), although not all of these have been confirmed as true phosphorylation targets or SH2 motifs (Obenauer et al., 2003). The SH2 domain, a protein binding domain of approximately 100 amino acids, binds preferentially to an SH2 binding motif, a short amino acid sequence containing a site of tyrosine phosphorylation which is directly amino terminal to 1 or 2 negatively charged amino acids followed by a hydrophobic amino acid (Songyang et al., 1993). It is these surrounding amino acids that allow SH2 domains to display selectivity in their binding partners (Cantley et al., 1991). The known N-terminal AFAP1 SH2 binding motif, Y^{93}YEEA, shares 100% identity with both AFAP1L1 and AFAP1L2 while the C-terminal SH2 binding motif, Y^{451}DYI, shares 75% identity as both AFAP1L1 and AFAP1L2 share a YDYV sequence. A third SH2 binding motif in AFAP1L2, Y^{54}IYM, is responsible for PI3K interaction and is not conserved in AFAP1 or AFAP1L1.

The leucine zipper of AFAP1 is a helical repeat of leucine residues that is responsible for intramolecular regulation of AFAP1 by contacting the PH1 domain and keeping AFAP1 is a closed conformation and also for contacting other AFAP1 molecules when AFAP1 is activated by PKCα so as to aid in multimerization and bundling of actin filaments (Qian et al., 2004). AFAP1L1 contains a similar but less conserved leucine

zipper motif which may still allow AFAP1L1 to multimerize with other AFAP1L1 molecules or perhaps other AFAP family members. AFAP1L2 contains a coiled-coil region corresponding to the AFAP1 leucine zipper and actin binding domain and may instead use an N-terminal actin binding domain to localize to branched filamentous (F)-actin networks in the lamellipodium (Lodyga et al., 2010; Xu et al., 2007).

VI: The AFAP Family: Cellular Function

AFAP1 was first identified as a tyrosine-phosphorylated protein that had the ability to interact with activated forms of the non-receptor protein tyrosine kinase cSrc (Flynn et al., 1993). This interaction was dependent upon the integrity of both the cSrc SH2 and SH3 domains as well as their binding motifs within AFAP1 (Kanner et al., 1991; Kmiecik and Shalloway, 1987; Reynolds et al., 1989). When cells become transformed by cSrc, AFAP1 undergoes a change in localization from actin filaments into podosomes (Flynn et al., 1993; Linder and Kopp, 2005). In its inactive form, cSrc is found in the perinuclear region while activated cSrc moves to the cell periphery where association with the plasma membrane is necessary for cell transformation (Yeatman, 2004). This movement is dependent upon the Rho family members where active cSrc can be moved to focal adhesions by RhoA, to lamellipodia by Rac1 and to filopodia by Cdc42 (Sandilands and Frame, 2008). With regard to AFAP1, consitutively active RhoA was able to largely overcome actin filament rearrangement seen upon activation of AFAP1. Therefore, the ability of AFAP to bind to and activate cSrc may be in a Rho-dependent manner (Baisden et al., 2001a). To determine the temporal binding of AFAP1 and cSrc, a temperature-sensitive form of constitutively active viral Src, LA29, which

resides on perinuclear vesicles at non-permissive temperatures (39.5°C) and induces cellular transformation at permissive temperature (35°C), was utilized. Upon cell stimulation with PKCα at non-permissive temperature, AFAP1 was seen to move to perinuclear vesicles and colocalize with cSrc. AFAP1 and cSrc, which is activated by the binding of AFAP1, can then move to sites of actin rearrangement in transformed cells when shifted to the permissive temperature (Walker et al., 2007). Interestingly, a polymorphic variant of AFAP1 exists that is an efficient cSrc activator. A nonsynonymous single nucleotide polymorphism (SNP) in the PH2 domain resulting in a Ser[403] to Cys[403] amino acid change found in approximately 21% of both tumor and adjacent normal tissues and 28% of normal tissues is more efficient at activating cSrc when cSrc levels are high. The PH2 domain of AFAP1 does not contain the necessary conserved basic residues for lipid binding and is instead predicted by molecular modeling to be in contact with a water molecule 71% of the time. This Ser[403] to Cys[403] change was predicted to interrupt the PH2 domain's ability to interact with a water molecular possibly leading to a destabilized structure of the PH2 domain which may alter its binding capacity and thus create a more constitutively active AFAP molecule (Clump et al., 2010).

Stimulation of cells with PKCα induces a cell morphology like that seen with activated cSrc. PKCα binds to the PH1 domain of AFAP1 and phosphorylates a key serine residue in the substrate domain, Ser[277] (Dorfleutner et al., 2008; Gatesman et al., 2004). AFAP1 is normally held in an inhibitory intramolecular interaction by binding of the leucine zipper to the PH1 domain (Qian et al., 2004). The binding of PKCα to the PH1 domain disrupts this interaction and AFAP is able to interact with cSrc. Mutants of

AFAP1 that contain deletions of the PH1 domain are unable to bind PKCa and cannot associate with and activate cSrc (Gatesman et al., 2004). When cSrc is not present, PKCa is unable to direct this change thus indicating that cSrc is necessary for actin filament rearrangement.(Gatesman et al., 2004).

The leucine zipper of AFAP1 is located in the C-terminal portion of the AFAP1 molecule juxtaposed to an actin binding domain and holds AFAP1 into an autoinhibitory conformation by contacting the PH1 domain. Deletion or modification of the leucine zipper destabilizes this intramolecular interaction, resulting in a conformational change in AFAP1 that allows cSrc activation and more efficient bundling of actin filaments (Qian et al., 1998; Qian et al., 2004). Similarly, when this inhibition is released by binding of PKCa, the leucine zipper may then promote increased actin filament crosslinking activity through interactions with other AFAP1 molecules (Qian et al., 2004). This is witnessed by the interaction of AFAP1 in cell lysates with GST proteins expressing only the AFAP1 leucine zipper. Upon expression of AFAP1 constructs lacking the leucine zipper, AFAP1 is shown to have an intrinsic ability to activate cSrc and localize to podosomes because the intramolecular inhibition of AFAP1 is not present (Qian et al., 1998). Therefore, the leucine zipper plays a key role in the ability of AFAP1 to multimerize and subsequently cross-link actin filaments. As the neighboring actin binding domain, which has been shown to be necessary and sufficient for the binding of AFAP1 to actin, of AFAP1 contacts actin filaments, the leucine zipper allows multimerization of AFAP1 molecules and thus bundles the actin filaments together (Qian et al., 2000). Gel filtration experiments have shown that AFAP1 has the ability to exist in large multimeric complexes that are mediated by the leucine zipper. Upon deletion of the leucine zipper,

AFAP1 complexes are smaller and predicted to exist as dimers or trimers. These smaller complexes allow AFAP1 to more efficiently cross-link actin from a meshwork into a bundle. Co-pelleting assays of recombinant AFAP1 showed a direct association of AFAP1 with actin filaments, and this association was not inhibited by deletion of the leucine zipper (Qian et al., 2002).

Knockdown of AFAP1 by short hairpin RNA (shRNA) in the MDA-MB-231 breast cancer cell line, while having no gross effect on cell morphology, resulted in a loss of stress filament integrity and reduced cell adhesion. While cortical actin structure was maintained in AFAP1 shRNA knockdown cells, stress fibers could not be detected by phalloidin immunostaining, possibly due to the loss of cross-linking of actin filaments by the lack of AFAP1. The role of AFAP1 in the integrity of stress filaments was confirmed by an inducible shRNA system in which AFAP1 levels were rescued or knocked down due to the absence or presence of doxycycline. When AFAP1 levels are knocked down, phalloidin readily decorates cortical actin while stress fibers cannot be detected. Upon rescue of AFAP1 levels by relief of doxycycline, stress fibers reappear. The ability of AFAP1 molecules to bundle actin by virtue of their actin binding domains and leucine zippers may be responsible for the lack of stress fiber formation in AFAP1 knockdown cells. Stimulation of wild-type MDA-MB-231 cells with lysophosphatidic acid (LPA), a Rho activator, induces actin filament cross-linking into stress fibers and more pronounced focal adhesions. When MDA-MB-231 AFAP1 knockdown cells are treated with LPA, stress fiber formation is inhibited and focal contacts do not form (Dorfleutner et al., 2007). AFAP1 knockdown cells are slower to adhere to fibronectin, however, given enough time, they will reach adhesion levels similar to cells expressing AFAP1 and with

no detectable difference in cell survival. It is possible that this difference in cell adhesion may be due to the role of AFAP1 in cellular adhesion structures. While wild-type cells readily form adhesion structures when plated onto fibronectin, AFAP1 knockdown cells have an impaired ability to form focal contacts and have no stress fiber formation in the same time span that wild-type cells have matured their focal contacts. The formation of stress fibers and their role in cell contractility is essential for the maturation of focal adhesions and thus the loss of AFAP1 may affect the adherence ability of the cell (Parsons et al., 2010). As cellular adhesion is a prerequisite for cell migration, knockdown of AFAP1 in prostate cancer cell line PC3 impaired cellular motility as witnessed by their lack of migration in a modified Boyden chamber (Zhang et al., 2007). While the knockdown of AFAP1 in MDA-MB-231 cells did not affect the expression levels of α5 integrin, β1 integrin, paxillin or vinculin, PC3 cells showed decreased levels of β1 integrin when AFAP1 expression was lost (Dorfleutner et al., 2007; Zhang et al., 2007). β1 integrin expression can be restored in PC3 AFAP1 knockdown cells by reintroduction of wild-type AFAP1. While the AFAP1 mutant defective for binding to cSrc, AFAP1[71A] , also has the ability to restore β1 integrin expression, AFAP1 mutants lacking the PH1 domain failed to restore β1 integrin expression and stress filament and focal contact expression in PC3 AFAP1 knockdown cells. This indicates that it is not the altered AFAP1/cSrc interaction that is affected when AFAP1 is knocked down in PC3 cells, but that the role of AFAP1 in focal contact formation through stress fiber formation and cell contractility may be regulated by the ability of AFAP1 to bind to PKC (Zhang et al., 2007). While AFAP1 has been shown to bind to multiple PKC family members, $\alpha,\beta,\gamma\iota$, it is the activation and binding of PKCα which phosphorylates AFAP1 on serine

residues in the substrate domain (Dorfleutner et al., 2008; Qian et al., 2002). Mutation of key serine residues revealed that Ser^{277} is phosphorylated by PKCa and this is either independent or upstream of cSrc activation. A phospho-specific antibody created against the AFAP1 $pSer^{277}$ site showed that in resting A7r5 cells, while AFAP1 localizes to stress filaments, there is no evidence for phosphorylation of Ser^{277}. Stimulation of A7r5 cells with the phorbol ester phorbol 12,13-dibutyrate (PDBu), a PKCa activator, induced phospho-GFP-AFAP1 to localize strongly to podosomes. When GFP-AFAP1 is overexpressed in A7r5 cells, the majority of cells display stress fibers while some display both stress fibers and small podosomes and a very small percentage spontaneously produced podosomes. If A7r5 cells express GFP-AFAP1^{S277A}, a mutant of AFAP1 in which Ser^{277} has been mutated to an alanine, there is in increase in the number of cells that spontaneously form podosomes and it is reasoned that phosphorylation of AFAP1 at Ser^{277} plays a role in podosome formation (Dorfleutner et al., 2008). When podosome lifespan was calculated, A7r5 cells expressing AFAP1^{S277A} mutants showed decreased podosome turnover and longer average lifespan of podosomes as compared to wildtype AFAP1 (Dorfleutner et al., 2008). Due to its ability to act as an actin bundling protein by virtue of the leucine zipper and actin binding domain, AFAP1 may regulate the construction/deconstruction of podosomes in response to cellular adhesion signals such as PKC activation.

AFAP1L2 was also initially discovered as a cSrc binding partner and has been found to be able to interact with the SH2 and SH3 domain of both cSrc and GTPase activating protein (GAP) (Han et al., 2004). Similarly to AFAP1, AFAP1L2 has the ability to interact with cSrc when overexpressed, activates cSrc as witnessed by an

overall increase in tyrosine phosphorylation seen when both AFAP1L2 and cSrc are overexpressed, and has been shown to be a phosphorylation target of cSrc. Deletions in AFAP1L2 that remove the N-terminal SH3 and SH2 binding motifs (XB130DN) reduce the interaction between AFAP1L2 and cSrc and result in a decrease in cSrc phosphorylation at Tyr[416] and overall cellular phosphotyrosine when mutant AFAP1L2 and cSrc are overexpressed. Similarly, siRNA knockdown of AFAP1L2 decreases the tyrosine phosphorylation of cSrc at Tyr[416] (Xu et al., 2007).

Using a luciferase assay, it was shown that co-expression of AFAP1L2 and cSrc had the ability to increase the activation of activating protein-1 (AP-1) and serum response element (SRE) reporters which may indicate that the AFAP1L2 and cSrc interaction is important in gene regulation. Expression of XB130DN with cSrc greatly decreased SRE activity in comparison with wild-type AFAP1L2 but there was still some activation of SRE in comparison with cSrc expression alone. This indicates that there are domains in AFAP1L2 other than the N-terminal SH3 and SH2 domains that can be involved in cSrc activation and SRE activity (Xu et al., 2007). Stimulation of cells by epidermal growth factor (EGF) showed enhanced SRE activation in cells expressing wild-type AFAP1L2 and cSrc and this activation was blunted when cSrc was expressed with XB130DN, suggesting a role for AFAP1L2 in signal transduction from EGF leading to cellular processes such as mitosis, transformation or survival. Knockdown of AFAP1L2 by siRNA was able to blunt the effects of EGF stimulation and arrest cells in the G1 phase, thus supporting the conclusion that AFAP1L2 plays a role in EGF signaling. In comparison, the interleukin-8 (IL-8) promoter which contains AP-1 sites was linked to a luciferase reporter and, similarly, showed increased activation when

AFAP1L2 and cSrc were overexpressed. This was also blunted by the expression of XB130ΔN (Xu et al., 2007).

Unlike AFAP1, AFAP1L2 has not been shown to interact with stress fibers. Immunofluoresence of GFP-tagged, His-tagged and endogenous AFAP1L2 in multiple cell lines indicated that AFAP1L2 has a cytosolic distribution in resting cells but can translocate to the lamellipodium under cell stimulation in a Rac-dependent manner. AFAP1L2 contains a coiled-coil in its C-terminus where AFAP1 contains an actin binding domain, and deletion of this coiled-coil domain had no effect on the translocation of AFAP1L2 to branched F-actin networks. Instead of an actin binding domain similar to AFAP1, AFAP1L2 may contain an actin binding domain 2 (ABD2) similar to that of the actin binding domain of talin in its N-terminus. Deletion of the N-terminus inhibited AFAP1L2 translocation to the lamellipodia while expression of a tagged peptide of the first 167 amino acids of the N-terminus was sufficient for translocation (Lodyga et al., 2010).

Tissue analysis of AFAP1L2 showed localization mainly to the thyroid and spleen and it was noted that AFAP1L2 expression was also found in human papillary thyroid carcinoma (PTC) at or below normal thyroid levels (Lodyga et al., 2009). Because a central feature of PTC is dysregulation of the Rearranged in Transformation/papillary thyroid carcinoma (RET/PTC) tyrosine kinase and AFAP1L2 is a potential substrate for kinase activity, AFAP1L2 and RET/PTC were transiently overexpressed in human embryonic kidney 293 (HEK) cells and an increase in tyrosine phosphorylation of AFAP1L2 dependent on the kinase activity of RET/PTC was seen. Although AFAP1L2 has been shown to interact with and be a substrate for cSrc, the presence or absence of

cSrc had no effect on RET/PTC phosphorylation or interaction with AFAP1L2 (Lodyga et al., 2009). Point mutations within the predicted SH2 binding motifs of AFAP1L2 revealed that Tyr^{54} is a major, although not the sole, target for RET/PTC phosphorylation. This SH2 binding motif, YIYM, resembles that of a PI3K binding motif, YXXM in which X is any amino acid. Indeed, expression of AFAP1L2 with RET/PTC greatly increased the association of AFAP1L2 with the p85 regulatory subunit of PI3K in immunoprecipitation experiments and this interaction was dependent upon the phosphorylation status of Y^{54} (Lodyga et al., 2009). Since the serine/threonine kinase Akt is a well known downstream target of PI3K signaling, Akt activation was tested in response to overexpression of AFAP1L2 and RET/PTC to see if AFAP1L2 played a role in the Akt pathway. Akt phosphorylation was reduced by either mutation of AFAP1L2 so that it may not be phosphorylated at Y^{54} or inhibiting the PI3K pathway. siRNA knockdown of AFAP1L2 also lead to a decrease in Akt phosphorylation while it had no effect on other PI3K targets such as extracellular signal-regulated kinase (ERK). As Akt activation is known to lead to cell survival, AFAP1L2 was knocked down in TPC1 thyroid papillary carcinoma cells to look at the effect of AFAP1L2 on tumor cell survival (Marte and Downward, 1997). Knockdown of AFAP1L2 in TPC1 cells showed a higher amount of cells in the G_0/G_1 phase and decreased proliferation of TPC1 tumor cells. As the proliferation of TPC1 has been attributed to the RET/PTC recombination, AFAP1L2 may be playing an important role in the progression of tumor cells by linking together the RET/PTC and Akt pathways. Downregulation of AFAP1L2 has been shown in TPC1 cells to decrease cell migration and invasion. Therefore, it is hypothesized that AFAP1L2 may translocate to the cell periphery upon stimulation where it may interact with PI3K

and thus link the RET/PTC kinase to the Akt pathway and subsequent migration and invasion (Lodyga et al., 2010).

Although the function of AFAP1L1 can be hypothesized based upon its sequence and similarity with AFAP1 and AFAP1L2, no cellular data was available to determine the function of AFAP1L1.

VII: The AFAP Family: Pathology and Disease

As cSrc is highly active in diseases such as cancer and AFAP1 is a known cSrc activator, it is probable that AFAP1 and possibly other AFAP family members have a role in disease progression. Dorfleutner et al. have shown that expression of AFAP1 is necessary for stress fiber formation and the efficient formation of focal adhesions (Dorfleutner et al., 2007). The phosphorylation of AFAP1 at Ser^{277} plays a role in the stability and lifespan of podosomes (Dorfleutner et al., 2008). In cases where AFAP1 is overexpressed, it has been shown that increased levels of AFAP1 correlate with progressive stages of disease. Through immunohistochemical staining, Zhang et al. showed that normal prostate tissues have high AFAP1 levels in prostate smooth muscle (Zhang et al., 2007). AFAP1 was low in the luminal secretory layer and the basal layer of prostatic glands. In benign prostatic hyperplasia, the luminal secretory layer of cells again had low AFAP1 immunoreactivity while basal layer cells had weak reactivity. In contrast, prostate cancer tissues exhibited strong AFAP1 immunoreactivity in all prostate tissues. Levels of AFAP1 immunostaining correlated with increasing Gleason scores for grading the aggressiveness of tumors, and this correlation was seen in normal versus tumorigenic cells lines as well. The downregulation of AFAP1 through RNA interference

(RNAi) slowed the proliferation of the prostate cancer cell line PC3 in soft agar, possibly through a decrease in anchorage-independent growth. These same cells implanted orthotopically into nude mice produced either no or significantly smaller tumors as compared to parental lines, indicating that AFAP1 played an important role in the progression of prostate tumors. The migration of PC3 cells lacking AFAP1 expression was markedly decreased when compared to parental lines, and this was thought to be due to a loss of focal contacts in AFAP1 knockdown cells as these cells also resulted in a loss of b1 integrin expression. Although the reason for high expression levels of AFAP1 in prostate cancer is still unknown, it is increased integrin signaling to the extracellular matrix through AFAP1 that is thought to mediate prostate cancer growth and invasive potential (Zhang et al., 2007).

In addition to prostate cancer, Clump et al. have also shown that AFAP1 levels, although not detectable in normal ovarian tissue, are increased in ovarian cancer and also correlate with progression of disease (Clump et al., 2010). A polymorphic variant of AFAP1 in which a key serine residue in the PH2 domain is mutated to a cysteine, AFAP1^{403C}, is found in approximately 20-30% of both normal and ovarian cancer tissues. This variant has an increased capacity for cSrc activation and again, although the reason for increased AFAP1 expression in ovarian cancer is also unknown, overexpression of this polymorphic variant can result in increased cSrc activation and a more aggressive cancer (Clump et al., 2010). Through these experiments, it is plausible to assume that AFAP1 could be used in both prostate and ovarian cancers as a marker for tumor progression and aggressiveness.

AFAP1L2 is also known to play role in cancer. AFAP1L2 expression is high in normal thyroid tissue and has been shown to be expressed in high levels in human papillary thyroid carcinoma (Lodyga et al., 2009). A genetic hallmark of PTC are various rearrangements of the RET tyrosine kinase with portions of various other proteins, most often coiled-coil domain containing protine 6 (CCDC6), which is a coiled-coil domain containing protein that may act in tumor suppression through apoptosis (Leone et al., 2010). AFAP1L2 has multiple sites for tyrosine phosphorylation and Y^{54} was shown to be the major target of RET/PTC kinase. The amino acid sequence surrounding Y^{54} predicts a motif that is homologous for a binding site of the p85 subunit of PI3-kinase and these proteins are able to interact in an endogenous setting. Akt is a serine/threonine kinase downstream of PI3k activation and the interaction of AFAP1L2 and PI3k greatly increased the activation of Akt. Aberrant activation of this pathway can lead to cell proliferation and survival in cancer cells such as in thyroid papillary carcinoma where high levels of AFAP1L2 are found. Therefore, both AFAP1 and AFAP1L2 have been shown to play a role in cancer and possible cell survival and metastasis.

As for AFAP1L1, online database screens can assess the level of AFAP1L1 expressed in normal and cancerous tissues. The National Cancer Institute Serial Analysis of Gene Expression (SAGE) database predicts AFAP1L1 levels in normal tissues to be high in heart, ovary, uterus, skin and vasculature. In comparison, AFAP1L1 levels are predicted to be high in various cancers such as brain, breast, liver, prostate, skin, stomach, and thyroid cancers. The Oncomine database predicts the highest AFAP1L1 expression levels as compared to normal to be found in esophageal cancer, glioblastoma and neuroblastoma. With regard to various cancers, regardless of the origin, it appears

that AFAP1L1 expression is upregulated, not downregulated, compared to levels in normal tissues. Therefore, AFAP1L1 may play a role in the formation or progression of different cancers by virtue of its signaling capabilities.

IX: Summary

The AFAP family represents a family of adaptor proteins with a variety of roles in signal transduction. AFAP1 is well-known as a cSrc activator and actin cross linking protein and its role in stress fiber formation is indispensable. AFAP1 also plays a key role in podosome formation and turnover. As AFAP1 levels are increased in various cancers, the inherent abilities of AFAP1 may play a role in the progression of disease. AFAP1L2, like AFAP1, is also a known cSrc binding protein and has been shown to link the RET/PTC kinase rearrangement to the Akt pathway by virtue of PI3K. AFAP1L2 localizes to the lamellipodium of motile cells in a Rac-dependent manner and its expression is necessary for efficient motility and invasion. As AFAP1L2 has been shown to be highly expressed in papillary thryroid carcinoma, it too provides an interesting target in the progression of disease. Although the role of AFAP1L1 can be inferred by its similarity with AFAP1 and AFAP1L2, no data is available on the role of AFAP1L1 in the cell. The purpose of this study was to elucidate the function of AFAP1L1 by virtue of its cellular and tissue localization and its interaction with cellular binding partners.

REFERENCES

Baisden, J.M., Gatesman, A.S., Cherezova, L., Jiang, B.H., and Flynn, D.C. (2001a). The intrinsic ability of AFAP-110 to alter actin filament integrity is linked with its ability to also activate cellular tyrosine kinases. Oncogene *20*, 6607-6616.

Baisden, J.M., Qian, Y., Zot, H.M., and Flynn, D.C. (2001b). The actin filament-associated protein AFAP-110 is an adaptor protein that modulates changes in actin filament integrity. Oncogene *20*, 6435-6447.

Brandt, D., Gimona, M., Hillmann, M., Haller, H., and Mischak, H. (2002). Protein kinase C induces actin reorganization via a Src- and Rho-dependent pathway. J Biol Chem *277*, 20903-20910.

Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R., and Soltoff, S. (1991). Oncogenes and signal transduction. Cell *64*, 281-302.

Clump, D.A., Clem, R., Qian, Y., Guappone-Koay, A., Berrebi, A.S., and Flynn, D.C. (2003). Protein expression levels of the Src activating protein AFAP are developmentally regulated in brain. J Neurobiol *54*, 473-485.

Clump, D.A., Yu, J.J., Cho, Y., Gao, R., Jett, J., Zot, H., Cunnick, J.M., Snyder, B., Clump, A.C., Dodrill, M.*, et al.* (2010). A Polymorphic Variant of AFAP-110 Enhances cSrc Activity. Transl Oncol *3*, 276-285.

Dehal, P., and Boore, J.L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol *3*, e314.

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature *439*, 965-968.

DiNitto, J.P., and Lambright, D.G. (2006). Membrane and juxtamembrane targeting by PH and PTB domains. Biochim Biophys Acta *1761*, 850-867.

Dorfleutner, A., Cho, Y., Vincent, D., Cunnick, J., Lin, H., Weed, S.A., Stehlik, C., and Flynn, D.C. (2008). Phosphorylation of AFAP-110 affects podosome lifespan in A7r5 cells. J Cell Sci *121*, 2394-2405.

Dorfleutner, A., Stehlik, C., Zhang, J., Gallick, G.E., and Flynn, D.C. (2007). AFAP-110 is required for actin stress fiber formation and cell adhesion in MDA-MB-231 breast cancer cells. J Cell Physiol *213*, 740-749.

Duncan, J.S., Turowec, J.P., Vilk, G., Li, S.S., Gloor, G.B., and Litchfield, D.W. (2010). Regulation of cell proliferation and survival: convergence of protein kinases and caspases. Biochim Biophys Acta *1804*, 505-510.

Fincham, V.J., Chudleigh, A., and Frame, M.C. (1999). Regulation of p190 Rho-GAP by v-Src is linked to cytoskeletal disruption during transformation. J Cell Sci *112 (Pt 6)*, 947-956.

Flajnik, M.F., and Kasahara, M. (2001). Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. Immunity *15*, 351-362.

Flynn, D.C., Koay, T.C., Humphries, C.G., and Guappone, A.C. (1995). AFAP-120. A variant form of the Src SH2/SH3-binding partner AFAP-110 is detected in brain and contains a novel internal sequence which binds to a 67-kDa protein. J Biol Chem *270*, 3894-3899.

Flynn, D.C., Leu, T.H., Reynolds, A.B., and Parsons, J.T. (1993). Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. Mol Cell Biol *13*, 7892-7900.

Frame, M.C., Fincham, V.J., Carragher, N.O., and Wyke, J.A. (2002). v-Src's hold over actin and cell adhesions. Nat Rev Mol Cell Biol *3*, 233-245.

Gatesman, A., Walker, V.G., Baisden, J.M., Weed, S.A., and Flynn, D.C. (2004). Protein kinase Calpha activates c-Src and induces podosome formation via AFAP-110. Mol Cell Biol *24*, 7578-7597.

Gimona, M., Buccione, R., Courtneidge, S.A., and Linder, S. (2008). Assembly and biological role of podosomes and invadopodia. Curr Opin Cell Biol *20*, 235-241.

Guappone, A.C., and Flynn, D.C. (1997). The integrity of the SH3 binding motif of AFAP-110 is required to facilitate tyrosine phosphorylation by, and stable complex formation with, Src. Mol Cell Biochem *175*, 243-252.

Guappone, A.C., Weimer, T., and Flynn, D.C. (1998). Formation of a stable src-AFAP-110 complex through either an amino-terminal or a carboxy-terminal SH2-binding motif. Mol Carcinog *22*, 110-119.

Hai, C.M., Hahne, P., Harrington, E.O., and Gimona, M. (2002). Conventional protein kinase C mediates phorbol-dibutyrate-induced cytoskeletal remodeling in a7r5 smooth muscle cells. Exp Cell Res *280*, 64-74.

Han, B., Bai, X.H., Lodyga, M., Xu, J., Yang, B.B., Keshavjee, S., Post, M., and Liu, M. (2004). Conversion of mechanical force into biochemical signaling. J Biol Chem *279*, 54793-54801.

Ho, B., Hou, G., Pickering, J.G., Hannigan, G., Langille, B.L., and Bendeck, M.P. (2008). Integrin-linked kinase in the vascular smooth muscle cell response to injury. Am J Pathol *173*, 278-288.

Kanner, S.B., Reynolds, A.B., Vines, R.R., and Parsons, J.T. (1990). Monoclonal antibodies to individual tyrosine-phosphorylated protein substrates of oncogene-encoded tyrosine kinases. Proc Natl Acad Sci U S A *87*, 3328-3332.

Kanner, S.B., Reynolds, A.B., Wang, H.C., Vines, R.R., and Parsons, J.T. (1991). The SH2 and SH3 domains of pp60src direct stable association with tyrosine phosphorylated proteins p130 and p110. EMBO J *10*, 1689-1698.

Kasahara, M. (2007). The 2R hypothesis: an update. Curr Opin Immunol *19*, 547-552.

Kmiecik, T.E., and Shalloway, D. (1987). Activation and suppression of pp60c-src transforming ability by mutation of its primary sites of tyrosine phosphorylation. Cell *49*, 65-73.

Kouzarides, T., and Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. Nature *336*, 646-651.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Larsson, C. (2006). Protein kinase C and the regulation of the actin cytoskeleton. Cell Signal *18*, 276-284.

Le Clainche, C., and Carlier, M.F. (2008). Regulation of actin assembly associated with protrusion and adhesion in cell migration. Physiol Rev *88*, 489-513.

Leone, V., Mansueto, G., Pierantoni, G.M., Tornincasa, M., Merolla, F., Cerrato, A., Santoro, M., Grieco, M., Scaloni, A., Celetti, A., *et al.* (2010). CCDC6 represses CREB1 activity by recruiting histone deacetylase 1 and protein phosphatase 1. Oncogene *29*, 4341-4351.

Linder, S. (2009). Invadosomes at a glance. J Cell Sci *122*, 3009-3013.

Linder, S., and Aepfelbacher, M. (2003). Podosomes: adhesion hot-spots of invasive cells. Trends Cell Biol *13*, 376-385.

Linder, S., and Kopp, P. (2005). Podosomes at a glance. J Cell Sci *118*, 2079-2082.

Lodyga, M., Bai, X.H., Kapus, A., and Liu, M. (2010). Adaptor protein XB130 is a Rac-controlled component of lamellipodia that regulates cell motility and invasion. J Cell Sci *123*, 4156-4169.

Lodyga, M., De Falco, V., Bai, X.H., Kapus, A., Melillo, R.M., Santoro, M., and Liu, M. (2009). XB130, a tissue-specific adaptor protein that couples the RET/PTC oncogenic kinase to PI 3-kinase pathway. Oncogene *28*, 937-949.

Marte, B.M., and Downward, J. (1997). PKB/Akt: connecting phosphoinositide 3-kinase to cell survival and beyond. Trends Biochem Sci *22*, 355-358.

Martiny-Baron, G., and Fabbro, D. (2007). Classical PKC isoforms in cancer. Pharmacol Res *55*, 477-486.

Mattila, P.K., and Lappalainen, P. (2008). Filopodia: molecular architecture and cellular functions. Nat Rev Mol Cell Biol *9*, 446-454.

Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic Acids Res *31*, 3635-3641.

Ohno, S. (1970). Evolution by Gene Duplication. London, Allen and Unwin New York, Spring-Verlag.

Parsons, J.T., Horwitz, A.R., and Schwartz, M.A. (2010). Cell adhesion: integrating cytoskeletal dynamics and cellular tension. Nat Rev Mol Cell Biol *11*, 633-643.

Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. Science *300*, 445-452.

Pebusque, M.J., Coulier, F., Birnbaum, D., and Pontarotti, P. (1998). Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. Mol Biol Evol *15*, 1145-1159.

Pellegrin, S., and Mellor, H. (2007). Actin stress fibres. J Cell Sci *120*, 3491-3499.

Qian, Y., Baisden, J.M., Cherezova, L., Summy, J.M., Guappone-Koay, A., Shi, X., Mast, T., Pustula, J., Zot, H.G., Mazloum, N., *et al.* (2002). PC phosphorylation increases the ability of AFAP-110 to cross-link actin filaments. Mol Biol Cell *13*, 2311-2322.

Qian, Y., Baisden, J.M., Westin, E.H., Guappone, A.C., Koay, T.C., and Flynn, D.C. (1998). Src can regulate carboxy terminal interactions with AFAP-110, which influence self-association, cell localization and actin filament integrity. Oncogene *16*, 2185-2195.

Qian, Y., Baisden, J.M., Zot, H.G., Van Winkle, W.B., and Flynn, D.C. (2000). The carboxy terminus of AFAP-110 modulates direct interactions with actin filaments and regulates its ability to alter actin filament integrity and induce lamellipodia formation. Exp Cell Res *255*, 102-113.

Qian, Y., Gatesman, A.S., Baisden, J.M., Zot, H.G., Cherezova, L., Qazi, I., Mazloum, N., Lee, M.Y., Guappone-Koay, A., and Flynn, D.C. (2004). Analysis of the role of the leucine zipper motif in regulating the ability of AFAP-110 to alter actin filament integrity. J Cell Biochem *91*, 602-620.

Rathinam, R., and Alahari, S.K. (2010). Important role of integrins in the cancer biology. Cancer Metastasis Rev *29*, 223-237.

Ren, R., Mayer, B.J., Cicchetti, P., and Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. Science *259*, 1157-1161.

Reynolds, A.B., Kanner, S.B., Wang, H.C., and Parsons, J.T. (1989). Stable association of activated pp60src with two tyrosine-phosphorylated cellular proteins. Mol Cell Biol *9*, 3951-3958.

Ridley, A.J., Schwartz, M.A., Burridge, K., Firtel, R.A., Ginsberg, M.H., Borisy, G., Parsons, J.T., and Horwitz, A.R. (2003). Cell migration: integrating signals from front to back. Science *302*, 1704-1709.

Roberts, E.C., Shapiro, P.S., Nahreini, T.S., Pages, G., Pouyssegur, J., and Ahn, N.G. (2002). Distinct cell cycle timing requirements for extracellular signal-regulated kinase and

phosphoinositide 3-kinase signaling pathways in somatic cell mitosis. Mol Cell Biol *22*, 7226-7241.

Saltel, F., Daubon, T., Juin, A., Ganuza, I.E., Veillat, V., and Genot, E. (2010). Invadosomes: Intriguing structures with promise. Eur J Cell Biol.

Sandilands, E., and Frame, M.C. (2008). Endosomal trafficking of Src tyrosine kinase. Trends Cell Biol *18*, 322-329.

Small, J.V., and Resch, G.P. (2005). The comings and goings of actin: coupling protrusion and retraction in cell motility. Curr Opin Cell Biol *17*, 517-523.

Small, J.V., Rottner, K., Kaverina, I., and Anderson, K.I. (1998). Assembling an actin cytoskeleton for cell attachment and movement. Biochim Biophys Acta *1404*, 271-281.

Small, J.V., Stradal, T., Vignal, E., and Rottner, K. (2002). The lamellipodium: where motility begins. Trends Cell Biol *12*, 112-120.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., *et al.* (1993). SH2 domains recognize specific phosphopeptide sequences. Cell *72*, 767-778.

Sparks, A.B., Rider, J.E., Hoffman, N.G., Fowlkes, D.M., Quillam, L.A., and Kay, B.K. (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. Proc Natl Acad Sci U S A *93*, 1540-1544.

Walker, V.G., Ammer, A., Cao, Z., Clump, A.C., Jiang, B.H., Kelley, L.C., Weed, S.A., Zot, H., and Flynn, D.C. (2007). PI3K activation is required for PMA-directed activation of cSrc by AFAP-110. Am J Physiol Cell Physiol *293*, C119-132.

Weed, S.A., and Parsons, J.T. (2001). Cortactin: coupling membrane dynamics to cortical actin assembly. Oncogene *20*, 6418-6434.

Xu, J., Bai, X.H., Lodyga, M., Han, B., Xiao, H., Keshavjee, S., Hu, J., Zhang, H., Yang, B.B., and Liu, M. (2007). XB130, a novel adaptor protein for signal transduction. J Biol Chem *282*, 16401-16412.

Yamaguchi, H., Pixley, F., and Condeelis, J. (2006). Invadopodia and podosomes in tumor invasion. Eur J Cell Biol *85*, 213-218.

Yeatman, T.J. (2004). A renaissance for SRC. Nat Rev Cancer *4*, 470-480.

Yilmaz, M., and Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer Metastasis Rev *28*, 15-33.

Zamir, E., and Geiger, B. (2001). Molecular complexity and dynamics of cell-matrix adhesions. J Cell Sci *114*, 3583-3590.

Zhang, J., Park, S.I., Artime, M.C., Summy, J.M., Shah, A.N., Bomser, J.A., Dorfleutner, A., Flynn, D.C., and Gallick, G.E. (2007). AFAP-110 is overexpressed in prostate cancer and contributes to tumorigenic growth by regulating focal contacts. J Clin Invest *117*, 2962-2973.

Figure 1A

```
AFAP1              ------MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQE 54
AFAP1L1            MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
AFAP1L2           MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
Ciona.intestinalis  ---MNEWTHVLKEKAAAFTEADNLPAPPVAQRPASSMPEHGAVKRVGSKSDLPVFAINDK 57
Consensus                 *** E    * D*   * **     ***        *       *

AFAP1             TANSLP-----------------------APPQMPLPEIP-QPWLP-PDSGPP----P 83
AFAP1L1           DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPP----P 116
AFAP1L2           TINKQQNAESQGKAPEEQGLLP-------NGEPSQHSSAPQKSLPDLPPPKMIPERKQLA 113
Ciona.intestinalis  RIDDGP---------------------------------------------------- 63
Consensus              **

AFAP1             LPTSSLPEGYYEEAVPLSPGKAPEYITSN------------------------------ 112
AFAP1L1           LPNKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELK 176
AFAP1L2           IPKTESPEGYYEEAEPYDTSLNE------------------------------------ 136
Ciona.intestinalis  LTSPTSPSNGLNNIG--------------------------------------------- 78
Consensus          ***   P**  **

AFAP1             ---YDSDAMSSSYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQW 169
AFAP1L1           SSYNDSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQW 236
AFAP1L2           ----DGEAVSSSYESYDEED-GSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQW 191
Ciona.intestinalis  -----KGLKSAVTDIFKKKDNMKKKISMAEQLG-----------ADICGSLNVYS-DGKW 121
Consensus                 S*   * *****  K  *    Q             * IC* L   * G*W

AFAP1             TKLLCVIKDTKLLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVL 229
AFAP1L1           AKQLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVL 296
AFAP1L2           AKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVL 251
Ciona.intestinalis  PKKLCTIRDNTLNVFG--KDESPEQSVILHGCDLTPGFGDP-VKKFVFKLTKNKQDLLLM 178
Consensus          *K L *I**  L  *  KD**P*  * L  ***      *   K** ***T    * ***

AFAP1             AVQSKEQAEQWLKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVV 289
AFAP1L1           ALQSREQAEEWLKVIREVSKPVGGAEGVEVPRSP-VLLCKLDLDKRLSQEKQTSDSDSVG 355
AFAP1L2           GLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTP--------DAQRFNCQKPDIAEKYLS 303
Ciona.intestinalis  EASNSAEMGKWVGMLIAETG----------------------CAEMPDHVPEQENVYLE 215
Consensus             **  *  *W* **                          **  *          *

AFAP1             ENGITTCNGK----EQVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKK-PSTDEQTSSA 344
AFAP1L1           VGDNCSTLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSST 415
AFAP1L2           ASEYGSSVDG-------HPEVPETKDVKKKCS--AGLKLSNLMNLGRKK--STSLEP--V 350
Ciona.intestinalis  TDTYESVMNT------------------------------------------------- 225
Consensus           *    *   *

AFAP1             EEDVPTCGYLNVLSNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSK 404
AFAP1L1           EEEVPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPR 475
AFAP1L2           ERSLETSSYLNVLVNSQWKSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPD 410
Ciona.intestinalis  ---------------------------ARKVFHDTVKMKKERNPHSSSSQTAPPPIPEE 257
Consensus                                     ***D  *    * *  *   * P

AFAP1             HPLTFRLLRNGQEVAVLEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQ 464
AFAP1L1           HPFAFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVS 535
AFAP1L2           HLYSFRILHKGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVS 470
Ciona.intestinalis  DDSIYLEVVASAPPPTLATTPTKAHKQQKGVNKDCEPVASTRKQKPRNGAARSRSGSSEQ 317
Consensus          *   *   *       * L ** **  * G*          *       *    *    *

AFAP1             TAKQTFCFMNRRVISANPYLGGTSNG----------------YAHPSGTALHYDDVPCI 507
AFAP1L1           AGRNSFLYAR-----------SCQNQ----------------WPEPR----VYDDVPYE 563
AFAP1L2           AAKNSLLLMQRKFSEPNTYIDGLPSQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNE 530
Ciona.intestinalis  SGGESTVNTN------------------------------------------------K 328
Consensus          ** **    *

AFAP1             NGSL------------------------------------------------------ 511
AFAP1L1           KMQD------------------------------------------------------ 567
AFAP1L2           RESDRVYLDLTPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCP 590
Ciona.intestinalis  KKHK------------------------------------------------------ 332
Consensus          *

AFAP1             -------------KGKKPPVASNGVTGKGKTLSSQPKKADPAAVVKRTGS---------N 549
AFAP1L1           ------------EEPERPTGAQ-VKRHASSCSEKSHRVDPQVKVKRHAS---------S 604
```

```
AFAP1L2            ENLGEQQLESLEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVT 650
Ciona.intestinalis -------------QDKERGIHKKAAEVDDERAKQHVRTINVISQVPTQKS---------- 369
Consensus                       *  *        *  *  *       *        *          S

AFAP1              AAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDLRAAIEVNAGRKPQA 609
AFAP1L1            ANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGAK-LK 663
AFAP1L2            SAEIKLGKNRTEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLA 710
Ciona.intestinalis ----------VSEEGYNLKSRENALKARKQEILQNLKVLRVKKTELNRYVEAAKSAKEKC 419
Consensus                    **  *  *    *** L  *K* *  *L  LR ** *L*  *     * *

AFAP1              ILEEKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGLAIEPKSGTSSPQSP 669
AFAP1L1            ALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSK-----PKS- 717
AFAP1L2            SLEQKLKEIDEECRGEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTT--------- 761
Ciona.intestinalis NLIRPLDEVKENYTKLENELVNIERELRNVRRLSNTAIEQPSNIRPKTKVG--------- 470
Consensus           L * *  ** *      E * ***E **  V*    * *        *      *

AFAP1              VFRHRTLENSPISSCDTSDTEGPVPVNSAAVLKKSQAAPGSSPCRGHVLRKAKEWELKNG 729
AFAP1L1            ----GETANKPQNSVP----EQPLPVNCVSELRKRSPSIVAS-NQGRVLQKAKEWEMKKT 768
AFAP1L2            ---HLENVSPRPKAVTPASAPDCTPVNSATTLKNRPLSVVVT-GKGTVLQKAKEWEKKGA 817
Ciona.intestinalis -------------------ETAIPT------------------GKVASRAKIFENLEA 491
Consensus                               P*                   G V  *AK *E

AFAP1              T- 730
AFAP1L1            --
AFAP1L2            S- 818
Ciona.intestinalis KK 493
Consensus
```

Figure 1B

```
AFAP1.PH1domain      ------------------------------------------------------------
AFAP1L1.PH1domain    ------------------------------------------------------------
AFAP1L2.PH1domain    ------------------------------------------------------------
Ciona.intestinalis   MNEWTHVLKEKAAAFTEADNLPAPPVAQRPASSMPEHGAVKRVGSKSDLPVFAINDKRID 60
Consensus

AFAP1.PH1domain      ------------------------------------------------DAKICAFLLRKKRFG 15
AFAP1L1.PH1domain    ------------------------------------------------ECRICAFLLRKKRFG 15
AFAP1L2.PH1domain    ------------------------------------------------DARICAFLWRKKWLG 15
Ciona.intestinalis   DGPLTSPTSPSNGLNNIGKGLKSAVTDIFKKKDNMKKKISMAEQLGADICGSLNVYS-DG 119
Consensus                                                         * IC* L   *  G

AFAP1.PH1domain      QWTKLLCVIKDTKLLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPL 75
AFAP1L1.PH1domain    QWAKQLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVL 75
AFAP1L2.PH1domain    QWAKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVI 75
Ciona.intestinalis   KWPKKLCTIRDNTLNVFG--KDESPEQSVILHGCDLTPGFGDP-VKKFVFKLTKNKQDLL 176
Consensus            *W*K L *I**  L  *   KD**P* * L ***     *   K** ***T    * *

AFAP1.PH1domain      VLAVQSKEQAEQWLKVIKEAYS--------------------------------------- 97
AFAP1L1.PH1domain    VLALQSREQAEEWLKVIREVSK--------------------------------------- 97
AFAP1L2.PH1domain    VLGLQSKDQAEQWLRVIQEVSG--------------------------------------- 97
Ciona.intestinalis   LMEASNSAEMGKWVGMLIAETGCAEMPDHVPEQENVYLETDTYESVMNTARKVFHDTVKM 236
Consensus            ** ** * *W* **

AFAP1.PH1domain      ------------------------------------------------------------
AFAP1L1.PH1domain    ------------------------------------------------------------
AFAP1L2.PH1domain    ------------------------------------------------------------
Ciona.intestinalis   KKERNPHSSSSQTAPPPIPEEDDSIYLEVVASAPPPTLATTPTKAHKQQKGVNKDCEPVA 296
Consensus

AFAP1.PH1domain      ------------------------------------------------------------
AFAP1L1.PH1domain    ------------------------------------------------------------
AFAP1L2.PH1domain    ------------------------------------------------------------
Ciona.intestinalis   STRKQKPRNGAARSRSGSSEQSGGESTVNTNKKKHKQDKERGIHKKAAEVDDERAKQHVR 356
Consensus

AFAP1.PH1domain      ------------------------------------------------------------
AFAP1L1.PH1domain    ------------------------------------------------------------
AFAP1L2.PH1domain    ------------------------------------------------------------
Ciona.intestinalis   TINVISQVPTQKSVSEEGYNLKSRENALKARKQEILQNLKVLRVKKTELNRYVEAAKSAK 416
Consensus

AFAP1.PH1domain      ------------------------------------------------------------
AFAP1L1.PH1domain    ------------------------------------------------------------
AFAP1L2.PH1domain    ------------------------------------------------------------
Ciona.intestinalis   EKCNLIRPLDEVKENYTKLENELVNIERELRNVRRLSNTAIEQPSNIRPKTKVGETAIPT 476
Consensus

AFAP1.PH1domain      -----------------
AFAP1L1.PH1domain    -----------------
AFAP1L2.PH1domain    -----------------
Ciona.intestinalis   GKVASRAKIFENLEAKK 493
Consensus
```

Figure 1. Similarity between human AFAP family members and *Ciona intestinalis*

(A) The 493 amino acid sequence of a predicted AFAP homolog in *Ciona intestinalis* was compared by ClustalW2 alignment (50) with the 730 amino acid sequence of AFAP1, the 768 amino acid sequence of AFAP1L1 and the 818 amino acid sequence of AFAP1L2. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

(B) The highly conserved PH1 domains of the human AFAP family members were compared with the amino acid sequence of the *Ciona intestinalis* predicted AFAP homolog using ClustalW2 analysis (50). Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 2A

Figure 2B

LEGEND

── x1 branch length		
── x10 branch length	**Gene ID** current gene	■ speciation node
── x100 branch length	**Gene ID** within-sp. paralog	■ duplication node
		■ ambiguous node

collapsed sub-tree

◀ collapsed (current gene)

◀ collapsed (paralog)

AA alignment match/mismatch

AA consensus > 66% (mis)match

AA consensus > 33% (mis)match

AA alignment gap

52
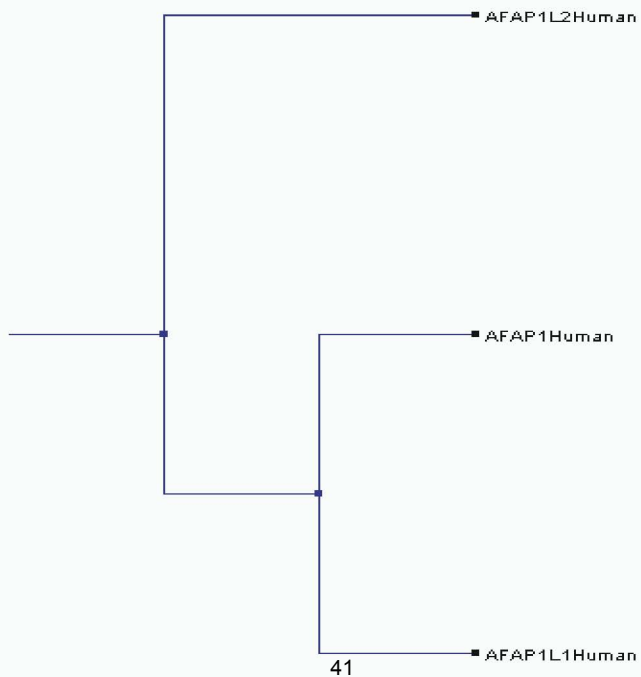
Figure 2. Cladistic analysis of AFAP family members

(A) Full length amino acid sequences of AFAP family members were compared by ClustalW2 alignment (50) and a cladogram created so as estimate a shared ancestry between AFAP family members.

(B) Full length amino acid sequences of all AFAP family members across species that are deposited in the Ensembl database were used to create a detailed phylogram of all AFAP family member sequences known. Red nodes indicate a duplication event while blue nodes indicate a speciation event. The alignment of shared amino acid similarity between all AFAP family member sequences is shown.

Figure 3

Figure 3. Phylogenetic analysis of AFAP1

Full length amino acid sequences for AFAP1 across human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), mouse (*Mus musculus*), *Ciona intestinalis*, macaque (*Macaca mulatta*), *Xenopus tropicalis*, zebrafish (*Danio rerio*) and Anole lizard (*Anolie carolinesis*) were compared by Clustal W2 alignment (50) and a detailed phylogram created to compare the sequential and temporal relationship of ancestry in AFAP1 sequences.

Figure 4

```
AFAP1.human        MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQETANSLP 60
AFAP1.chimpanzee   MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQETANSLP 60
AFAP1.mouse        MEELIVELRLFLELLDHEYLTSTVREKKAVLTNILLRLQSSKGFEVKDHAQKAEANN-LP 59
AFAP1.lizard       MEELIVELQLFLQLLDHEYLTSTVREKKAVLTNILLRIQSSKDFDLKEKVQKQEVVNSLP 60
AFAP1.zebrafish    MEELLAELRVFLELLDREYLTAGVREKKQQILNILHRVLATR-----EPSCKTEIHTSLP 55
Consensus          MEEL**EL**FL*LLD*EYLT* VREKK * NIL R* ***       *   K E  * LP

AFAP1.human        APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
AFAP1.chimpanzee   APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
AFAP1.mouse        APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 119
AFAP1.lizard       APPQMPLPEIPQQWLPPDNGPPPLPSSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
AFAP1.zebrafish    APPQMPLPEIPHPWMPPDNGPPPLPSSSLPEGYYEEAVPLGPGKAPEYITSNYDSDAMSS 115
Consensus          APPQMPLPEIP* W*PPD*GPPPLP*SSLPEGYYEEAVPL*PGKAPEYITSNYDSDAMSS

AFAP1.human        SYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 180
AFAP1.chimpanzee   SYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 180
AFAP1.mouse        SYESYDEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 179
AFAP1.lizard       SYESYDEEDEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKENK 180
AFAP1.zebrafish    SYESYDEEEEDGKGQKMRHQWPSEEASMDLVKDARICAFLLRKKRFGQWTKLLCVIKDNK 175
Consensus          SYESYDEE*EDGKG*K RHQWPSEEASMDLVKDA*ICAFLLRKKRFGQWTKLLCVIK**K

AFAP1.human        LLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 240
AFAP1.chimpanzee   LLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 240
AFAP1.mouse        LLCYKSSKDQQPQMELPLQGCSITYIPRDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 239
AFAP1.lizard       LLCYKSSKDQQPQMELLLNGCSIIYIPKDSKKKKHELKITHQGQDALVLAVQSKEQADQW 240
AFAP1.zebrafish    LLCYKSSKDQTPQMELLLSGCSITHIPKDGKKKKHELKIVHQGADALVLAVQSKEQAEQW 235
Consensus          LLCYKSSKDQ PQMEL L*GC*I *IP*D*KKKKHELKI**QG D*LVLAVQSKEAQ*QW

AFAP1.human        LKVIKEAYSGCSGPVDSECPPPPSS--PVHKAELEKKLSSERPSSDGEGVVENGIT-TCN 297
AFAP1.chimpanzee   LKVIKEAYSGCSGPVDSECPPPPSS--PVHKAELEKKLSSERPSSDGEGVVENGIT-TCN 297
AFAP1.mouse        LKVIKEAYSGCSGPVDPECSPPPSTSAPVNKAELEKKLSSERPSSDGEGGVENGVT-TCN 298
AFAP1.lizard       LKIIKEVCSNCVGTTDSDG---PSSNSPVHKTELEKKLSSERPSSDGEGTVENGIASICN 297
AFAP1.zebrafish    LKVMKEVCSNGNGVVDCDG---AGSGSTVHKAELEKKLSCDRPSSDGEPCHENGIS---D 289
Consensus          LK**KE* S*  G *D *    *** *V*K*ELEKKLS**RPSSDGE   ENG**   *

AFAP1.human        GKEQVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVL 357
AFAP1.chimpanzee   GKEQVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVL 357
AFAP1.mouse        GKEQAKRKKPSKSEAKGTVSKVTGKKITKIIGLGKKKPSTDEQTSSAEEDVPTCGYLNVL 358
AFAP1.lizard       GKEQVKRKKSSKTDSKSTVSKVTGKKITKIIGLGKKKPSTDEQTSSAEEDIPTCGYLNVL 357
AFAP1.zebrafish    GKDPAKGKKNSKSEQKGTVGRVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVL 349
Consensus          GK* *K KK SK** K*TV**VTGKITTKII*LGKKKPSTDEQTSSAEE* PTCGYNNVL

AFAP1.human        SNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSKHPLTFRLLRNGQE 417
AFAP1.chimpanzee   SNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDCKHPLTFRLLRNGQE 417
AFAP1.mouse        SNSRWRERWCRVKDSKLILHKDRADLKTHLVSIPLRGCEVIPGLDSKHPLTFRLLRNGQE 418
AFAP1.lizard       SNNRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSKHPLTFRLLRNAQE 417
AFAP1.zebrafish    SNNRWRERWCQLKDNQLFLHKDRADLKTHMASLPLRGCEVIPGLDSKHPFAFRLLRNGQE 409
Consensus          SN*RWRERWV**KD**L**HKDR*DLKTH**S*PLRGCEVIPGLD*KHP**FRLLRN*QE

AFAP1.human        VAVLEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQTAKQTFCFMNRRV 477
AFAP1.chimpanzee   VAVLEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQTAKQTFCFMNRRV 477
AFAP1.mouse        VAVLEASSSEDMGRWIGILLAETGSSTDPGALHYDYIDVEMSANVIQTAKQTFCFMNRRA 478
AFAP1.lizard       VAVLEASSSEDMGRWIGILLAESGSSADPGTLHYDYIDVDITASVIQAAKQTFCFMNRRV 477
AFAP1.zebrafish    VAVLEASSSESMGRWLGVLLAETGSTTDPAALHYDYIDVETTANVIQLAKQSFCFTSKRA 469
Consensus          VAVLEASSSE*MGRW*G*LLAE*GS**DP *LHYDYIDV* *A*NIQ AKQ*FCF **R*

AFAP1.human        ISANPYLGGTSNGYAHPSGTALHYDDVPCINGSLKGKKPPVASNGVTGKGKTLSSQPKK- 536
AFAP1.chimpanzee   ISANPYLGGTSNGYAHPSGTALHYDDVPCINGSLKGKKPPVASNGVTGKGKTLSSQPKK- 536
AFAP1.mouse        VSTSPYLGSLSNGYAHPSGTALHYDDVPCVNGSLKNKKPPASSNGVPVKGKAPSSQQKK- 537
AFAP1.lizard       LSTNPYRGNPTNGYACPSGMALHYDDVPCINGSFKGKKVSTATNGVVGKGRTLNNPQKK- 536
AFAP1.zebrafish    VSPNPYLDNPVNGYACPTGVALHYDDVPCINGTMKGKK-GLITNGFGAKKLDKNQPKKAN 528
Consensus          *S**PY **  NGYA P*G ALHYDDVPC*NG**K*KK    *NG* K    **  K
```

```
AFAP1.human          -ADPAAVVKRTGSNAAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDL 595
AFAP1.chimpanzee     -ADPAAVVKRMGSNAAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDL 595
AFAP1.mouse          -VETAGGVKRTASNAEQYKYGKNRVEADAKRLQSKEEELLKRKEALRNRLAQLRKERKDL 596
AFAP1.lizard         -SEFSSNVKRSTSSAEQYRYGKNRVEADAKKLQTKEEELLKKKEALRNRLAQLRKERKDL 595
AFAP1.zebrafish      GISSTLPVKRNNSSVDQYKYGKNRVEADAKKLQAKEEELMRKKQEIRNRLTQLKKDRKDL 588
Conensus              *  *   VKR  S** QY*YGKNRVEADAK*LQ*KEEEL***K* *RNRL*QL*K*RKDL


AFAP1.human          RAAIEVNAGRKPQAILEEKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGL 655
AFAP1.chimpanzee     RAAIEVNAGRKPQAILEEKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGL 655
AFAP1.mouse          RAAIEVNAGRKTQAALEDKLKRLEEECKQREAERVSLELELTEVKESLKKALAGGVTLGL 656
AFAP1.lizard         RAALEANVGRKPLIILEDKLKKLEEECKLKESERVSLELELTEVKENLKKALAGGITLGL 655
AFAP1.zebrafish      RTALENNTAKRSQASLTERLKKVEDECKLKEEERVSLELELTEVKESLKKALNGGVTLGL 648
Consensus            R*A*E N*****   L **LK**E*EC* *E ERVSLELELTEVKE*LKKAL GG*TLGL

AFAP1.human          AIEPKSGTSSPQSPVFRHRTLENSPISSCDTSDTEG-PVPVNSAAVLKKSQAAPGSSPCR 714
AFAP1.chimpanzee     AIEPKSGTSSPQSPVFRHRTLENSPISSCDTSDTEG-PVPVNSAAVLKKSQAAPGSSPCR 714
AFAP1.mouse          AIEPRSGTSSPQSPVFRHRTLENSPISSCDTSDAEG-PLPVNSAAVLKKSQPSSGSSPCR 715
AFAP1.lizard         AIEPKSGTSTAQSPVLKHQTLENSPISSCDTSDTET-SVPVNSAVVMKR-HSSSSSSPCR 713
AFAP1.zebrafish      TIEPKTGSSSPQSPVLMRRTVDNSPISSCNTSDTETCSLPVNSASLLRR-QTQQKASPVR 707
Consensus            *IEP**G*S**QSPV* **T**NSPISSC*TSD*E  **PVNSA **** **   *SP R

AFAP1.human          GHVLRKAKEWELKNGT 730
AFAP1.chimpanzee     GHVLRKAKEWELKNGT 730
AFAP1.mouse          GHVLQKAKEWELKNGT 731
AFAP1.lizard         GHVLQKAKEWELKNGT 729
AFAP1.zebrafish      GHVLRKAKEWEMKSGT 723
Consensus            GHVL*KAKEWE*K*GT
```

Figure 4. Comparison of AFAP1 sequences across species

The 730 amino acid sequence of human AFAP1 was compared by ClustalW2 alignment (50) with the 730 amino acid sequence of chimpanzee AFAP1, the 731 amino acid sequence of mouse AFAP1, the 729 amino acid sequence of Anole lizard AFAP1 and the 723 amino acid sequence of zebrafish AFAP1. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 5

```
AFAP1L1.human        -MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNT 59
AFAP1L1.chimpanzee   -MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNT 59
AFAP1L1.mouse        -MDRSRVLEQLIPELTGLLSLLDHEYLSDSTLEKKMAVASLLQSLQPLPAKEVSFLYVNT 59
AFAP1L1.lizard       --FLSTVLDQLLPELNVLLKLLDHEYLSSTTMEKKTAVSNILQKLQPPTGKDVNYMYMNT 58
AFAP1L1.zebrafish    MEINSKPMELLVTELNMLLKLLDHETLSCATEEKKMAVKNLLRQLQPS-VTAKDYMYVNT 59
Consensus                  *   ** L**EL* LL*LLDHE LS *T EKK AV **L**LQP   *   ***Y*NT


AFAP1L1.human        ADLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLR-NAADLPPPLP 118
AFAP1L1.chimpanzee   ADLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLR-NAADLPPPLP 118
AFAP1L1.mouse        ADLHSGPSFVESLFEEFDCDLSDLRDMS-DDGEPSKGASPEPTKSPSLRSAAADVPPPLP 118
AFAP1L1.lizard       ETLHNGTSFVESLFEGFDCDLSNLQDMQEDEVDTKEGISLELSKSQLAKSISGEPPPPLP 118
AFAP1L1.zebrafish    SVYRNGTSFVESLFETFDCDLGDLKVEMEDQKKEPE------ANHTVTKPSKTDSPPPLP 113
Consensus                   **G*SFVESLFE FDCDL**L*    D* * *       **     *  * PPPLP


AFAP1L1.human        NKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSS 178
AFAP1L1.chimpanzee   NKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSS 178
AFAP1L1.mouse        NKPPPEDYYEEALPLGPGKSPEYISSHNGCSPAQSIVDGYYEDADNSYPTTRMNGELKNS 178
AFAP1L1.lizard       TTPPPEDYYEEALPLGPGKAPEYISHSNSSPPNSIEDGYYEDADSNYPVTRMNGEQKNS 178
AFAP1L1.zebrafish    NTPPPEDYYEEAVPLSPGKMPEYITSRSSSSPPNSIEDGYYEDAENNYPTTQVNGRRKNS 173
Consensus            **PPPEDYYEEA*PL*PGK PEYI*S****SP**SI EDGYEDA***YP*T**NG* K*S


AFAP1L1.human        YNDSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAK 238
AFAP1L1.chimpanzee   YNDSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAK 238
AFAP1L1.mouse        YNDSDAMSSSYESYDEEEEEKGRQPKHQWPSEEASMHLVRDCRICAFLLRKKRFGQWAK 238
AFAP1L1.lizard       YNDSDAMSSSYESYDEEEEDGKGQRLTHQWPSEEASMNLVKDCRICAFLLRKKRFGQWAK 238
AFAP1L1.zebrafish    YNDSDALSSSYESYDEEEEE-KGQRLTHQWPSEENSMAPVRDCHICAFLLRKKRFGQWAK 232
Consensus            YNDSDA*SSSYESYDEEEE* K* *   HQWPSEE SM  V**C*ICAFLLRKKFRGQWAK


AFAP1L1.human        QLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLAL 298
AFAP1L1.chimpanzee   QLTVIREDQLLCYKSSKDQQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLAL 298
AFAP1L1.mouse        QLTVIKEEQLLCYKSSKDRQPHLRLALDVCTVIYVPKDSRHKRHELRFSQGATEVLVLAL 298
AFAP1L1.lizard       QLTVIKDNKLLCYKSSKHRQPHLEVPLSVCNVVYVPKDGRRKKHELRFSLPGAEALVLAV 298
AFAP1L1.zebrafish    QLTVIRENRLQCYKSSKDQSPYTDIPLSLCSVIYVPKDGRRKKHELRFTLPGGEALVLAV 292
Consensus            QLTVI****L CYKSSK***P*  **L* C***YVPKD*R*K*HELRF*  * E*LVLA*


AFAP1L1.human        QSREQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGD 358
AFAP1L1.chimpanzee   QSREQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSMGVGD 358
AFAP1L1.mouse        QSREQAEEWLKVIREVSRPIVGAEGLEVPRSPVILCKADQDKRLSQEKQNSDSDSLGMND 358
AFAP1L1.lizard       QSKEQAEEWLKVMKEVSN---GQSGTEVLTSPMLTCKMDHDKRSSQDKHTSDSDSVATAE 355
AFAP1L1.zebrafish    QSKEQAEKWLHVVRDVTG---QGNGLDSPSSPMIPKKIELDKWCSAEKQTSDSDSMPSGE 349
Consensus            QS*EQAE*WL*V***V*       *G *   SP**  K * DK  S *K**SDSDS*   *


AFAP1L1.human        NCSTLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEE 418
AFAP1L1.chimpanzee   NCSTLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEE 418
AFAP1L1.mouse        SGSTLGRREACEHGKGKKNSLAELKGSMSRAAGRKITRIISFSKKKALSEDLQTFSSEDE 418
AFAP1L1.lizard       NCSSMTRREAQEQGKGKKSGLAELKGSVSRAAGRKITRIISFSKKKPSPEDTQTSSTEED 415
AFAP1L1.zebrafish    S-----ARDIRENGKPKRGALSELTGTVSRAAGRKITRIISFSKRKP-PLPGDSRSSFDH 403
Consensus            *      R*  **GK K***L*EL*G**SRAAGRKITRII*FSK*K* *   ** S* **


AFAP1L1.human        VPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPF 478
AFAP1L1.chimpanzee   VPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPF 478
AFAP1L1.mouse        VPCCGYLNVLVNQGWKERWCRLRCNTLYFHKDRTDLHTHVNSIALRGCEVAPGFGPRHPF 478
AFAP1L1.lizard       IPCCGYLNVLVNQCWKERWCRLKGNTLYFHKDRTDLRTHVNAIVLRGCEVAPGLGPKHPL 475
AFAP1L1.zebrafish    DPRCGYVGVLVNRCWREHWCRVRAGSLYLYQEKGEQRVPHTTVGLKGCEVVPGLGPKHPF 463
Consensus             P CGY**VLVN* W*E*WCR** **LY***** * **  *** L*GCEV*PG*GP*HP*


AFAP1L1.human        AFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGR 538
AFAP1L1.chimpanzee   AFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGR 538
AFAP1L1.mouse        AFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSKVTPEALHYDYVDVETLTSIVSAGR 538
AFAP1L1.lizard       AFRILRNGQEVSALEANSYEDLGRWLGLLLVETGSQTAPEALHYDYVDVEKIANIINAVR 535
AFAP1L1.zebrafish    ALRILKGGAEVAALEASCSEDMGRWLGVLLAETGSSADPESLHYDYVDVETIANIRTAAR 523
Consensus            A*RIL** EV* LEA** ED*GRWLG*LL*E GS * PE*LHYDYVDVE****I *A R
```

```
AFAP1L1.human       NSFLYARSCQNQWPEPR--VYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQV 596
AFAP1L1.chimpanzee  NSFLYARSCQNQWPEPR--VYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQV 596
AFAP1L1.mouse       NSFLYAQSCQDQWPEPR--IYDEVPYEKVQDEEPQRPTGAQVKRHASSCSEKSHRADPQV 596
AFAP1L1.lizard      HSYMWASSSVENQGDSSRVLYDEVPYEKVELEKSRWPSGTQVKRHGSSCSEKSRRVDPQV 595
AFAP1L1.zebrafish   HSFLWATSTDSR------TYDEVPFETIEQENERLRGRAQTKRRSSFSSSDTGKPSPQI 576
Consensus           *S***A S  **        YD*VP*E*** E* *    *Q*KR**S *S*** * *PQ*


AFAP1L1.human       KVKRHASSANQY-KYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIR 655
AFAP1L1.chimpanzee  KVKRHASSANQY-KYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIR 655
AFAP1L1.mouse       KVKRHASSANQY-KYGKNRAEEDARRYLVEKERLEKEKETIRTELTALRQEKKELKEAIR 655
AFAP1L1.lizard      KVKRHASNANNY-RYGKNRAEEDARRFLTEKDKLEKEKASIRTEIMGLRKEKRELREAMK 654
AFAP1L1.zebrafish   TLKRHGSNANQYGRYGKTRAQEDARRYLKEKEDLETEIDSIRTVLVALRKKKREAKEKMK 636
Consensus           **KRH*S*AN*Y *YGK*RA*EDARR*L EK* LE*E  *IRT * *LR**K*E *E **

AFAP1L1.human       SSPGAKLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKP 715
AFAP1L1.chimpanzee  SSPGAKLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSMSSKP 715
AFAP1L1.mouse       NNPGAKSKALEEAVATLEAQCRAKEEQRIDLELKLVAVKERLQQSLAGGPALGLSVSNKN 715
AFAP1L1.lizard      SSSGKDLKELEHRVAALEEQCKANEARRVDLELKLTEVKDRLKQSLAGGPALGLTVTTKP 714
AFAP1L1.zebrafish   SATDKQKLTLEECVTKLEDSCRVKEGERVDLELKLTQVKENLKKSLAGG---EMEAPTES 693
Consensus           * ** *   LE* V* LE *C***E *R*DLEELK* VK**L**SLAGG    *  ***

AFAP1L1.human       KSGETANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAKEWEMKKT- 768
AFAP1L1.chimpanzee  KSGETANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAKEWEMKKT- 768
AFAP1L1.mouse       KSQDTTNKPQSNAPEQSLPVNCVSELRKRSPSIVTSNQGRVLQKAKEWEMKKT- 768
AFAP1L1.lizard      ENKDVAIQPNGTPPEHLVPVNCAAEMRKRSPSLLPANKGNVLQKAK-------- 760
AFAP1L1.zebrafish   KPAHKTQRTEAQYMESFLPVNCASEMRKRPPSIYASTKGNVMQKAKEWESKKGT 747
Consensus           * * * ***    E  *PVNC**E*RKR*PS* ****G*V*QKAK
```

Figure 5. Comparison of AFAP1L1 sequences across species

The 768 amino acid human AFAP1L1 sequence was compared using ClustalW2 alignment (50) with the 768 amino acid sequence of chimpanzee AFAP1L1, the 768 amino acid sequence of mouse AFAP1L1, the 760 amino acid sequence of Anole lizard AFAP1L1 and the 747 amino acid sequence of zebrafish AFAP1L1. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 6

```
AFAP1L2.human        MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
AFAP1L2.chimpanzee   MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
AFAP1L2.mouse        ------------------------------------------------------MNKV 4
AFAP1L2.lizard       -----ALEQLLTELEDFLKILDKENLSSTAIVKKSFLADLLRLCTKSNGGDEEYIYMNKV 55
AFAP1L2.zebrafish    MDKHKVLEQLLEQLQKFLKILDVEKLSGNAKVQKGLLMELLQSYKSSNGGDEEYIYMNKV 60
Consensus                                                                  MNLV

AFAP1L2.human        TINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPPPKMIPERKQLAIPKTESP 120
AFAP1L2.chimpanzee   TINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPPPKMIPERKQLAIPKTESP 120
AFAP1L2.mouse        SVNGEQNSASPDKVPEEQGPLTNGEPSQHSSAPQKSLPDLPPPKMIPERKQPTVPKIESP 64
AFAP1L2.lizard       SIN-KQHGELEKSDKGHRDSLTNGDAEQHLSPPQKSLPDLPPSKIIPETKPYSGSKTESP 114
AFAP1L2.zebrafish    IVT----CQTQDKTNRDHRPEANGDPSKHIS--VKNPPEPPPPPRPVSKQKRAPAPASMEN 114
Consensus             **        *  **   *NG****H S   K* P* PP** *** K  * *    *

AFAP1L2.human        EGYYEEAEPYD-TSLNEDGEAVSSSYESYDEEDGSKGKS-APYQWPSPEAGIELMRDARI 178
AFAP1L2.chimpanzee   EGYYEEAEPYD-TSLNEDGEAVSSSYESYDEEDGSKGKS-APYQWPSPEAGIELMRDARI 178
AFAP1L2.mouse        EGYYEEAEPFD-RSINEDGEAVSSSYESYDEDENSKGKA-APYQWPSPEASIELMRDARI 122
AFAP1L2.lizard       EGYYEEAEPYG-ISLNDDGEAVSSSYESYDEEENTKGKS-APHQWPSPEASIELMKDARI 172
AFAP1L2.zebrafish    ESYYEDPQPYDPISINEDTEQLSSSYESYDEEEVTKGKSTAQHQWPSPEASIELMKDARI 174
Consensus            E*YYE***P**  S*N*D E *SSSYESYDE** *KGK* A *QWPSPEA*IELM*DARI

AFAP1L2.human        CAFLWRKKWLGQWAKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHK 238
AFAP1L2.chimpanzee   CAFLWRKKWLGQWAKQLCVIKDNRLMCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHK 238
AFAP1L2.mouse        CAFLWRKKWLGQWAKQLCVIRDTRLLCYKSSKDHSPQLDVNLRGSSVVHKEKQVRKKGHK 182
AFAP1L2.lizard       CAFLWRKKWLGQWAKQLCVIKDNRLLCYKTSKDHNPQLDVNLLGCSVIHKEKNVRKQEHK 232
AFAP1L2.zebrafish    CAFLWRKKWLGQWAKQLCVVREHRLLCYKSSKDQTPLLDISLLGCSVVYKEKQTKRKEHK 234
Consensus            CAFLWRKKWLGQWAKQLCV*** RL*CYK*SKD**P LD**L G*SV**KEK***** HK

AFAP1L2.human        LKITPMNADVIVLGLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIA 298
AFAP1L2.chimpanzee   LKITPMNADVIVLGLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIA 298
AFAP1L2.mouse        LKITPMNADVIVLGLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRLNCQKPDIA 242
AFAP1L2.lizard       LKIIPMNADVIVLGLQSKDQAEQWLRVIQETSGLYPDGWGEGNQYVPDSQRLSYPKVEAT 292
AFAP1L2.zebrafish    LKISPLGAEAIVLGLQSKEQAEQWLKVIQEISPKNTTGSD-----VTDSPTLICTKGEQS 289
Consensus            LKI P**A**IVLGLQSK*QAEQWL*VIQE S   * G *    **D*  *   K * *

AFAP1L2.human        EKYLSASEYGSSVDGHPEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSS 358
AFAP1L2.chimpanzee   EKYLSASEYGSSVDGHSEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSS 358
AFAP1L2.mouse        EKYLSAAEYGITINGHPEIPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPPERSLETSS 302
AFAP1L2.lizard       ERYSVASESGSSTDGHPELMETKDAKKKGTSGLKLSNLMNLGRKKSASLDSPERSLDTCT 352
AFAP1L2.zebrafish    ERYSVASESGSSTDSHAENMENKDVKKKYG---KFSNLMNIGKKKVCSLESPEKSVDTSG 346
Consensus            E*Y  A*E G * **H*E  E*KD*KKK     K*SNLMN*G*KK  SL** E*S**T*

AFAP1L2.human        YLNVLVNSQWKSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHLYSFRIL 418
AFAP1L2.chimpanzee   YLNVLVNSQWKSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHLYSFRIL 418
AFAP1L2.mouse        YLNVLVNSQWKSRWCFVRDSHLHFYQDRNRSKVAQQPLSLVGCDVLPDPSPDHLYSFRIL 362
AFAP1L2.lizard       YLNVLVNSQWKSRWCYVKDGQLHFYQDKNKTKTAQQPLNLIGCEIFPDPSPDHLYSFRIL 412
AFAP1L2.zebrafish    YLNVLVNTQWSRWCSVKDRQLWIYSDKSKGKVAQQPLSLEGCMVLPDPSPEHLYSFRIQ 406
Consensus            YLNVLVN*QW*SRWC V*D *L *Y*D*** K*VAQQP*L GC **PDPSP*HLYSFRI

AFAP1L2.human        HKGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVADDRVSCIVSAAKNSLLL 478
AFAP1L2.chimpanzee   HKGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVADDRVSCIVSAAKNSLLL 478
AFAP1L2.mouse        HNGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEELTYDYVDAERVSCIVSAAKTSLLL 422
AFAP1L2.lizard       HNGEERAMLEAKTCEEMGHWLGLLLSESGSKTDPEELTYDYVDADRVSCIVSAAKNSLLL 472
AFAP1L2.zebrafish    MDGEELAILEAKSSADMGHWLGLILSQTGTKTDPEDLSYDYVNSERISSIVNAAKTSMYL 466
Consensus             *GEE A LEAK** *MGHWLGL*LS**G*KTDPE***YDYV***R*S*IV*AAK*S* L

AFAP1L2.human        MQRKFSEPNTYIDGLP-SQDRQEELYDDVDLSELTA-AVEPTEEATP--VADDPNERESD 534
AFAP1L2.chimpanzee   MQRKFSEPNTYIDGLP-SQDRQEELYDDVELSELTA-AVEPTEEATP--VADDPNERESD 534
AFAP1L2.mouse        MQRKFSEPNTYIDGLP-SRDCQDDLYDDVEVSELIA-VVEPAEEAAP--AVDANSGSEPD 478
AFAP1L2.lizard       MQRKYSEPNTYIDNLPKEKEEQEELYDDVDVPESTMNKFNISHQEIP--LIESKSAGEQD 530
AFAP1L2.zebrafish    MQRRYSEPNTYTDSPPSDPQTCDDIYDDVPSTENEQEEVQEVQNGSEEGTVNAEEENGKD 526
Consensus            MQR**SEPNTY D* P * * ***YDDV *E     ** **       *  *   D
```

```
AFAP1L2.human       RVYLDLTPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCPENLG 594
AFAP1L2.chimpanzee  RVYLDLTPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADPGPGPTPDEPCIKCPENLG 594
AFAP1L2.mouse       RVYLDLTPVKSFLHSSSEAQAQASLPAVPHQDDVAETLTVDPKPGTTPEEPHTESPGDPE 538
AFAP1L2.lizard      RVYLDLTPVQSFVYSAGRKHGQLTPSVSPSLQRSVHKNSTDSDK----DLFTLYKEGELN 586
AFAP1L2.zebrafish   RVYLDLIPLRSFLHTSSAKTLGQNLPGDPHLPPAATLISP-PPPGPQLFGTVYNTPGLPT 585
Consensus           RVYLDL P**SF** **      * * *     *   *  *


AFAP1L2.human       EQQ-----LESLEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVK-DRLR 648
AFAP1L2.chimpanzee  EQQ-----LESLEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVK-DRLR 648
AFAP1L2.mouse       VQQRQPEVQESSEPIEPTPRITMVKLQAEQQRISFPANCPDTMASAPIAASPPVK-EKLR 597
AFAP1L2.lizard      RQS-----VEATDQLLP-LRTTTVKIQAQQQQIAFPQSAFEMKNATTIVAAPKEKGERPK 640
AFAP1L2.zebrafish   MIKYPIAADRTRNPEIQKKASPLPKTQSLPQTSAQTPQSPPTSRAR--AASTDRLLDRLK 643
Consensus               *      ** *           * K Q*  Q  * * **     *    A**    ** *

AFAP1L2.human       VT-----SAEIKLGKNRTEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKC 703
AFAP1L2.chimpanzee  VT-----SAEIKLGKNRTEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKC 703
AFAP1L2.mouse       VT-----SAEIKLGKNRTEAEVKRYTEEKERLERSKEEIRGHLAQLRREKRELKETLLRC 652
AFAP1L2.lizard      VA-----PVETKLGKNRTEAEVKRYSEERDRLEKEKEEIRSQLAQLRKDKRELKELLTNC 695
AFAP1L2.zebrafish   FSPAGPGSVEVKLGKNRTEADVRRYTDDRDRLEREREEVKNTLATLRKDRREVKDELSSC 703
Consensus           **      **E KLGKNRTEA*V*RY*****RLE***EE*** LA LR***RE*K* L  C

AFAP1L2.human       TDKEVLASLEQKLKEIDEECRGEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTTHL 763
AFAP1L2.chimpanzee  TDKEVLASLEQKLKEIDEECRGEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTTHL 763
AFAP1L2.mouse       TDKGVLAKLEQTLKKIDEECRMEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTTHL 712
AFAP1L2.lizard      TDKSILSTLEQNLKEIEEECKRKEDQRVDLELNLVEVKENLRKAESGPVTLGTAVDTTHL 755
AFAP1L2.zebrafish   QVQTELASLEARLKQMEETCREAERRRVEVELSLMEVKENLRKVESGPFTLGTTVDSSLL 763
Consensus              *  L**LE  LK***E C*  E *RV**EL***EVK*NL*K*E*GP*TLGT*VD** L

AFAP1L2.human       ENVSPRP--KAVTPASAP-------------DCTPVNSATTLKNRPLSVVVTGKGTVLQ 807
AFAP1L2.chimpanzee  ENVSPRP--KAVTPASAP-------------DCTPVNSATTLKNRPLSVMVTGKGTVLQ 807
AFAP1L2.mouse       DNMSPRPQPKAATPNPPP-------------DSTPVNSASVLKNRPLSVMVTGKGTVLQ 758
AFAP1L2.lizard      ENSSPKM--KVANPMNST-------------ESSPVNSAMALKNRPLSIMVTGKGTVLQ 799
AFAP1L2.zebrafish   DISVPKP-AAVSSPAPNTPNNTNINTPACTNNEDSPVNSATALKNRPPSVMAASKGNVLQ 822
Consensus           *   P*    * *P   *             * *PVNSA *LKNRP S*****KG*VLQ

AFAP1L2.human       KAKEWEKKGAS 818
AFAP1L2.chimpanzee  KAKEWEKKGAS 818
AFAP1L2.mouse       KAKEWEKKGAS 769
AFAP1L2.lizard      KAKEWEKKGAS 810
AFAP1L2.zebrafish   KAKEWEKKNTT 833
Consensus           KAKEWEKK***
```

Figure 6. Comparison of AFAP1L2 sequences across species

The 818 amino acid sequence of AFAP1L2 was compared by ClustalW2 alignment (50) with the 818 amino acid sequence of chimpanzee AFAP1L2, the 769 amino acid sequence of mouse AFAP1L2, the 810 amino acid sequence of Anole lizard AFAP1L2 and the 833 amino acid sequence of zebrafish AFAP1L2. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 7

```
AFAP1     ------MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQE 54
AFAP1L1   MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
AFAP1L2   MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
Consensus       *E*L* EL  *L**LD*E L**T* KK  ****L       * * *      *

AFAP1     TANSLP-----------------------APPQMPLPEIP-QPWLP-PDSGPP----P 83
AFAP1L1   DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPP----P 116
AFAP1L2   TINKQQNAESQGKAPEEQGLLP-------NGEPSQHSSAPQKSLPDLPPPKMIPERKQLA 113
Consensus      **                     P** * *    P L  * P      *

AFAP1     LPTSSLPEGYYEEAVPLSPGKAPEYITSN------------------------------ 112
AFAP1L1   LPNKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELK 176
AFAP1L2   IPKTESPEGYYEEAEPYDTSLNE----------------------------------- 136
Consensus *P**  PE*YYEEA P ***

AFAP1     ---YDSDAMSSSYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQW 169
AFAP1L1   SSYNDSDAMSSSYESYDEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQW 236
AFAP1L2   ----DGEAVSSSYESYDEED-GSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQW 191
Consensus     D**A*SSSYESYDEE*  *K* ** *QWPS EA***L*****ICAFL RKK *GQW

AFAP1     TKLLCVIKDTKLLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVL 229
AFAP1L1   AKQLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVL 296
AFAP1L2   AKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVL 251
Consensus *K L VI** *LLCYKSSKD**P** * L  *** *  K* **K*H*L**T   ** *VL

AFAP1     AVQSKEQAEQWLKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVV 289
AFAP1L1   ALQSREQAEEWLKVIREVSKPVGGAEGVEVPRSP-VLLCKLDLDKRLSQEKQTSDSDSVG 355
AFAP1L2   GLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTP--------DAQRFNCQKPDIAEKYLS 303
Consensus **QS**QAE*WL*VI*E*    *   *     *p         **** **     * *

AFAP1     ENGITTCNGK----EQVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKK-PSTDEQTSSA 344
AFAP1L1   VGDNCSTLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSST 415
AFAP1L2   ASEYGSSVDG-------HPEVPETKDVKKKCS--AGLKLSNLMNLGRKK--STSLEP--V 350
Consensus  *   *   *        * * ** * K * S  *G K*********KK * * **  *

AFAP1     EEDVPTCGYLNVLSNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSK 404
AFAP1L1   EEEVPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPR 475
AFAP1L2   ERSLETSSYLNVLVNSQWKSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPD 410
Consensus E***  **YLNVL N* W**RWC ** N L F**D*   **   ***L GCEV P* **

AFAP1     HPLTFRLLRNGQEVAVLEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQ 464
AFAP1L1   HPFAFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVS 535
AFAP1L2   HLYSFRILHKGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVS 470
Consensus H  *FR*L** *E*A LEA**SE*MG*W*G*LL E GS * PE * YDY*D**  *****

AFAP1     TAKQTFCFMNRRVISANPYLGGTSNG----------------YAHPSGTALHYDDVPCI 507
AFAP1L1   AGRNSFLYAR-----------SCQNQ----------------WPEPR----VYDDVPYE 563
AFAP1L2   AAKNSLLLMQRKFSEPNTYIDGLPSQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNE 530
Consensus ******   *         * *                      *P        D P

AFAP1     NGSL------------------------------------------------------- 511
AFAP1L1   KMQD------------------------------------------------------- 567
AFAP1L2   RESDRVYLDLTPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCP 590
Consensus                                                           * *

AFAP1     -------------KGKKPPVASNGVTGKGKTLSSQPKKADPAAVVKRTGS---------N 549
AFAP1L1   -------------EEPERPTGAQ-VKRHASSCSEKSHRVDPQVKVKRHAS---------S 604
AFAP1L2   ENLGEQQLESLEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVT 650
Consensus              * *     * * * **  S *  D* * *   *S         *

AFAP1     AAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDLRAAIEVNAGRKPQA 609
AFAP1L1   ANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGAK-LK 663
AFAP1L2   SAEIKLGKNRTEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLA 710
Consensus * * K GKNR*E ***R   **E*L K*KE *R *L  LR*E***L* **   ** *
```

52

```
AFAP1      ILEEKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGLAIEPKSGTSSPQSP 669
AFAP1L1    ALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSK-----PKS- 717
AFAP1L2    SLEQKLKEIDEECRGEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTT--------- 761
            LE* *  ** *CR *E *R**LEL**  VK* L*** AG  *LG *****

AFAP1      VFRHRTLENSPISSCDTSDTEGPVPVNSAAVLKKSQAAPGSSPCRGHVLRKAKEWELKNG 729
AFAP1L1    ----GETANKPQNSVP----EQPLPVNCVSELRKRSPSIVAS-NQGRVLQKAKEWEMKKT 768
AFAP1L2    ---HLENVSPRPKAVTPASAPDCTPVNSATTLKNRPLSVVVT-GKGTVLQKAKEWEKKGA 817
                 *   **             PVN*** L**   *   *  *G VL*KAKEWE K

AFAP1      T 730
AFAP1L1    -
AFAP1L2    S 818
```

Figure 7. Identity and similarity between human AFAP family members

The 730 amino acid sequence of AFAP1 was compared to the 768 amino acid sequence of AFAP1L1 and the 818 amino acid sequence of AFAP1L2 using ClustalW2 alignment (50). Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.
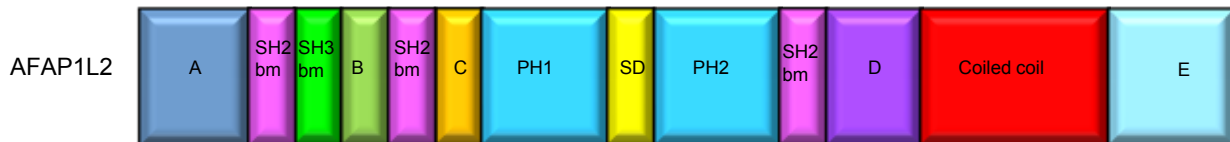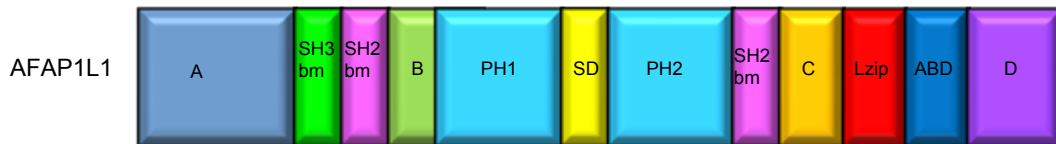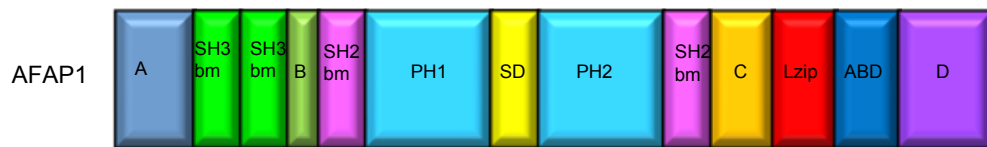
Figure 8A

Figure 8B

```
AFAP1.PH1      DAKICAFLLRKKRFGQWTKLLCVIKDTKLLCYKSSKDQQPQMELPLQGCNITYIPKDSKK 60
AFAP1L1.PH1    ECRICAFLLRKKRFGQWAKQLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRH 60
AFAP1L2.PH1    DARICAFLWRKKWLGQWAKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRK 60
Consensus      ***ICAFL RKK *GQW*K L VI** *LLCYKSSKD**P** * L  *** *  K* **

AFAP1.PH1      KKHELKITQQGTDPLVLAVQSKEQAEQWLKVIKEAYS 97
AFAP1L1.PH1    KRHELRFTQGATEVLVLALQSREQAEEWLKVIREVSK 97
AFAP1L2.PH1    KEHKLKITPMNADVIVLGLQSKDQAEQWLRVIQEVSG 97
Consensus      K*H*L**T   ** *VL**QS**QAE*WL*VI*E*




AFAP1.PH2      DVPTCGYLNVLSNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSKHP 60
AFAP1L1.PH2    EVPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHP 60
AFAP1L2.PH2    SLETSSYLNVLVNSQWKSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHL 60
Consensus      **  **YLNVL N* W**RWC ** N L F**D*   **   ***L GCEV P* ** H

AFAP1.PH2      LTFRLLRNGQEVAVLEASSSEDMGRWIGILLAETGSSTD 99
AFAP1L1.PH2    FAFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVT 99
AFAP1L2.PH2    YSFRILHKGEELAKLEAKSSEEMGHWLGLLLSE------ 93
Consensus       *FR*L** *E*A LEA**SE*MG*W*G*LL E
```

Figure 8. Comparison of AFAP family members

(A) Modular domain organization of AFAP family members was compared. SH3bm = SH3 binding motif, SH2bm = SH2 binding motif, PH1 = pleckstrin homology domain 1, PH2 = pleckstrin homology domain 2, SD = serine/threonine rich substrate domain, Lzip = leucine zipper, ABD = actin binding domain. Sequences that do not correlate with an identified type of modular domain or motif are labeled "A, B, C, D or E".

(B) The PH1 domains of AFAP family members were compared by ClustalW2 analysis (50). Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 9

AFAP1

AFAP1L1

AFAP1L2

56
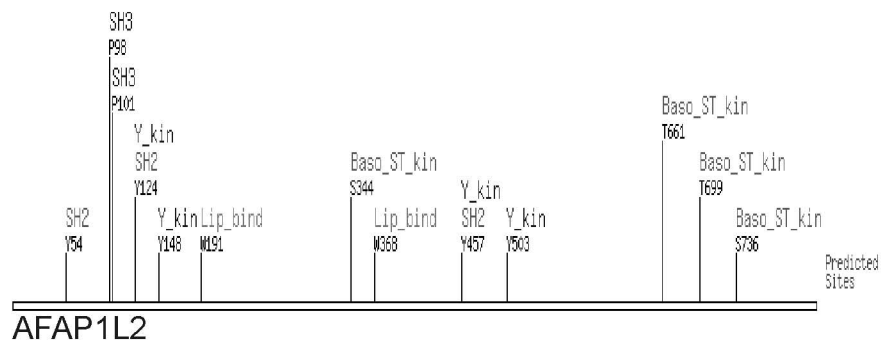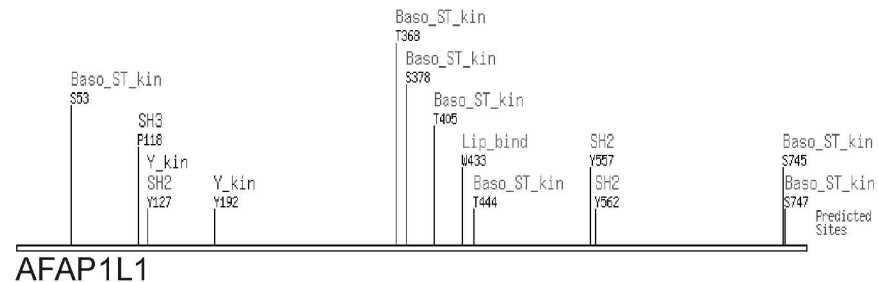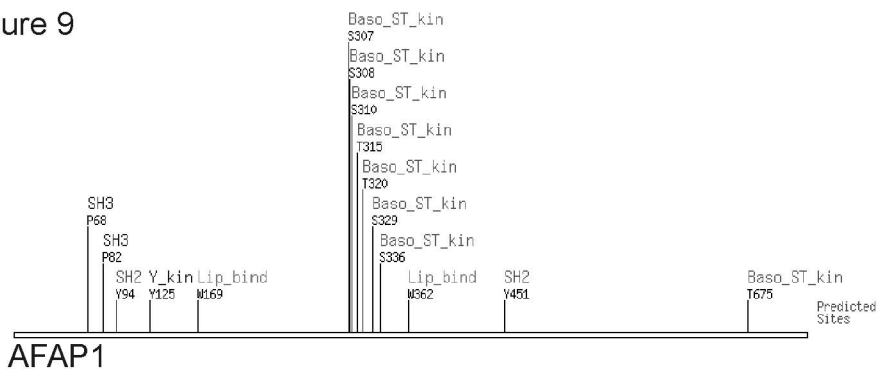
Figure 9. A Scansite scan determined potential sites for phosphorylation

The Scansite motif scanning progam (59) determined possible sites of protein interaction in AFAP family members. Predicted sites include Src homology 3 groups (SH3), Src homology 2 groups (SH2), tyrosine kinase groups (Y_kin), lipid binding groups (Lip_bind) and basophilic serine/threonine kinase groups (Baso_ST_kin).

CHAPTER 2

AFAP1L1 is a Novel Adaptor Protein of the AFAP Family that Interacts with Cortactin and Localizes to Invadosomes

Brandi N. Snyder[1,+], YoungJin Cho[2,+], Yong Qian[3], James Coad[4], Daniel C. Flynn[2,4] and Jess Cunnick[2]

[1]The Mary Babb Randolph Cancer Center and the Department of Cancer Biology, West Virginia University, Morgantown, WV, 26505
[2]The Department of Cell Biology, The Commonwealth Medical College, Scranton, PA, 18510
[3]Pathology and Physiology Research Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Morgantown, WV 26505
[4]Department of Pathology, West Virginia University, Morgantown, WV 26505

+These authors contributed equally to this work.

ABSTRACT

The actin-filament associated protein (AFAP) family of adaptor proteins consists of three members: AFAP1, AFAP1L1, and AFAP1L2/XB130 with AFAP1 being the best described as a cSrc binding partner and actin cross-linking protein. A homology search of AFAP1 recently identified AFAP1L1 which has a similar sequence, domain structure and cellular localization; however, based upon sequence variations, AFAP1L1 is hypothesized to have unique functions that are distinct from AFAP1. While AFAP1 has the ability to bind to the SH3 domain of the nonreceptor tyrosine kinase cSrc via an N-terminal SH3 binding motif, it was unable to bind cortactin. However, the SH3 binding motif of AFAP1L1 was more efficient at interacting with the SH3 domain of cortactin and not cSrc. AFAP1L1 was shown by fluorescence microscopy to decorate actin filaments and move to punctate actin structures and colocalize with cortactin, consistent with localization to invadosomes. Upon overexpression in A7r5 cells, AFAP1L1 had the ability to induce podosome formation and move to podosomes without stimulation. Immunohistochemical analysis of AFAP1L1 in human tissues shows differential expression when contrasted with AFAP1 with localization of AFAP1L1 to unique sites in muscle and the dentate nucleus of the brain where AFAP1 was not detectable. We hypothesize AFAP1L1 may play a similar role to AFAP1 in affecting changes in actin filaments and bridging interactions with binding partners, but we hypothesize that AFAP1L1 may forge unique protein interactions in which AFAP1 is less efficient, and these interactions may allow AFAP1L1 to affect invadosome formation.

INTRODUCTION

Adaptor proteins are non-enzymatic proteins that have the ability to link together different components of cellular signaling complexes through protein-binding motifs. Actin Filament-Associated Protein of 110 kilodaltons (kDa) (AFAP-110/AFAP1) is an adaptor protein with multiple protein binding motifs that is known to function as both an actin binding protein and a cSrc activating protein (Flynn et al., 1993; Qian et al., 2000). The AFAP1 protein binding motifs include two juxtaposed poly-proline rich Src homology 3 (SH3) binding motifs of approximately 10 amino acids with essential prolines at the amino acid number seven and ten positions within the motif (Mayer, 2001), two Src homology 2 (SH2) binding motifs containing a phosphotyrosine residue amino terminal to 1 or 2 negatively charged amino acids followed by a hydrophobic amino acid (Songyang et al., 1993), two pleckstrin homology (PH) domains, a substrate domain (SD) rich in serine and threonine residues that is a substrate for serine/threonine kinases, a helical leucine zipper (Lzip) with a heptad repeat of leucine residues (Kouzarides and Ziff, 1988) for intra- and inter-molecular interactions within itself and other AFAP1 molecules and an actin binding domain (ABD) which is both necessary and sufficient for AFAP1 to interact with actin filaments (Baisden et al., 2001a; Qian et al., 1998; Qian et al., 2000; Qian et al., 2004). AFAP1 is known to bind actin filaments through its carboxy terminal actin filament-binding domain and will multimerize through its leucine zipper, thus enabling it to crosslink actin filaments (Qian et al., 2002). Activation by the serine/threonine kinase PKCα directs AFAP1 to colocalize with cSrc in the perinuclear region of the cell and activate cSrc by binding to the SH3 domain via its N-terminal SH3 binding motif (Gatesman et al., 2004; Walker et al., 2007).

Subsequently, cSrc then has the ability to activate downstream cellular signals that affect cell adhesion, invasion and motility (Fincham et al., 1996; Frame and Brunton, 2002).

Upon cSrc binding and activation, AFAP1 and cSrc move to podosomes, adhesion structures found on the ventral membrane of cells which contain an F-actin rich core (Gatesman et al., 2004; Linder and Kopp, 2005). Podosomes also secrete proteases which enable cells to degrade the extracellular matrix, cross tissue barriers and invade. As cSrc activation is known to switch cells from a normal to an invasive phenotype, the activation and movement of cSrc and AFAP1 may be important steps toward promoting cellular changes in adhesion and invasive potential. Also involved in this process is the ability of a cell to assemble, disassemble and remodel its actin cytoskeleton (Yilmaz and Christofori, 2009). As AFAP1 has been shown to affect dynamic changes in actin filament cross-linking (Qian et al., 2004), AFAP1 may also play a r ole in actin cytoskeleton remodeling in addition to cSrc activation.

AFAP1 represents a f amily of three proteins that also include Actin Filament Associated Protein 1 Like 1 (AFAP1L1) and Actin Filament Associated Protein 1 Like 2 (AFAP1L2/XB130) so named as a family by t he Human Genome Project due to similarity in modular domain structure and amino acid sequence, most notably within their PH domains. There are 250 know n PH domain-containing proteins in the human genome and while the amino acid sequences of these PH domains are not well conserved, they are predicted to have a similar structure with a β-barrel with four β strands on one side and three β strands on the other connected by three variable loops connecting β1/β2, β3/β4 and β6/β7 (DiNitto and Lambright, 2006). However, within the AFAP family, the amino acid sequence, as well as st ructure and placement within the protein, are well

conserved between the PH domains, designating them as a related family of proteins. Overall, AFAP1L2/XB130 consists of 818 amino acids and is 35% identical (64% similar) to AFAP1. AFAP1L2/XB130 was also discovered as a cSrc binding protein and contains a shared functional domain structure to AFAP1 including one SH3 motif, three SH2 motifs, 2 PH domains, a coiled-coil region corresponding to the AFAP1 leucine zipper and a sequence similar to the AFAP1 actin binding domain (Xu et al., 2007). However, unlike AFAP1, AFAP1L2/XB130 did not appear to bind efficiently to actin filaments. AFAP1L2/XB130 does appear to have functions distinct from AFAP1 in that it acts as an intermediary between the RET/PTC kinase and PI-3kinase pathway in the thyroid (Lodyga et al., 2009).

In this report we characterize the third AFAP family member, AFAP1L1, which also contains a shared domain structure with AFAP1 and AFAP1L2/XB130. Cladistic analysis of the three AFAP family members (data not shown) indicated that AFAP1 and AFAP1L1 are more closely related to each other than to AFAP1L2/XB130, therefore we sought to compare and contrast AFAP1L1 to AFAP1 so as to characterize AFAP1L1 and determine whether it has functions that are shared or distinct relative to AFAP1.

MATERIALS AND METHODS

Cell Culture and Reagents

A7r5, rat aortic smooth muscle cells, were purchased from American Type Culture Collection and grown in Dulbecco's modified Eagle's medium (DMEM, Mediatech) supplemented with 10% fetal bovine serum (FBS, Atlanta Biologicals), 2mM L-Glutamine (Gibco) and penicillin/streptomycin (Mediatech) at 37°C with 5% CO_2. MDA-MB-435, MDA-MB-231, B1A (Dorfleutner et al., 2007), and Cos-1 cells were grown in Dulbecco's modified Eagle's medium (DMEM, Mediatech), 10% fetal bovine serum (FBS, Atlanta Biologicals), 2mM L-Glutamine (Gibco) and penicillin/streptomycin (Mediatech) at 37°C with 5% CO_2. MCF-10A cells were grown in complete Mammary Epithelium Basal Medium supplemented with MEGM SingleQuots (Lonza) and MCF-7 cells were cultured in Modified Eagle's medium containing 10% FBS and 10 µg/ml bovine insulin (Sigma).

Antibodies used for western blotting and immunofluorescence are as follows: anti-AFAP1 (F1) (Flynn et al., 1993), anti-AFAP1 (BD Transduction Labs), anti-cortactin (Millipore), and β̃actin (Sigma). Tetramethylrodamine-isothiocyanate (TRITC)-conjugated phalloidin was from Sigma and Alexa Fluor conjugated phalloidin and secondary antibodies were from Invitrogen. The antibody against AFAP1L1, 1L1-CT, was generated by ProSci Incorporated. Rabbits were immunized with AFAP1L1 peptide corresponding to amino acids 714-727. 1L1-CT antibody was affinity-purified by affinity absorption against the AFAP1L1 peptide attached to an Ultralink Iodoacetyl column (Pierce) according to manufacturer's protocol. Two additional antibodies against

AFAP1L1, Ab1 and Ab2, targeting either the C-terminus (amino acid 525- 659) or the N-terminus (amino acids 21-159) of 1L1, respectively, were from Sigma.

Constructs

The *afap1l1* cDNA sequence was purchased in two vectors from OpenBioSource. The coding sequence for AFAP1L1 amino acids 1 t hrough 340 w as identified in a pCMV-SPORT6 vector. The coding sequence for AFAP1L1 amino acids 273 t hrough 768 was identified in a pINCY vector. A BstYI restriction site, GATCC, in the overlap region was mutated to a BglII restriction site, GATCT, to create a unique restriction site using the Stratagene QuikChange Site-Directed Mutagenesis Kit according to manufacturer's protocol. AFAP1L1 N-terminal coding sequence was subcloned into pBluescript II K S (Stratagene) using HindIII and an engineered EcoRI restriction site. AFAP1L1 C-terminal coding sequence was subcloned into pBluescript II KS using engineered HindIII and EcoRI restriction sites. F ull length *afa1l1* was created by restriction digest of pBluescript containing each *afap1l1* coding sequence with the unique BglII site in the overlap region and a unique Sca1 site found in the vector followed by fusion of the two halves of pBluescript. The *afap1l1* full length sequence was confirmed by DNA sequencing.

Full length *afap1l1* was subcloned into pEGFP (Clonetech) using HindIII and EcoRI. Full length *afap1l1* was subcloned from pEGFP into pcDNA3.1(+) hygro (Invitrogen) using HindIII and Kpn1. GFP-AFAP1 was previously described by (Qian et al., 2000).

Transfection

For antibody characterization and GST pull down overexpression studies respectively, Cos-1 and 293T cells were transiently transfected with 5μg of either GFP-AFAP1 or GFP-AFAP1L1 using Lipofectamine and Plus reagent according to manufacturer's protocol. For confocal overexpression studies, mouse embryo fibroblasts (MEF) were transfected with 5μg of either GFP-AFAP1L1 or untagged AFAP1L1 (in pcDNA3.1) using Lipofectamine and Plus reagent.

To determine endogenous AFAP1L1 localization to invadopodia, MDA-MB-435 cells were transfected with cSrc527F plasmid using Lipofectamine and Plus reagent (Invitrogen) according to the manufacturer's instructions. A7r5 cells were transfected with GFP-AFAP1 or GFP-AFAP1L1 in increasing amounts from 0.1 μg to 1.0 μg using Lipofectamine and Plus reagent per well of a 6 well plate. Total DNA concentration for dose response transfections was kept constant using an empty pcDNA3.1 vector to keep the total DNA transfected at 1.0 μg to maintain equal transfection efficiency.

Immmunoblotting

Cos-1 cells transiently expressing GFP-AFAP1 or GFP-AFAP1L1 were lysed in 2X SDS buffer (125mM Tris-HCl pH6.8, 20% glycerol, 4% SDS). Cell lines MCF-10A, MCF-7, MDA-MB-231, MDA-MB-435, B1A, and Cos-1 were lysed in 2X SDS buffer. Protein concentration was determined using a BCA Protein Assay Kit (Pierce) according to manufacturer's protocol. 50μg of total lysate was resolved by 8 % SDS-PAGE. Proteins were transferred to polyvinylidene fluoride (PVDF) membrane (Immobilon-P, Millipore) using semi-dry electroblotting. Proteins were detected by incubation with

either anti-1L1-CT (ProSci) 1:250, anti-1L1-Ab1 (Sigma) 1:1000, anti-1L1-Ab2 (Sigma) 1:500, anti-AFAP1 (BD Transduction Labs) 1:10000, anti-AFAP1 (F1) 1:20000, anti-GFP (Zymed) 1:1000 or anti-β-actin (Sigma) 1:10000 in 5% powdered milk (TBS, 0.05% Tween-20) followed by incubation with 1:3000 dilution of donkey anti-mouse or donkey anti-rabbit horseradish peroxidase conjugated antibodies (GE Healthcare Bio-Sciences). Chemiluminescence was visualized with Pierce ECL Western Blotting Substrate.

Immunofluorescence

MDA-MB-435, MEF and A7r5 cells were grown on fibronectin-coated coverslips (50μg/ml) overnight at 37°C. For GFP-AFAP1L1 overexpression and overexpressed untagged AFAP1L1, MEF and A7r5 cells were fixed with 3.7% formaldehyde, permeabilized with 0.2% triton X-100 and incubated with or without anti-cortactin antibody (1:200). Actin filaments were decorated with TRITC-phalloidin (1:600) or Alexa Flour conjugated phalloidin (1:100). For endogenous AFAP1L1 staining, MDA-MB-435 were fixed and permeabilized in 3.7% formaldehyde containing 0.1% triton X-100, incubated with Alexa Flour conjugated phalloidin (1:100) in 3% BSA, and then stained with anti-AFAP1L1 antibodies Ab1 or Ab2 (Sigma, 2 μg/ml) and anti-cortactin antibodies. Fluorescence conjugated secondary antibodies (Invitrogen) were used at a 1:200 dilution. Coverslips were mounted on slides using Prolong Gold (Invitrogen) and imaged with a Zeiss LSM 150 microscope or Nikon eclipse Ti inverted epi-fluorescence microscope. Podosomes were confirmed by anti-cortactin localization along the basolateral membrane via confocal microscopy (0.5 μm scanning thickness). Images

were further processed with Zeiss LSM Image Browser (Zeiss), CorelDRAW12, Nikon NIS element software (Nikon) and Adobe Photoshop.

Immunoprecipitation

Cos-1 cells were lysed in RIPA buffer (50mM Tris pH7.5, 150mM NaCl, 2mM EDTA, 1% Igepal, 0.25% sodium deoxycholate, 10mM β-glycerol, 1mM sodium vanadate, 5μg/μl aprotinin, 5μg/μl leupeptin, 1mM PMSF). 1 mg of Cos-1 lysate was used for each control IgG, anti-1L1-CT and AFAP1 (F1) pull down. Lysates were pre-cleared with agarose A beads for one hour. The precleared lysates were incubated on ice for one hour with 2μg IgG, 1L1-CT or F1 antibodies, and then 50% slurry agarose A beads were added and incubated for one hour. Proteins were eluted from beads using 2X Laemmli sample buffer (LSB) with 200mM dithiothreitol (DTT), separated by 8% SDS-PAGE and processed for western blot analysis. 293T cells were transfected with human cortactin cDNA with either EGFP AFAP1 or AFAP1L1 construct using Polyfect (Qiagen) according to the manufacturer's specification. Forty-eight hours post transfection, cells were lysed in RIPA buffer and 2 mg of cell lysates were used for immunoprecipitation with either GFP antibodies (Polyclonal Av antibodies, Clonetech) or cortactin antibodies (clone 4F11, Millipore) as described above.

Immunohistochemical methods

Immunohistochemistry was performed on de-identified human brain, breast and colon tissue slices obtained from the West Virginia University Department of Pathology Tissue Bank. Tissues were paraffin-embedded and cut into 5 μm-thick sections and

mounted on positive-charge coated slides. Tissue sections were dried overnight in a 45ºC oven, deparaffinized, rehydrated, and subjected to heat-induced epitope retrieval for two hours in 1 mM citrate buffer (pH 6.00) in an 80º C water bath. Endogenous peroxide activity was blocked with 3% hydrogen peroxide and was followed by treatment with a serum-free protein blocker to block non-specific binding. Following each step of the immunoreaction except the protein blocker, sections were rinsed in Tris-buffered saline with Tween-20. Tissues were incubated for two hours with either anti-AFAP1 antibody (F1) at a dilution of 1:50 or with anti-1L1-CT diluted 1:50 in 10% normal horse serum. Negative controls (i.e., preimmune serum or normal rabbit IgG) were incubated in 10% normal horse serum. Sections were incubated with biotinylated-linked secondary antibody, followed by t reatment with streptavidin peroxidase and stained with diaminobenzidine substrate-chromogen solution. Counterstaining was performed with hematoxylin, followed by a water rinse and bluing solution. Tissues were then dehydrated and mounted with coverslips.

Affinity precipitation with GST fusion proteins

293T cells transiently expressing GFP-AFAP1 or GFP-AFAP1L1 were lysed in Modified RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 2mM EDTA, 1% Igepal, 0.25% sodium deoxycholate, 10 m M β-glycerol, 1 m M sodium vanadate, 5µg/µl aprotinin, 5µg/µl leupeptin and 1µM PMSF). GFP-AFAP1 and GFP-AFAP1L1 lysates were pre-cleared using 50 µg of GST protein bound to Glutathione Sepharose 4B coated beads (GE Healthcare) and 1mg of each pre-cleared lysate was incubated at 4°C for 1 hour with either 50 µg control GST, 50 µg GST-Src SH3 domain or 50 µg GST-cortactin

SH3 domain bound t o GST beads. Proteins were eluted from beads by boi ling for 5 minutes in 2X Laemmli sample buffer with 200mM DTT and further processed with PAGE and western blot analysis. Western blot results were examined with ImageJ densitometry analysis so as to provide a ratio between the amount of AFAP1 and AFAP1L1 pulled down with each GST or GST fusion protein.

RESULTS

Identification of AFAP1L1

AFAP1L1 was discovered during a homology search of the human genome using sequences of the AFAP1 PH domains and was subsequently referred to as the third member of the AFAP family by the Human Genome Project. The *afap1l1* gene is located on chromosome 5q33.1 and consists of 19 exons that are predicted to encode 768 amino acids in the open reading frame (Figure 1). All AFAP family members contain at least one predicted N-terminal SH3 binding motif, at least one predicted N-terminal SH2 binding motif, two PH domains separated by a region rich in serine and threonine residues and a C-terminal SH2 binding motif (Figure 2A). One major difference is found in the C-terminus where AFAP1 and AFAP1L1 contain a leucine zipper and actin binding domain while AFAP1L2/XB130 contains a related coiled-coil in this respective region (Figure 2B) (Baisden et al., 2001b; Xu et al., 2007).

There appear to be differences in the SH3 binding motif of AFAP1L1 compared to AFAP1. AFAP1 contains two juxtaposed SH3 binding motifs, PPQMPLPEIP and PPDSGPPPLP, beginning at amino acids 65 and 76 respectively (Guappone and Flynn, 1997). AFAP1 will bind cSrc using the N-terminal P^{71}EIP of the first SH3 binding motif (Guappone and Flynn, 1997), while AFAP1L2/XB130 has been reported to also bind cSrc through its SH3 binding motif, identified as PDLPPPKMIP (Xu et al., 2007). However, the AFAP1L1 coding sequence contains only one predicted SH3 binding motif in its amino terminus, DLPPPLPNKP, which is not consistent with a consensus cSrc SH3 binding motif but may have the ability to bind to other SH3 domains.

AFAP1 contains two SH2 binding motifs that interact with cSrc, YYEEA in the N-terminal portion of the protein and YDYI in the C-terminal portion, which are both sites for tyrosine phosphorylation (Guappone et al., 1998). AFAP1L1 shares 100% sequence identity with both AFAP1 and AFAP1L2/XB130 in its predicted N-terminal SH2 binding motif, YYEEA, and is highly similar in its C-terminal SH2 binding motif YDYV (Guappone et al., 1998; Xu et al., 2007). Thus, these may be potential sites for tyrosine phosphorylation.

The N-terminal and C-terminal PH domains of the AFAP family share sequence similarity and, thus, can be predicted to share structural similarity. The amino acid sequence intervening between the PH domains of AFAP1 (substrate domain, SD, Figure 2B) is rich in serine and threonine residues (Qian et al., 2002). Serine[277] in AFAP1 is a known target of phosphorylation by PKCα and plays a role in the ability of AFAP1 to regulate podosome formation and lifespan (Dorfleutner et al., 2008). A FAP1L1 and AFAP1L2/XB130 also have multiple serine and threonine residues in the sequence flanked by their PH domains, although these sequences have not been confirmed to be phosphorylated by serine/threonine protein kinases.

The leucine zipper of AFAP1 is essential for intra-molecular regulation of AFAP1 and interaction with other AFAP1 molecules and is directly adjacent to the actin binding domain (Qian et al., 2004). AFAP1L1 contains a similar sequence that may have the ability to act as a leucine zipper or coiled-coil motif as well as a putative actin binding domain which is similar to the known actin filament binding domain found in AFAP1 (Qian et al., 2000). AFAP1L2/XB130 contains a coiled-coil motif in its C-terminus and a sequence that is similar to the AFAP1 actin binding domain, but AFAP1L2/XB130 has

not been demonstrated to bind actin filaments (Xu et al., 2007). An amino acid sequence alignment of the three family members found that AFAP1 and AFAP1L1 overall share 44% identity and 71% similarity while AFAP1L1 and AFAP1L2/XB130 share 31% identity and 61% similarity and their sequences predict conserved functional domains within the proteins (Figure 2B).

A novel antibody is specific for AFAP1L1

A rabbit polyclonal antibody was generated against the carboxy terminal amino acids 714-727 of human AFAP1L1. We refer to this antibody as 1L1-CT. We also utilized two commercially available AFAP1L1 antibodies, called Ab1 (epitope defined as amino acids 525-659) and Ab2 (epitope defined as amino acids 21-159). The ability of these antibodies to detect AFAP1L1 was studied using lysates of human breast lines MCF-10A, MCF-7, MDA-MB-231, B1A (a variant of MDA-MB-231 cells that had AFAP1 knocked down using shRNA (Dorfleutner et al., 2007)), MDA-MB-435 as well as monkey kidney cell line Cos-1. All three AFAP1L1 antibodies detected a distinct protein band by western blot analysis with a M_r of 115 kDa (Figure 3A, result from Ab2 as a r epresentative example) with MDA-MB-435 having the highest endogenous expression levels of AFAP1L1 among the tested cell lines. The M_r of AFAP1L1 was further confirmed by s iRNA knockdown against AFAP1L1 (data not shown). The expression pattern of AFAP1L1 across breast cancer cell lines was further confirmed in messenger RNA levels through quantitative RT-PCR analysis (data not shown). To test the ability of the 1L1-CT antibody to immunoprecipitate endogenous AFAP1L1, Cos-1 cells were lysed and 1mg of lysate (1µg/µl) was incubated with either 5µg 1L1-CT, 5µg

AFAP1 antibody F1 or 5μg control rabbit IgG and the immunoprecipitate resolved by

SDS-PAGE (Figure 3B).   Upon western blot analysis with 1L1-CT, it was apparent

that the 1L1-CT antibody specifically immunoprecipitated a protein with a M_r of 115kDa,

consistent with the predicted M_r for AFAP1L1 (Figure 3B, left panel, middle lane); while

the AFAP1 antibody F1 did not immunoprecipitate a protein of this M_r (Figure 3B, left

panel, right lane) that could be detected by the 1L1-CT antibody. To test the specificity

of F1 antibody for endogenous AFAP1, Cos-1 cells were lysed and 1.5mg (1μg/μl) of

lysate was incubated with rabbit IgG, F1 or 1L1-CT. F1 antibody specifically

immunoprecipitated the expected AFAP1 protein with an M_r of 110 kDa (Figure 3B,

right panel, right lane) and did not immunoprecipitate AFAP1L1 (Figure 3B, right panel,

middle lane), indicating that these two antibodies are specific for endogenous protein. To

determine if 1L1-CT and F1 antibodies can specifically recognize overexpressed protein,

plasmids encoding GFP-AFAP1L1 and GFP-AFAP1 were transfected into Cos-1 cells.

Western blot analysis of these cells (Figure 3C) showed that 1L1-CT antibody detected

overexpressed GFP-AFAP1L1 but not overexpressed GFP-AFAP1 (Figure 3C, left

panel). AFAP1 antibody F1 detected GFP-AFAP1 but not GFP-AFAP1L1 (Figure 3C,

middle panel). These data indicate that the AFAP1L1 antibody 1L1-CT and AFAP1

antibody F1 have specificity for AFAP1L1 and AFAP1, respectively. These data also

indicated that 1L1-CT antibody is specific to the AFAP1L1 of human and monkey

species. 1L1-CT antibody did not recognize any distinct band in murine cell lysates (data

not shown), because the epitope is not conserved in murine AFAP1L1 (Figure 3D).

Sigma antibodies Ab1 and Ab2 also failed to recognize AFAP1L1 in murine cell lysates.

An alignment of the peptide sequence used to create 1L1-CT against the amino acid

sequence of human, chimpanzee, mouse and rat AFAP1L1 was shown to illustrate the similarity of the 1L1-CT binding site across species (Figure 3D).

Immunohistochemical analysis of AFAP1L1

The ability of the 1L1-CT antibody to specifically recognize endogenous AFAP1L1 in its native conformation by immunoprecipitation and to recognize AFAP1L1 by immunofluorescence (Supplemental Figure 1) demonstrated its usefulness for immunohistochemistry. To determine the tissue localization of AFAP1L1, immunohistochemistry was performed on human breast, colon and brain tissues using 1L1-CT antibody (Figure 4). AFAP1 or AFAP1L1 tissue localizations were compared using sequential sections of tissue immunolabeled with F1 or 1L1-CT antibodies, respectively, in order to contrast expression patterns. In the breast, AFAP1 was found to strongly associate with the ductal cells and also the breast microvasculature (Figure 4A panels a, c and e). AFAP1L1 was found in the contractile myoepithelial cell layer which surrounds the breast ducts and in the microvasculature (Figure 4A panels b, d and f), similar to AFAP1. In the colon (Figure 4B), AFAP1 was found in the epithelial mucous membrane as well as in the colonic crypts (Figure 4B panels a and g) which aid in mucous production and production of new epithelial cells for the intestinal surface (Sherwood, 2003). AFAP1L1 was also found in the mucous membrane and colonic crypts (Figure 4B panels b and h). Both AFAP1 and AFAP1L1 are found in the microvasculature (Figure 4B panels a-d) while AFAP1 can also be found in nerves that pass through the lamina propria (Figure 4B panel g). The lamina propria lies directly beneath the epithelial cell layer and consists of connective tissue, blood vessels, nerve

fibers and lymphatic ducts (Widmaier, 2003). The muscularis, layers of smooth muscle which provide movement of the colon, expresses both AFAP1 and AFAP1L1 (Figure 4B panels c-f) while AFAP1 is solely found in Auerbach's plexus (Figure 4B panel e), the nerves that innervate the muscle of the gut (Sherwood, 2003). While AFAP1 and AFAP1L1 have overlapping localization in human breast and colon tissue, human brain showed examples of differential expression patterns for these two proteins (Figure 4C). The cerebellum is located at the base of the skull and is an important part of the motor system. Divided into three layers, the cerebellar cortex is composed of a granular cell layer consisting of small granule cells which receive input from mossy fibers and extend into the molecular layer, the outermost portion of the cerebellar cortex which houses stellate and basket cells. Purkinje cells are large neurons acting as the sole output of motor coordination from the cerebellar cortex and are found in a single layer between the molecular and granular layers (Squire, 2008). AFAP1 is found in the microvasculature of the brain, in the molecular layer and meningeal vessels and to a slight level around the granule cells of the granular layer (Figure 4C panels a, c, e and g). AFAP1L1 is again found in low levels in the microvasculature of the brain but is localized around the Purkinje neurons and the granule cells of the granular layer (Figure 4C panels b, d, f and h). This immunolabeling is not inside of either the Purkinje neurons or the granule cell bodies but instead extends away from the cell body. Although the source of this immunolabeling is unknown, mossy fibers and climbing fibers extend into the granule cell layer and Purkinje cell layer respectively, synapsing onto granule and Purkinje cells (Squire, 2008). This may provide the source of immunolabeling for AFAP1L1 away from the granule and Purkinje cell body. Outside of the cerebellar cortex, AFAP1 and

AFAP1L1 are both found in glial cells (Figure 4C panels o-p) but have a differential expression pattern in the dentate nucleus (Figure 4C k-n), one of four deep cerebellar nuclei responsible for voluntary movements of the extremities (Squire, 2008). Again, AFAP1L1 is highly expressed away from the cell bodies within the dentate nucleus, while AFAP1 was not detected in this structure. Relative intensities of immunohistochemical signal for AFAP1 and AFAP1L1 are presented in Table 1.

AFAP1L1 subcellular localization

AFAP1 is known to bind actin filaments, move to actin rich structures called podosomes upon stimulation with the phorbol ester PDBu and also to be localized in podosome-like structures termed invadopodia in MDA-MB-231 cells, while AFAP1L2/XB130 is not known to bind actin filaments (Flynn et al., 1993; Gatesman et al., 2004; Xu et al., 2007). Thus, we sought to determine if AFAP1L1 had the potential to be localized with actin filaments and podosomes, similar to AFAP1 (Gatesman et al., 2004). For this analysis, we used the rat-derived A7r5 cell line, a well established cell model system for podosome formation (Hai et al., 2002). Overexpression of GFP-AFAP1L1 alone in A7r5 cells produced two distinct phenotypes (Figure 5A). GFP-AFAP1L1 expressing cells either have GFP-AFAP1L1 decorating actin stress fibers while cortactin is found in the cytoplasm and around the cell periphery (Figure 5A panels a-d) or both GFP-AFAP1L1 and cortactin colocalizing in punctate actin dots along the ventral membrane (Figure 5A and Supplemental Figure 2). The adaptor protein cortactin was used as a marker for podosome formation. Podosome formation was observed either around the cell periphery (5A panels e-h) or scattered across the ventral membrane (5A

panels i-l). Thus, AFAP1L1 decorates actin stress fibers and co-localizes with cortactin in podosomes. Overexpression of AFAP1L1 induced the formation of podosomes in a small percentage of cells.

Next, we utilized MDA-MB-435 cells to determine the subcellular localization of endogenous AFAP1L1 based on the high level of expression of AFAP1L1 in this cell line. Immunolabeling of MDA-MB-435 with Sigma Ab2 (Figure 5B, panel a-c) and Ab1 (data not shown) showed a population of AFAP1L1 decorating actin stress filaments while the rest of AFAP1L1 was detected diffusely across the cytoplasm. A similar pattern of expression was observed for the 1L1-CT antibody in MEF cells overexpressing untagged AFAP1L1 (Supplemental Figure 1).

To determine whether endogenous AFAP1L1 could co-localize to invadopodia, we transiently transfected a constitutively active Src construct (Src527F) into MDA-MB-435 to induce invadopodia, podosome-like structures found in cancer cells. The overexpression of Src527F induced punctate actin- and cortactin-rich structures in these cells (Figure 5C, b-c, f-g) accompanied by the loss of stress filaments, which were consistent with the appearance of typical invadopodia. The immunolabeling of endogenous AFAP1L1 (Figure 5C, a) showed discernable association with invadopodia, co-localizing with cortactin and actin (Figure 5C, a-d, marked with white arrow as an example of colocalization). These data indicate that AFAP1L1, like AFAP1, is associated with actin stress filaments and localizes to invadosomes (the collective term for podosomes and invadopodia) at both overexpressed and endogenous levels (also see Supplemental Figure 3) (Linder, 2009; Saltel et al., 2010).

Quantification of podosome formation in A7r5 expressing AFAP1 or AFAP1L1

In a previous study of AFAP1, expression of GFP-AFAP1 was sufficient to induce podosomes in a portion of the A7r5 cells in which it was expressed (Dorfleutner et al., 2008). Therefore, we sought to compare the potency of the GFP-AFAP1 and GFP-AFAP1L1 constructs to induce podosomes using a dose response analysis. A7r5 cells were transfected with equal amounts of total plasmid cDNA containing the indicated amount of GFP-AFAP1 or GFP-AFAP1L1 plasmid (Figure 6). Transfection of equal amounts of plasmid (combination of empty vector and GFP-AFAP) ensured that transfection efficiencies were identical for each dose of plasmid. Forty eight hours after transfection, the cells expressing either GFP-AFAP1 or GFP-AFAP1L1 were assessed for podosome formation. Reduction of plasmid DNA resulted in a corresponding reduction of GFP-AFAP1 and GFP-AFAP1L1 expression (Figure 6A). The reduction of expression was also reflected in a reduction in the percent of cells exhibiting podosomes (Figure 6B). When 1μg and 0.35 μg of plasmid DNA was used, the percentage of GFP-AFAP1 expressing cells exhibiting podosomes dropped from 11.4% to 7.1% while in GFPAFAP1L1 expressers the change was more modest, 9.2% to 8.0%. At 0.1 μg of DNA the number of podosome cells appeared to increase. This increase resulted from the variability in expression levels among cells and the dimness of the low expressers which made these cells difficult to detect by immunofluorescence. Overall, this experiment demonstrated that under identical conditions of transfection, expression, processing and immunofluorescence analysis, expression of GFP-AFAP1 and GFP-AFAP1L1 induce podosome formation in A7r5 cells.

Interaction of AFAP1L1 and cortactin

To determine if AFAP1L1 might have a differential ability to interact with proteins relative to AFAP1, we screened SH3 domains from various proteins for the binding to AFAP1L1. A Panomics array screening of SH3 domains showed that the SH3 domain of cortactin and the SH3 binding motif of AFAP1L1 had a strong potential to interact (data not shown). To confirm these screening results, GST fusion proteins containing the cortactin SH3 domain and the Src SH3 domain were incubated with 1mg of cell lysate containing either overexpressed GFP-AFAP1L1 or GFP-AFAP1. As expected from the known binding of AFAP1 to Src, AFAP1 showed a strong interaction with the Src SH3 domain with a 14-fold increase in the ratio of AFAP1 pulled down by GST-Src SH3 as compared to GST control while precipitation by the cortactin SH3 domain was undetectable (Figure 7A). However, AFAP1L1 was more efficiently precipitated by GST-SH3-cortactin than GST-SH3-Src with an approximately 9-fold increase in the ratio of AFAP1L1 pulled down by GST-cortactin SH3 compared to GST control (Figure 7B). These data indicate that the SH3 binding motifs of AFAP1 and AFAP1L1 have the ability to interact with different affinities toward different SH3 binding partners. To further examine the potential binding of AFAP1L1 with cortactin, we transiently overexpressed either GFP-AFAP1 or GFP-AFAP1L1 with cortactin in 293T cells. G FP-AFAP1L1 co-immunoprecipitated with cortactin (Figure 7C, upper panel) more efficiently than did GFP-AFAP1 (the amounts of immunoprecipitated GFP proteins were relatively equal Figure 7C, lower panel). Conversely, cortactin specifically co-immunoprecipitated GFP-AFAP1L1, but not GFP-AFAP1 (Figure 7D, upper panel). These data suggest that cortactin can form a complex with AFAP1L1, but is either unable

to form a complex with AFAP1 or its ability to bind in a complex with AFAP1 is below detection limits. This interaction is hypothesized to occur through SH3 interactions based on our affinity precipitations with the SH3 domains and suggests a unique function for AFAP1L1 that is distinct from AFAP1.

DISCUSSION

AFAP family amino acid sequence

AFAP1 is a well-studied cSrc binding partner and actin filament cross-linking protein that is the prototype member of what is predicted to be a family of three proteins which also includes AFAP1L1 and AFAP1L2/XB130. There are no published reports on AFAP1L1 and the goal of this study was to analyze some of the similarities and differences between AFAP1 and AFAP1L1 in order to understand what functions AFAP1L1 may have relative to AFAP1. The AFAP family consists of three family members that share similarity in protein binding motifs including both PH domains. Analysis of the overall amino acid structure of previously described AFAP1 and newly presented AFAP1L1 show 44% identity and 71% similarity between the two proteins. AFAP1 contains two juxtaposed SH3 binding motifs in its N-terminus, PPQMPLPEIP and PPDSGPPPLP, of which the first is necessary for efficient cSrc binding. Mutation of a necessary proline, Pro[71], to an alanine, Ala[71], in the AFAP1 SH3 binding motif prevented cSrc binding (Guappone and Flynn, 1997). While AFAP1L1 only contains one putative SH3 binding motif, DLPPPLPNKP, this sequence is not consistent with the conserved consensus cSrc SH3 binding motif. Rather, the AFAP1L1 SH3 binding motif more closely resembles that of a cortactin SH3 domain binding site, which preferentially binds a +PPΨPXKPXWL motif where + is a basic residue, Ψ is an aliphatic residue, and X is any amino acid (Sparks et al., 1996).

In addition to the SH3 binding motif, cSrc also has the ability to phosphorylate and bind two SH2 binding motifs in AFAP1. In AFAP1, an N-terminal YYEEA sequence adjacent to the SH3 binding motifs and a C-terminal YDYI sequence have been shown to

be phosphorylated by cSrc (Guappone et al., 1998). AFAP1L1 contains the conserved YYEEA sequence in its N-terminus adjacent to the SH3 binding motif and a similar YDYV sequence in its C-terminus. Although we did not analyze AFAP1L1 as a possible target for phosphorylation by tyrosine kinases, these similar SH2 binding motifs may be potential targets for tyrosine phosphorylation. A ScanSite motif scan comparison between AFAP1 and AFAP1L1 predicts these sites for tyrosine phosphorylation and also predicts similar sites for serine/threonine phosphorylation as well (data not shown). The amino acid sequences between the PH1 and PH2 domains of AFAP1 and AFAP1L1 are rich in serine and threonine residues that are predicted to be sites of serine and threonine phosphorylation. In particular, AFAP1 is known to be phosphorylated on serine 277 by PKCα and this phosphorylation is important for podosome turnover (Dorfleutner et al., 2008). Although we have not validated AFAP1L1 as a PKC substrate, AFAP1L1 also contains a similar PKCα phosphorylation site.

The conservation of domain structure, overall sequence similarity and high similarity of sequence between the two PH domains indicates that AFAP1, AFAP1L1 and AFAP1L2/XB130 are members of a family of proteins, referred to here and in genome databases as the AFAP family. The amino terminal PH1 domain of AFAP1 has been known to function in intra-molecular regulation of AFAP1 (Qian et al., 2004). In addition, the PH1 domain is a binding partner for PKCα and phospholipids (Cunnick and Flynn, manuscript in preparation) (Qian et al., 2002).

The actin binding domain within AFAP1 is necessary and sufficient for the binding of AFAP1 to actin filaments (Qian et al., 2000). Comparison of the AFAP1 actin binding domain sequence to the corresponding sequence in AFAP1L1 does not indicate a

high level of similarity although our data indicates that AFAP1L1 also has the ability to associate with actin filaments. We predict that the sequences within the C-terminus of AFAP1L1 that are similar to the known actin filament binding domain in AFAP1 may promote actin filament association, however, as the binding affinity of AFAP1 for actin is relatively low, we predict that the putative actin binding domain of AFAP1L1 may have a lower affinity for actin. S tudies are underway to determine the mechanism of actin filament binding and if the leucine zipper/coiled-coil motif adjacent to the putative actin filament binding domain can regulate self-association and actin filament cross-linking, similar to AFAP1 (Qian et al., 2004).

AFAP1L1 western blot analysis

AFAP1L1 can be found in human breast and breast cancer cell lines, as well as the Cos-1 monkey cell line. The predicted molecular weight of AFAP1L1 is 85kDa, but it is detected as a single protein band by w estern blot analysis with a M_r of 115 kD a, Similarly, AFAP1 has a molecular weight of 82 kDa but is detected on western blot as a 110 kDa protein.. F or AFAP1, it was hypothesized that the difference in molecular weight and M_r might be due to overall charge and the shape of the protein, which may also apply to AFAP1L1. T he AFAP1L1 specific antibodies, 1L1-CT, Ab1 and Ab2 antibodies (latter two from Sigma), have the ability to specifically detect AFAP1L1 but not AFAP1 by western blot. In our hands, all three antibodies could detect AFAP1L1 by immunoprecipitation, immunofluorescence and western blot; however, the 1L1-CT antibody was determined to be less efficient in detecting denatured AFAP1L1 by western blot, indicating that this antibody may preferentially recognize a conformational epitope

and be less efficient in detecting denatured AFAP1L1, relative to Ab1 and Ab2, which appear to detect denatured AFAP1L1 efficiently. Conversely, the anti-AFAP1 antibody F1 is able to specifically immunoprecipitate AFAP1 but does not immunoprecipitate AFAP1L1.

AFAP1L1 tissue expression pattern

AFAP1 and AFAP1L1 have similar distribution in breast and colon tissue with the exception that AFAP1 can be found in the nerves that innervate the gut while AFAP1L1 was not detected there. While immunolabeling of the brain shows similarities between AFAP1 and AFAP1L1 in the granule cell layer and glial cells, differences occur in the brain where AFAP1L1 is found in the Purkinje cell layer and the dentate nucleus, while AFAP1 was not detected here. The dentate nucleus is known to be involved in motor coordination; alterations in the neuronal composition of signaling to and from the dentate have been implicated in diseases with motor dysfunction such as Alzheimer's disease, autism spectrum disorders and Pick's disease (Braak et al., 1999; Fukutani et al., 1999; Lotspeich and Ciaranello, 1993). Consistent immunolabeling of AFAP1L1 between the granule cell layer, Purkinje layer and dentate nucleus indicates that the AFAP1L1 immunohistochemical signal is not in the cell bodies of the granule cells, Purkinje neurons or cells of the dentate nucleus, but is instead found extending away from the cell body. Thus, we hypothesize that AFAP1L1 may be playing a role in either the axon terminals of afferent neurons or the dendritic spines of efferent neurons in these areas, both of which have a dynamic actin cytoskeleton and are rich in proteins involved in cytoskeletal remodeling, such as cortactin (Hering and Sheng, 2003). As we see the

most marked difference between AFAP1 and AFAP1L1 expression in the dentate nucleus, we speculate that AFAP1L1 may play a role in the association of proteins involving signaling to or from the dentate nucleus. Notably, podosomes have recently been shown to be involved in maturation of the neuromuscular junction and a similar mechanism may be involved in neuron-neuron interactions (Proszynski et al., 2009).

AFAP1L1 subcellular localization and podosome formation

GFP-AFAP1L1 and endogenous AFAP1L1 colocalized with actin filaments and invadosomes similar to the subcellular localization pattern of AFAP1. Overexpression of GFP-AFAP1L1 induced podosome formation in the absence of extracellular signals (e.g., phorbol ester). A direct comparison of the ability of GFP-AFAP1 and GFP-AFAP1L1 to induce podosomes demonstrated that both proteins had similar capacities to induce these structures. Thus, like AFAP1, AFAP1L1 may have the ability to interact with proteins involved in podosome formation such as f-actin, cSrc or possibly cortactin. A panomics array of various SH3 domains was scanned for the ability to bind the AFAP1L1 SH3 binding motif, and these data indicated that the AFAP1L1 SH3 binding motif and the cortactin SH3 domain showed the strongest interaction among those SH3 domains surveyed (data not shown). GST fusion proteins containing either the cortactin SH3 domain or Src SH3 domain show that AFAP1 interacted with GST-Src SH3 much better than with GST-cortactin SH3. However, with AFAP1L1 the reverse was true. AFAP1L1 interacted with GST-cortactin-SH3 better than GST-Src SH3. These data indicate that the SH3 binding motifs of AFAP1 and AFAP1L1 likely have different affinities for different SH3 binding partners. The co-immunoprecipitation of cortactin with full length

AFAP1L1 and not with AFAP1 further supports the possibility of differential characteristics of these two proteins and also suggests the interaction of AFAP1L1 with cortactin as a possible mechanism of AFAP1L1 localization to invadosomes. Investigations are underway to determine if the SH3 binding motif of AFAP1L1 mediates the binding of AFAP1L1 to cortactin and to determine the functional relevance of this interaction.

Conclusion

The goal of this study was to characterize the AFAP family member AFAP1L1 and to determine if it had functions common and distinct to AFAP1. Our data indicate that AFAP1L1 has strong similarity and conservation of domain structure with AFAP1. Also similar to AFAP1, it has an ability to associate with actin filaments, can be found in actin-rich structures such as invadosomes, and is capable of independently inducing podosomes upon overexpression. AFAP1L1 shares expression patterns with AFAP1 in several tissues. However, AFAP1L1 did display some unique properties. For example, unlike AFAP1, AFAP1L1 was found in the dentate nucleus and its expression appeared to occur along neuronal processes. Interestingly, it is in these processes that podosomes are hypothesized to be associated with synaptic connections (Proszynski et al., 2009), and these same connections in the dentate nucleus contain cortactin binding proteins (Hering and Sheng, 2003). As podosomes play a role in enabling cells to traverse and cross tissue barriers, we hypothesize that AFAP1L1 may play a unique role in the innervation of the dentate nucleus. Although it may have a weak affinity, AFAP1L1 is not predicted to be a strong binding partner for cSrc through SH3 interactions; however, it does interact with

cortactin. As a potential binding partner for actin filaments and cortactin, AFAP1L1 may associate with podosomes via interactions with these proteins and regulate podosome formation in cells, including neurons within the dentate nucleus.

ACKNOWLEDGEMENTS

REFERENCES

Baisden, J.M., Gatesman, A.S., Cherezova, L., Jiang, B.H., and Flynn, D.C. (2001a). The intrinsic ability of AFAP-110 to alter actin filament integrity is linked with its ability to also activate cellular tyrosine kinases. Oncogene *20*, 6607-6616.

Baisden, J.M., Qian, Y., Zot, H.M., and Flynn, D.C. (2001b). The actin filament-associated protein AFAP-110 is an adaptor protein that modulates changes in actin filament integrity. Oncogene *20*, 6435-6447.

Braak, E., Arai, K., and Braak, H. (1999). Cerebellar involvement in Pick's disease: affliction of mossy fibers, monodendritic brush cells, and dentate projection neurons. Exp Neurol *159*, 153-163.

DiNitto, J.P., and Lambright, D.G. (2006). Membrane and juxtamembrane targeting by PH and PTB domains. Biochim Biophys Acta *1761*, 850-867.

Dorfleutner, A., Cho, Y., Vincent, D., Cunnick, J., Lin, H., Weed, S.A., Stehlik, C., and Flynn, D.C. (2008). Phosphorylation of AFAP-110 affects podosome lifespan in A7r5 cells. J Cell Sci *121*, 2394-2405.

Dorfleutner, A., Stehlik, C., Zhang, J., Gallick, G.E., and Flynn, D.C. (2007). AFAP-110 is required for actin stress fiber formation and cell adhesion in MDA-MB-231 breast cancer cells. J Cell Physiol *213*, 740-749.

Fincham, V.J., Unlu, M., Brunton, V.G., Pitts, J.D., Wyke, J.A., and Frame, M.C. (1996). Translocation of Src kinase to the cell periphery is mediated by the actin cytoskeleton under the control of the Rho family of small G proteins. J Cell Biol *135*, 1551-1564.

Flynn, D.C., Leu, T.H., Reynolds, A.B., and Parsons, J.T. (1993). Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. Mol Cell Biol *13*, 7892-7900.

Frame, M.C., and Brunton, V.G. (2002). Advances in Rho-dependent actin regulation and oncogenic transformation. Curr Opin Genet Dev *12*, 36-43.

Fukutani, Y., Cairns, N.J., Everall, I.P., Chadwick, A., Isaki, K., and Lantos, P.L. (1999). Cerebellar dentate nucleus in Alzheimer's disease with myoclonus. Dement Geriatr Cogn Disord *10*, 81-88.

Gatesman, A., Walker, V.G., Baisden, J.M., Weed, S.A., and Flynn, D.C. (2004). Protein kinase Calpha activates c-Src and induces podosome formation via AFAP-110. Mol Cell Biol *24*, 7578-7597.

Guappone, A.C., and Flynn, D.C. (1997). The integrity of the SH3 binding motif of AFAP-110 is required to facilitate tyrosine phosphorylation by, and stable complex formation with, Src. Mol Cell Biochem *175*, 243-252.

Guappone, A.C., Weimer, T., and Flynn, D.C. (1998). Formation of a stable src-AFAP-110 complex through either an amino-terminal or a carboxy-terminal SH2-binding motif. Mol Carcinog *22*, 110-119.

Hai, C.M., Hahne, P., Harrington, E.O., and Gimona, M. (2002). Conventional protein kinase C mediates phorbol-dibutyrate-induced cytoskeletal remodeling in a7r5 smooth muscle cells. Exp Cell Res *280*, 64-74.

Hering, H., and Sheng, M. (2003). Activity-dependent redistribution and essential role of cortactin in dendritic spine morphogenesis. J Neurosci *23*, 11759-11769.

Kouzarides, T., and Ziff, E. (1988). The role of the leucine zipper in the fos-jun interaction. Nature *336*, 646-651.

Linder, S. (2009). Invadosomes at a glance. J Cell Sci *122*, 3009-3013.

Linder, S., and Kopp, P. (2005). Podosomes at a glance. J Cell Sci *118*, 2079-2082.

Lodyga, **M.**, **De Falco**, **V.**, **Bai**, **X.H.**, **Kapus**, **A.**, **Melillo**, **R.M.**, **Santoro**, **M.**, **and Liu**, **M.** (2009). XB130, a tissue-specific adaptor protein that couples the RET/PTC oncogenic kinase to PI 3-kinase pathway. Oncogene *28*, 937-949.

Lotspeich, **L.J.**, **and Ciaranello**, **R.D.** (1993). The neurobiology and genetics of infantile autism. Int Rev Neurobiol *35*, 87-129.

Mayer, **B.J.** (2001). SH3 domains: complexity in moderation. J Cell Sci *114*, 1253-1263.

Proszynski, **T.J.**, **Gingras**, **J.**, **Valdez**, **G.**, **Krzewski**, **K.**, **and Sanes**, **J.R.** (2009). Podosomes are present in a postsynaptic apparatus and participate in its maturation. Proc Natl Acad Sci U S A *106*, 18373-18378.

Qian, **Y.**, **Baisden**, **J.M.**, **Cherezova**, **L.**, **Summy**, **J.M.**, **Guappone-Koay**, **A.**, **Shi**, **X.**, **Mast**, **T.**, **Pustula**, **J.**, **Zot**, **H.G.**, **Mazloum**, **N.**, *et al.* (2002). PC phosphorylation increases the ability of AFAP-110 to cross-link actin filaments. Mol Biol Cell *13*, 2311-2322.

Qian, **Y.**, **Baisden**, **J.M.**, **Westin**, **E.H.**, **Guappone**, **A.C.**, **Koay**, **T.C.**, **and Flynn**, **D.C.** (1998). Src can regulate carboxy terminal interactions with AFAP-110, which influence self-association, cell localization and actin filament integrity. Oncogene *16*, 2185-2195.

Qian, **Y.**, **Baisden**, **J.M.**, **Zot**, **H.G.**, **Van Winkle**, **W.B.**, **and Flynn**, **D.C.** (2000). The carboxy terminus of AFAP-110 modulates direct interactions with actin filaments and regulates its ability to alter actin filament integrity and induce lamellipodia formation. Exp Cell Res *255*, 102-113.

Qian, **Y.**, **Gatesman**, **A.S.**, **Baisden**, **J.M.**, **Zot**, **H.G.**, **Cherezova**, **L.**, **Qazi**, **I.**, **Mazloum**, **N.**, **Lee**, **M.Y.**, **Guappone-Koay**, **A.**, **and Flynn**, **D.C.** (2004). Analysis of the role of the leucine zipper motif in regulating the ability of AFAP-110 to alter actin filament integrity. J Cell Biochem *91*, 602-620.

Saltel, **F.**, **Daubon**, **T.**, **Juin**, **A.**, **Ganuza**, **I.E.**, **Veillat**, **V.**, **and Genot**, **E.** (2010). Invadosomes: Intriguing structures with promise. Eur J Cell Biol.

Sherwood, **L.** (2003). Human Physiology: From Cells to Systems Fifth Edition. *Brooks Cole, Kentucky*.

Songyang, **Z.**, **Shoelson**, **S.E.**, **Chaudhuri**, **M.**, **Gish**, **G.**, **Pawson**, **T.**, **Haser**, **W.G.**, **King**, **F.**, **Roberts**, **T.**, **Ratnofsky**, **S.**, **Lechleider**, **R.J.**, *et al.* (1993). SH2 domains recognize specific phosphopeptide sequences. Cell *72*, 767-778.

Sparks, **A.B.**, **Rider**, **J.E.**, **Hoffman**, **N.G.**, **Fowlkes**, **D.M.**, **Quillam**, **L.A.**, **and Kay**, **B.K.** (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. Proc Natl Acad Sci U S A *93*, 1540-1544.

Squire, **L.B.**, **D.**; **Bloom**, **F.**; **Du Lac**, **S.**; **Ghosh A.**;**Spitzer**, **N.** (2008). Fundamental Neuroscience Third Edition. Elsevier, Inc, Massachusettes.

Walker, **V.G.**, **Ammer**, **A.**, **Cao**, **Z.**, **Clump**, **A.C.**, **Jiang**, **B.H.**, **Kelley**, **L.C.**, **Weed**, **S.A.**, **Zot**, **H.**, **and Flynn**, **D.C.** (2007). PI3K activation is required for PMA-directed activation of cSrc by AFAP-110. Am J Physiol Cell Physiol *293*, C119-132.

Widmaier, **E.R.**, **H.**; **Strang**, **K.** (2003). Vander, Sherman, Luciano's Human Physiology: The Mechanisms of Body Function Ninth Edition. McGraw-Hill, New York.

Xu, **J.**, **Bai**, **X.H.**, **Lodyga**, **M.**, **Han**, **B.**, **Xiao**, **H.**, **Keshavjee**, **S.**, **Hu**, **J.**, **Zhang**, **H.**, **Yang**, **B.B.**, **and Liu**, **M.** (2007). XB130, a novel adaptor protein for signal transduction. J Biol Chem *282*, 16401-16412.

Yilmaz, **M.**, **and Christofori**, **G.** (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer Metastasis Rev *28*, 15-33.

Table 1

AFAP1 and AFAP1L1 immunohistochemical signal in human tissue

	AFAP1	AFAP1L1
Breast Lobule	++++	+
Breast Duct	++++	+
Breast Microvasculature	+++	+
Colon Muscularis	++	+++
Colon Mucosa	+++	+++
Colon Lamina Propria	-	-
Cerebellar Cortex	+	+
Cerebellar Granule Cell Layer	++	++++
Cerebellar Purkinje Layer	+	++++
Meningeal Vessels	++++	+
Dentate Nucleus	-	++++
Glial Cells	++	+++

Table 1: AFAP1 and AFAP1L1 immunohistochemical signal intensity in human tissue

The intensity of AFAP1 staining in human breast, colon and brain tissues was analyzed. Staining in breast ducts was scored as ++++ for highest level of intensity. Staining in the dentate nucleus was scored as – for no level of staining. Other tissues labeled for AFAP1 were compared using this scale. AFAP1L1 staining intensity was analyzed in a similar manner, designating AFAP1L1 staining in the dentate nucleus as ++++ and lamina propria as – for no level of staining. Other tissues labeled for AFAP1L1 were compared using this scale.

Figure 1

```
ATG GAC CGA GGC CAG GTG CTG GAG CAG CTG CTC CCA GAG CTC ACC GGG CTG CTC AGC CTC CTG GAC CAC GAG
 M   D   R   G   Q   V   L   E   Q   L   L   P   E   L   T   G   L   L   S   L   L   D   H   E
TAC CTC AGC GAT ACC ACC CTG GAA AAG AAG ATG GCC GTG GCC TCC ATC CTG CAG AGC CTG CAG CCC CTT CCA
 Y   L   S   D   T   T   L   E   K   K   M   A   V   A   S   I   L   Q   S   L   Q   P   L   P
GCA AAG GAG GTC TCC TAC CTG TAT GTG AAC ACA GCA GAC CTC CAC TCG GGG CCC AGC TTC GTG GAA TCC CTC
 A   K   E   V   S   Y   L   Y   V   N   T   A   D   L   H   S   G   P   S   F   V   E   S   L
TTT GAA GAA TTT GAC TGT GAC CTG AGT GAC CTT CGG GAC ATG CCA GAG GAT GAT GGG GAG CCC AGC AAA GGA
 F   E   E   F   D   C   D   L   S   D   L   R   D   M   P   E   D   D   G   E   P   S   K   G
GCC AGC CCT GAG CTA GCC AAG AGC CCA CGC CTG AGA AAC GCG GCC GAC CTG CCT CCA CCG CTC CCC AAC AAG
 A   S   P   E   L   A   K   S   P   R   L   R   N   A   A   D   L   P   P   P   L   P   N   K
CCT CCC CCT GAG GAC TAC TAT GAA GAG GCC CTT CCT CTG GGA CCC GGC AAG TCG CCT GAG TAC ATC AGC TCC
 P   P   P   E   D   Y   Y   E   E   A   L   P   L   G   P   G   K   S   P   E   Y   I   S   S
CAC AAT GGC TGC AGC CCC TCA CAC TCG ATT GTG GAT GGC TAC TAT GAG GAC GCA GAC AGC AGC TAC CCT GCA
 H   N   G   C   S   P   S   H   S   I   V   D   G   Y   Y   E   D   A   D   S   S   Y   P   A
ACC AGG GTG AAC GGC GAG CTT AAG AGC TCC TAT AAT GAC TCT GAC GCA ATG AGC AGC TCC TAT GAG TCC TAC
 T   R   V   N   G   E   L   K   S   S   Y   N   D   S   D   A   M   S   S   S   Y   E   S   Y
GAT GAA GAG GAG GAG GAA GGG AAG AGC CCG CAG CCC CGA CAC CAG TGG CCC TCA GAG GAG GCC TCC ATG CAC
 D   E   E   E   E   E   G   K   S   P   Q   P   R   H   Q   W   P   S   E   E   A   S   M   H
CTG GTG AGG GAA TGC AGG ATA TGT GCC TTC CTG CTG CGG AAA AAG CGT TTC GGG CAG TGG GCC AAG CAG CTG
 L   V   R   E   C   R   I   C   A   F   L   L   R   K   K   R   F   G   Q   W   A   K   Q   L
ACG GTC ATC AGG GAG GAC CAG CTC CTG TGT TAC AAA AGC TCC AAG GAT CGG CAG CCA CAT CTG AGG TTG GCA
 T   V   I   R   E   D   Q   L   L   C   Y   K   S   S   K   D   R   Q   P   H   L   R   L   A
CTG GAT ACC TGC AGC ATC ATC TAC GTG CCC AAG GAC AGC CGG CAC AAG CGT CAC GAG CTG CGT TTC ACC CAG
 L   D   T   C   S   I   I   Y   V   P   K   D   S   R   H   K   R   H   E   L   R   F   T   Q
GGG GCT ACC GAG GTC TTG GTG CTG GCA CTG CAG AGC CGA GAG CAG GCC GAG GAG TGG CTG AAG GTC ATC CGA
 G   A   T   E   V   L   V   L   A   L   Q   S   R   E   Q   A   E   E   W   L   K   V   I   R
GAA GTG AGC AAG CCA GTT GGG GGA GCT GAG GGA GTG GAG GTC CCC AGA TCC CCA GTC CTC CTG TGC AAG TTG
 E   V   S   K   P   V   G   G   A   E   G   V   E   V   P   R   S   P   V   L   L   C   K   L
GAC CTG GAC AAG AGG CTG TCC CAA GAG AAG CAG ACC TCA GAT TCT GAC AGC GTG GGT GTG GGT GAC AAC TGT
 D   L   D   K   R   L   S   Q   E   K   Q   T   S   D   S   D   S   V   G   V   G   D   N   C
TCT ACC CTT GGC CGC CGG GAG ACC TGT GAT CAC GGC AAA GGG AAG AAG AGC AGC CTG GCA GAA CTG AAG GGA
 S   T   L   G   R   R   E   T   C   D   H   G   K   G   K   K   S   S   L   A   E   L   K   G
TCA ATG AGC AGG GCT GCG GGC CGC AAG ATC ACC CGT ATC ATT GGC TTC TCC AAG AAG AAG ACA CTG GCC GAT
 S   M   S   R   A   A   G   R   K   I   T   R   I   I   G   F   S   K   K   K   T   L   A   D
GAC CTG CAG ACG TCC TCC ACC GAG GAG GAG GTT CCC TGC TGT GGC TAC CTG AAC GTG CTG GTG AAC CAG GGC
 D   L   Q   T   S   S   T   E   E   E   V   P   C   C   G   Y   L   N   V   L   V   N   Q   G
TGG AAG GAA CGC TGG TGC CGC CTG AAG TGC AAC ACT CTG TAT TTC CAC AAG GAT CAC ATG GAC CTG CGA ACC
 W   K   E   R   W   C   R   L   K   C   N   T   L   Y   F   H   K   D   H   M   D   L   R   T
CAT GTG AAC GCC ATC GCC CTG CAA GGC TGT GAG GTG GCC CCG GGC TTT GGG CCC CGA CAC CCA TTT GCC TTC
 H   V   N   A   I   A   L   Q   G   C   E   V   A   P   G   F   G   P   R   H   P   F   A   F
AGG ATC CTG CGC AAC CGG CAG GAG GTG GCC ATC TTG GAG GCA AGC TGT TCA GAG GAC ATG GGT CGC TGG CTC
 R   I   L   R   N   R   Q   E   V   A   I   L   E   A   S   C   S   E   D   M   G   R   W   L
GGG CTG CTG CTG GTG GAG ATG GGC TCC AGA GTC ACT CCA GAG GCG CTG CAC TAT GAC TAC GTG GAT GTG GAG
 G   L   L   L   V   E   M   G   S   R   V   T   P   E   A   L   H   Y   D   Y   V   D   V   E
ACC TTA ACC AGC ATC GTC AGT GCT GGG CGC AAC TCC TTC CTA TAT GCA AGA TCC TGC CAG AAT CAG TGG CCT
 T   L   T   S   I   V   S   A   G   R   N   S   F   L   Y   A   R   S   C   Q   N   Q   W   P
GAG CCC CGA GTC TAT GAT GAT GTT CCT TAT GAA AAG ATG CAG GAC GAG GAA GAG CCC GAG CGC CCC ACA GGA GCC
 E   P   R   V   Y   D   D   V   P   Y   E   K   M   Q   D   E   E   E   P   E   R   P   T   G   A
CAG GTG AAG CGT CAC GCC TCC TCC TGC AGT GAG AAG TCC CAT CGT GTG GAC CCG CAG GTC AAA GTC AAA CGC
 Q   V   K   R   H   A   S   S   C   S   E   K   S   H   R   V   D   P   Q   V   K   V   K   R
CAC GCC TCC AGT GCC AAT CAA TAC AAG TAT GGC AAG AAC CGA GCC GAG GAG GAT GCC CGG AGG TAC TTG GTA
 H   A   S   S   A   N   Q   Y   K   Y   G   K   N   R   A   E   E   D   A   R   R   Y   L   V
GAA AAA GAG AAG CTG GAG AAA GAG AAA GAG ACG ATT CGG ACA GAG CTG ATA GCA CTG AGA CAG GAG AAG AGG
 E   K   E   K   L   E   K   E   K   E   T   I   R   T   E   L   I   A   L   R   Q   E   K   R
GAA CTG AAG GAA GCC ATT CGG AGC AGC CCA GGA GCA AAA TTA AAG GCT CTG GAA GAA GCC GTG GCC ACC CTG
 E   L   K   E   A   I   R   S   S   P   G   A   K   L   K   A   L   E   E   A   V   A   T   L
GAA GCT CAG TGT CGG GCA AAG GAG GAG CGC CGG ATT GAC CTG GAG CTG GTG GCT GTG AAG GAG CGC
 E   A   Q   C   R   A   K   E   E   R   R   I   D   L   E   L   K   L   V   A   V   K   E   R
TTG CAG CAG TCC CTG GCA GGA GGG CCA GCC CTG GGG CTC TCC GTG AGC AGC AAG CCC AAG AGT GGG GAA ACT
 L   Q   Q   S   L   A   G   G   P   A   L   G   L   S   V   S   S   K   P   K   S   G   E   T
GCA AAT AAA CCC CAG AAC AGC GTT CCA GGA CAA CCT CTC CCT GTC AAC TGT GTT TCT GAG CTG AGG AAG AGG
 A   N   K   P   Q   N   S   V   P   E   Q   P   L   P   V   N   C   V   S   E   L   R   K   R
AGC CCA TCC ATC GTA GCC TCC AAC CAA GGA AGG GTG CTA CAG AAA GCC AAG GAA TGG GAA ATG AAG AAG ACC
 S   P   S   I   V   A   S   N   Q   G   R   V   L   Q   K   A   K   E   W   E   M   K   K   T
TAG
STOP
```

92

Figure 1: AFAP1L1 sequence

AFAP1L1 coding sequence was divided into codons with corresponding amino acid sequence. Start and stop codons are indicated in bold type.

Figure 2A

```
AFAP1     ------MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQE 54
AFAP1L1   MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
AFAP1L2   MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
Consensus         E  EL   L  LD E L  T   KK         L

AFAP1     TANSLE-----------------------APPQMPLPEIP-QPW-P-PDSGPE----E 83
AFAP1L1   DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAAQLPE----E 116
AFAP1L2   TINKQQNAESQGKAPEEQGLLP-------NGEPSQHSSAPQKSLPDLPPPKPIPERKQLA 113
Consensus                                   P            P L     P

AFAP1     LPTSSLPEGYYEEAVPLSPGKAPEYITSN------------------------------ 112
AFAP1L1   LPNKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELK 176
AFAP1L2   IPKTESPEGYYEEAEFYDTSLNE------------------------------------ 136
Consensus   P     PE YYEEA P

AFAP1     ---YDSDAMSSSYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQW 169
AFAP1L1   SSYNDSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQW 236
AFAP1L2   ----DGEAVSSSYESYDEED-GSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQW 191
Consensus       D  A SSSYESYDEE      K      QWPS EA  L    ICAFL RKK  GQW

AFAP1     TKLLCVIKDTKLLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVL 229
AFAP1L1   AKQLTVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVL 296
AFAP1L2   AKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVL 251
Consensus   K L VI    LLCYKSSKD  P     L          K    K H L  T        VL

AFAP1     AVQSKEQAEQWLKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVV 289
AFAP1L1   ALQSREQAEEWLKVIREVSKPVGGAEGVEVPRSP-VLLCKLDLDKRLSQEKQTSDSDSVG 355
AFAP1L2   GLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTP--------DAQRFNCQKPDIAEKYLS 303
Consensus    QS  QAE WL VI E                  P

AFAP1     ENGITTCNGK----EQVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKK-PSTDEQTSSA 344
AFAP1L1   VGDNCSTLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSST 415
AFAP1L2   ASEYGSSVDG-------HPEVPETKDVKKKCS--AGLKLSNLMNLGRKK--STSLEP--V 350
Consensus                       K    S    G K           KK

AFAP1     EEDVPTCGYLNVLSNSRWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSK 404
AFAP1L1   EEEVPCCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPR 475
AFAP1L2   ERSLETSSYLNVLVNSQWSRWCSVRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPD 410
Consensus   E       YLNVL N  W  RWC   N L F  D       L GCEV  P

AFAP1     HPLTFRLLRNGQEVAVLEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQ 464
AFAP1L1   HPFAFRILRNRQEVAILEASSSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVS 535
AFAP1L2   HLYSFRILHKGEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVS 470
Consensus   H   FR L    E A LEA  SE MG W G LL  E GS    PE    YDY D

AFAP1     TAKQTFCFMNRRVISANPYLGGTSNG----------------YAHPSGTALHYDDVPCI 507
AFAP1L1   AGRNSFLYAR----------SCQNQ----------------WPEPR----VYDDVPYE 563
AFAP1L2   AAKNSLLLMQRKFSEPNTYIDGLPSQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNE 530
Consensus        F                                         P         D P

AFAP1     NGSL-------------------------------------------------------- 511
AFAP1L1   KMQD-------------------------------------------------------- 567
AFAP1L2   RESDRVYLDLTPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCP 590
Consensus

AFAP1     -------------KGKKPPVASNGVTGKGKTLSSQPKKADPAAVVKRTGS---------N 549
AFAP1L1   -------------EEPERPTGAQ-VKRHASSCSEKSHRVDPQVKVKRHAS---------S 604
AFAP1L2   ENLGEQQLESLEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVT 650
Consensus                             S        D         S

AFAP1     AAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDLRAAIEVNAGRRPQA 609
AFAP1L1   ANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQFKRELKFATRSSPCAK-LK 663
AFAP1L2   SAEIKLGKNRTEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLA 710
Consensus     K GKNR E    R    EKE L KKE   R L LRE      L

AFAP1     ILFSKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGLAIEPKSGTSSPQSP 669
AFAP1L1   ALFSAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSK-----PKS- 717
AFAP1L2   SLEQKLKEIDEECRGEESRRVDLELSIMEVKDNLKKAEAGPVTLGTTVDTT--------- 761
Consensus    LF          CR  E  R   LEL    VK  L     AG  LG

AFAP1     VFRHRTLENSPISSCDTSDTEGPVPVNSAAVLKKSQAAPGSSPCRGHVLRKAKEWELKNG 729
AFAP1L1   ----GETANKPQNSVP----EQPLPVNCVSELRKRSPSIVAS-NQGRVLQKAKEWEMKKT 768
AFAP1L2   ---HLENVSPRPKAVTPASAPDCTPVNSATTLKNRPLSVVVT-GKGTVLQKAKEWEKKGA 817
Consensus                       PVN    L         G VL KAKEWE K

AFAP1     T 730
AFAP1L1   -
AFAP1L2   S 818
Consensus
```

94
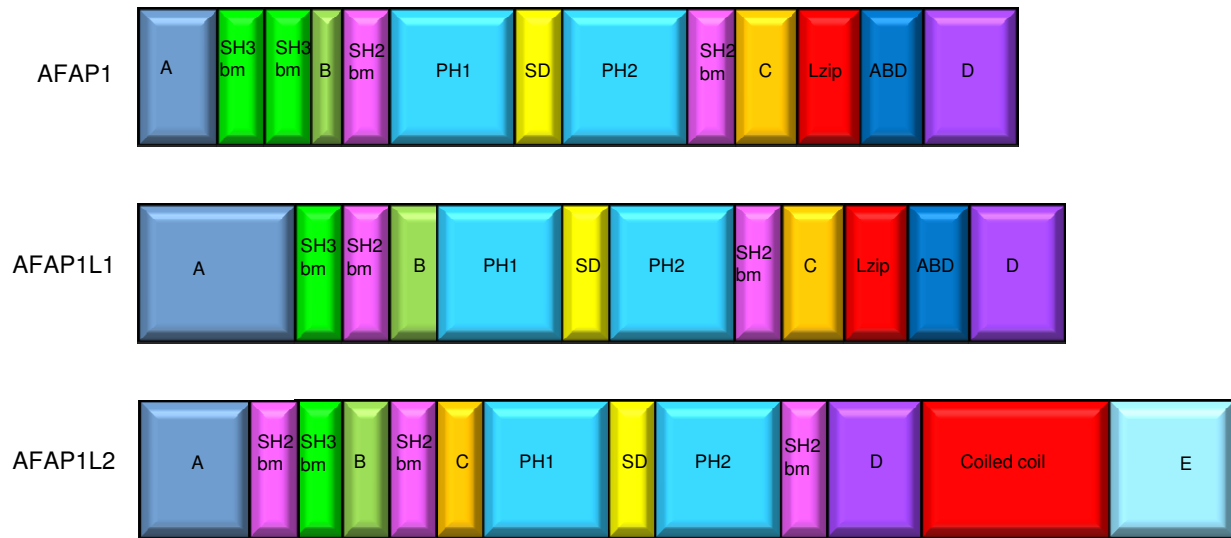
Figure 2B

Figure 2: AFAP family members share both sequence and domain similarity

(A) AFAP1, AFAP1L1 and AFAP1L2 amino acid sequences were compared using ClustalW2 alignment (Larkin et al., 2007). Consensus sequence between all three family members is labeled as consensus. Intron/exon boundaries are marked by red letters. Predicted SH3 binding motifs are highlighted in green, predicted SH2 binding motifs in pink, predicted PH domains in light blue, predicted Substrate Domain (SD) in yellow, predicted leucine zipper (AFAP1, AFAP1L1) and coiled coil (AFAP1L2) in red and predicted actin binding domain in dark blue. The AFAP1L1 peptide sequence used to create 1L1-CT antibody is underlined.

(B) Modular domain organization of AFAP family members was compared. SH3bm = SH3 binding motif, SH2bm = SH2 binding motif, PH1 = pleckstrin homology domain 1, PH2 = pleckstrin homology domain 2, SD = serine/threonine rich substrate domain, Lzip = leucine zipper, ABD = actin binding domain. Sequences that do not correlate with an identified type of modular domain or motif are labeled "A, B, C, D or E".

Figure 3

A

AFAP1L1 →

IB: AFAP1L1

IB: γ Tubulin

B

IP: IgG, AFAP1L1, AFAP1

AFAP1L1 →

IgG Heavy Chain →

AFAP1 →

IB: AFAP1L1   AFAP1

C

GFP-AFAP1L1 →
GFP-AFAP1 →

β−actin →

IB: 1L1-CT   IB: AFAP1 (F1)   IB: GFP
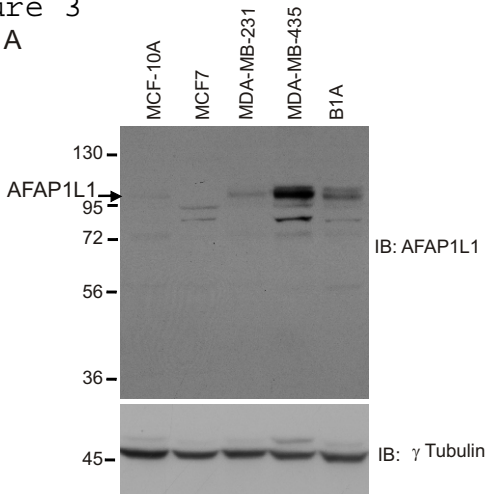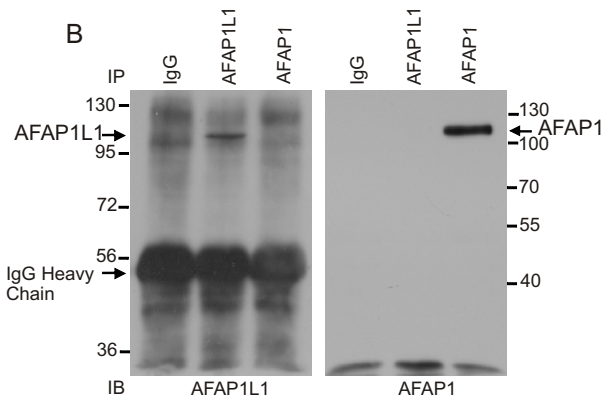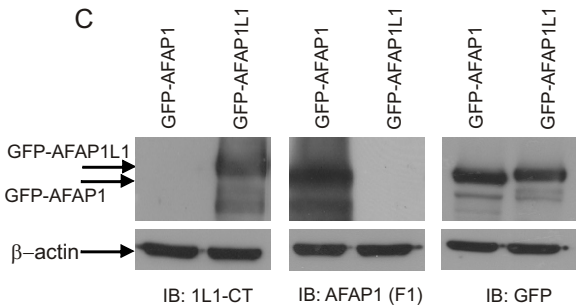
D

AFAP1L1 antibody peptide	KPKSGETANKPQNS
AFAP1L1 human	KPKSGETANKPQNS
AFAP1L1 chimpanzee	KPKSGETANKPQNS
AFAP1L1 mouse	KNKSQDTTNKPQSN
AFAP1L1 rat	KSKSQETTNKPQSS

97

Figure 3: A novel antibody specifically recognizes AFAP1L1

(A) Cell lines were lysed in 2X SDS buffer, resolved by 8% SDS-PAGE and transferred to PVDF. The Sigma antibody Ab2 specifically recognized a protein band of 115 kDa. Bands identified as AFAP1L1 and recognized by the Ab2 antibody (top panel) are indicated by an arrow. Gamma tubulin was used as a loading control. Not shown, but in an adjacent lane, was a lysate prepared from 293T cells transfected with untagged AFAP1L1 which was used in identifying the band corresponding to AFAP1L1.

(B) Endogenous AFAP1L1 and AFAP1 were specifically immunoprecipitated from Cos-1 cells using 5μg 1L1-CT and F1 polyclonal antibodies respectively and the resolved proteins detected with 1L1-CT (left panel) or AFAP1 monoclonal antibodies (BD Transduction, right panel). Rabbit IgG antibody was used as a control. Note the differences in the molecular weight markers for each western.

(C) GFP-AFAP1L1 and GFP-AFAP1 were overexpressed in Cos-1 cells and lysed in 2X SDS buffer. Lysates were resolved by 8% SDS-PAGE, transferred to PVDF membrane and immunoblotted with either AFAP1L1 (1L1-CT) or AFAP1 (F1) antibody. GFP (top right western) and β-actin (bottom westerns) were used as loading controls.

(D) The peptide sequence used to create antibody 1L1-CT was compared to analogous sequences in human, chimpanzee, mouse and rat to show similarity between antibody binding sites.

Figure 4
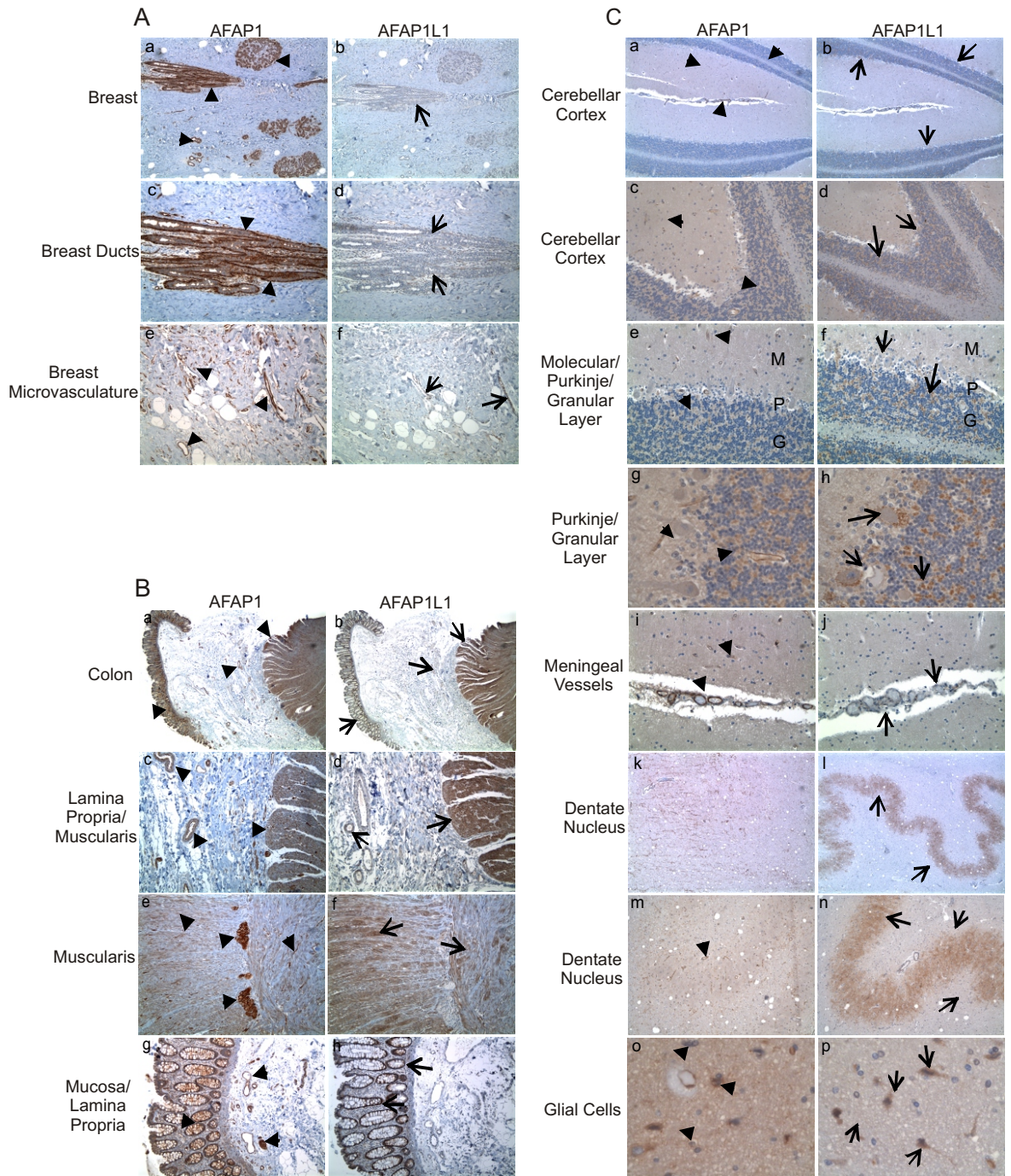
Figure 4: Immunohistochemical analysis of AFAP1L1 shows differential expression from AFAP1 in human tissue

Paraffin-embedded human breast (**A**), colon (**B**) and brain (**C**) tissues were analyzed for AFAP1 and AFAP1L1 localization using F1 and 1L1-CT antibodies respectively. Breast regions include breast ducts (4A panels a-d), breast lobules (4A panels a,b) and microvasculature (4A panels e,f). Colon regions include the mucosa (4B panels a,b,g,h), lamina propria (4B panels a-d, g,h) and muscularis (4B panels a-f). Brain regions include the cerebellar cortex (4C panels a-j), dentate nucleus (4C panels k-n) and glial cells (4C panels o,p). AFAP1L1 is designated by long thin arrows while AFAP1 is designated by arrowheads.

Figure 5

A

GFP-AFAP1L1 Cortactin Actin Merge

B

AFAP1L1

IgG control

Actin Merge

C

AFAP1L1 Cortactin Actin Merge

IgG control Cortactin Actin Merge
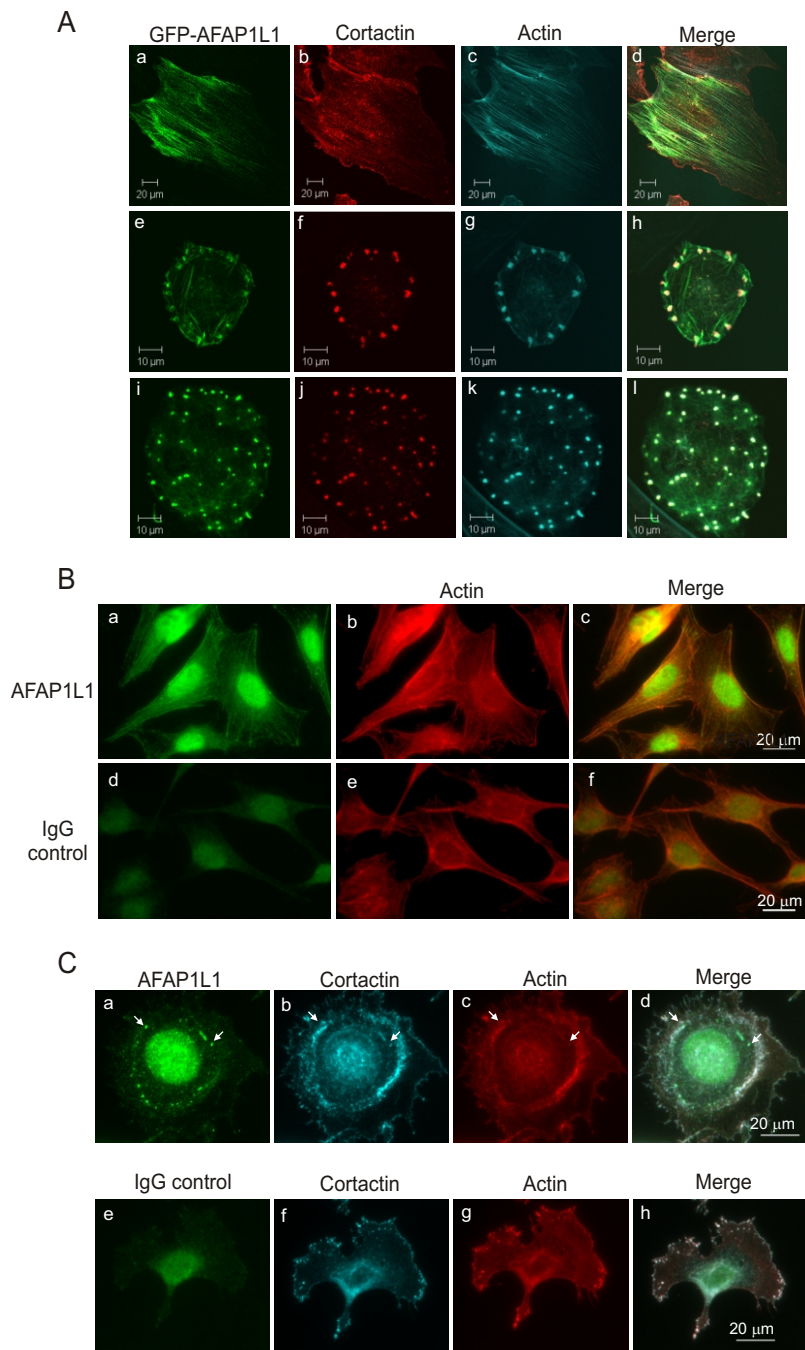
Figure 5: Subcellular localization of GFP-AFAP1L1 shows association with actin and invadosomes

(A) A7r5 cells transiently expressing GFP-AFAP1L1 were plated onto fibronectin-coated coverslips and immunolabeled for cortactin (Millipore). Actin was visualized with TRITC-phalloidin (Sigma). Representative images of cells with well-formed stress fibers or podosome formation are shown.

(B) MDA-MB-435 cells were plated on fibronectin coated coverslips and immunolabeled for endogenous AFAP1L1 (Sigma Ab2, panel a-c). Rabbit IgG was used as a control (panel d-f) and actin was visualized with AlexaFluor labeled phalloidin. Epifluorescence images of representative cells are shown.

(C) MDA-MB-435 cells were transfected with Src 527F construct, plated onto fibronectin coated coverslips and immunolabeled for AFAP1L1 (Sigma Ab1, panel a-d). Rabbit IgG was used as control antisera for AFAP1L1 antibody (panel e-h). Cortactin was immunolabeled with monoclonal anti-cortactin antibodies (4F11, Millipore) and actin was visualized by AlexaFluor labeled phalloidin (panel c, g). Examples of AFAP1L1 co-localizing to invadopodia, actin and cortactin rich punctate structures are marked with white arrows.

Figure 6
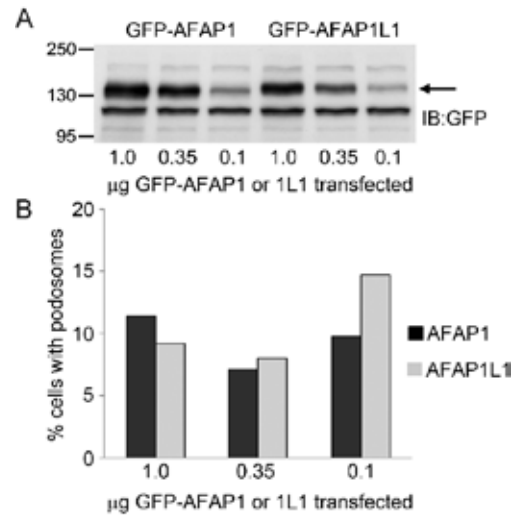
Figure 6: Podosome formation in A7r5 transfected with GFP-AFAP1 or GFP-AFAP1L1 plasmids.

A7r5 cells were transfected with the indicated amounts of plasmids encoding either GFP-AFAP1 or GFP-AFAP1L1 (in combination with pcDNA3.1) to bring the total amount of plasmid DNA to 1 µg for each transfection. Twenty four hours post-transfection, the cells were transferred to fibronectin-coated coverslips. Forty-eight hours after transfection, cells were processed for immunofluorescence analysis or used to prepare whole cell SDS lysates. (A) Equal amounts of SDS lysates were resolved by 8% SDS-PAGE and then transferred to PVDF membranes and subsequently probed with an antiserum that recognizes GFP. (B) Cells expressing GFP-AFAP1 or GFP-AFAP1L1 were assessed for podosome formation and the percentage of cells exhibiting podosomes for each transfection was calculated. 150 to 300 cells were counted for each transfection. Panel A and B represents one experiment out of two independently performed experiments.

Figure 7
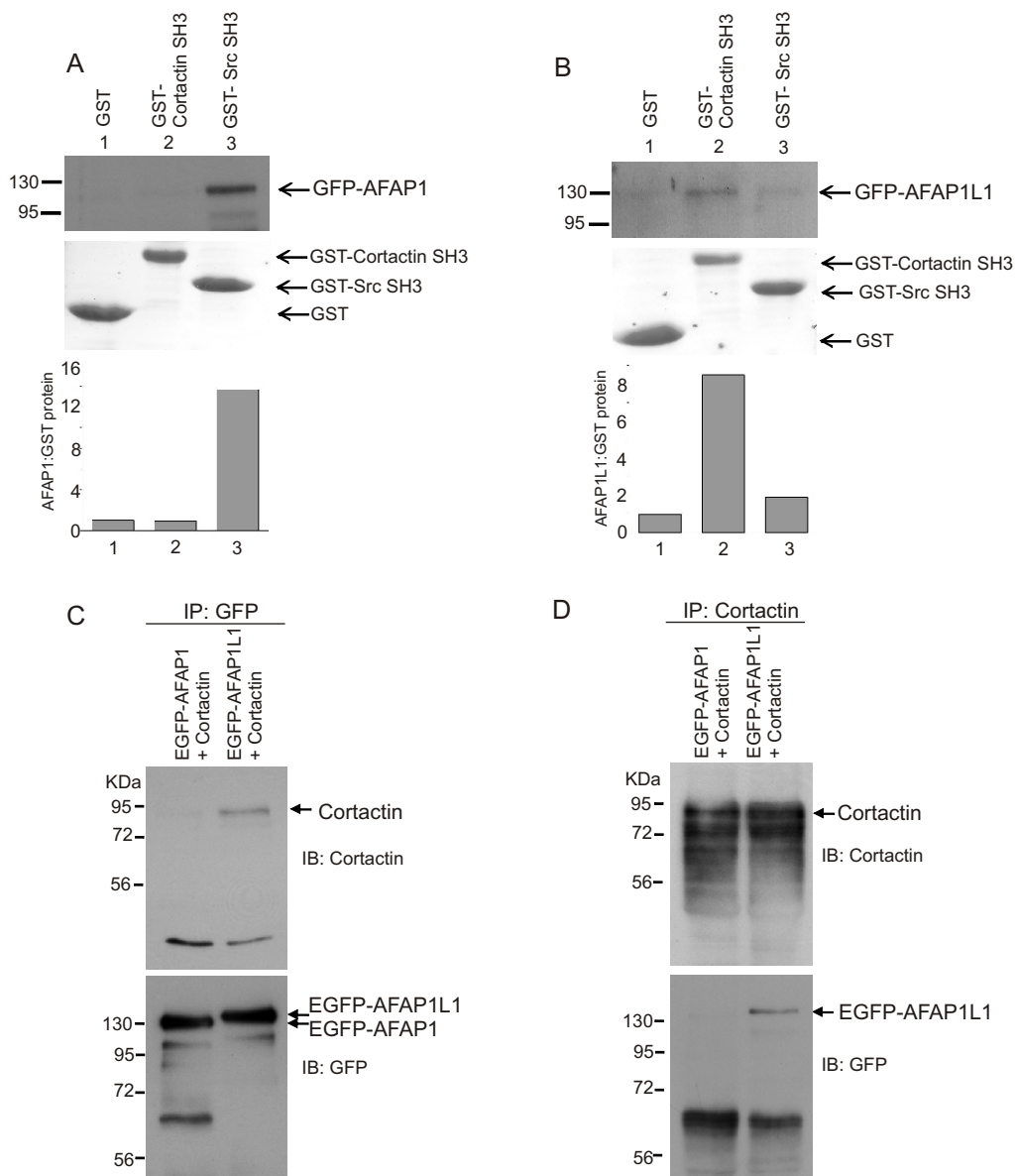
Figure 7: AFAP1L1 interacts with cortactin SH3 domain

(A) 1mg of lysate from 293T cells transiently expressing GFP-AFAP1 was incubated with 50μg of GST, GST-Src-SH3 domain or GST-cortactin-SH3 domain bound fusion protein to glutathione Sepharose 4B beads and probed for GFP through western blot analysis (upper panel). The lower panel represents a GelCode Blue Stain (ThermoScientific) of GST or GST fusion protein. A graph representing the ratio of AFAP1 pulled down by each GST fusion protein compared to GST control using scanning densitometry is shown.

(B) 1mg of lysate from 293T cells transiently expressing GFP-AFAP1L1 was incubated with 50μg of GST, GST-Src-SH3 domain or GST-cortactin-SH3 domain bound to glutathione Sepharose 4B beads and probed for GFP through western blot analysis (upper panel). The lower panel represents a GelCode Blue Stain (ThermoScientific) of GST or GST fusion protein. A graph representing the ratio of AFAP1L1 pulled down by each GST fusion protein compared to GST control using scanning densitometry is shown.

(C) 293T transiently expressing cortactin with either GFP-AFAP1 or GFP-AFAP1L1 were immunoprecipitated with anti-GFP antibody, probed for cortactin through western blot analysis (upper panel) and then re-probed for GFP tagged proteins (lower panel).

(D) 293T transiently expressing cortactin with either GFP-AFAP1 or GFP-AFAP1L1 were immunoprecipitated with anti-cortactin antibody, probed for GFP tagged proteins through western blot analysis (upper panel) and then re-probed for cortactin (lower panel).

Supplemental Figure 1
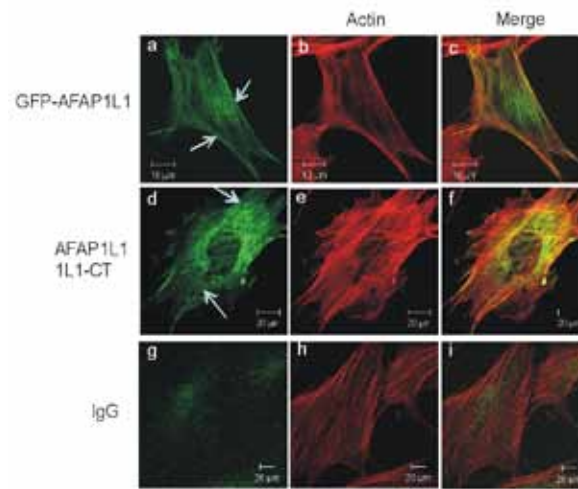
Supplemental Figure 1: AFAP1L1 overexpression in MEF cells

Mouse embryo fibroblast cells were transfected with 5 μg of either GFP-AFAP1L1 or untagged AFAP1L1 (in pcDNA3.1). Untagged AFAP1L1 was probed using 1L1-CT antibody (1:200). GFP-AFAP1L1 localized to stress filaments and cortical actin (panels a-c). Untagged AFAP1L1 also localized to stress filaments and showed some diffuse cytoplasmic staining (panels d-f). Actin was visualized by TRITC-phalloidin (Invitrogen). Rabbit IgG was used as a control (panels g-i).

Supplemental Figure 2
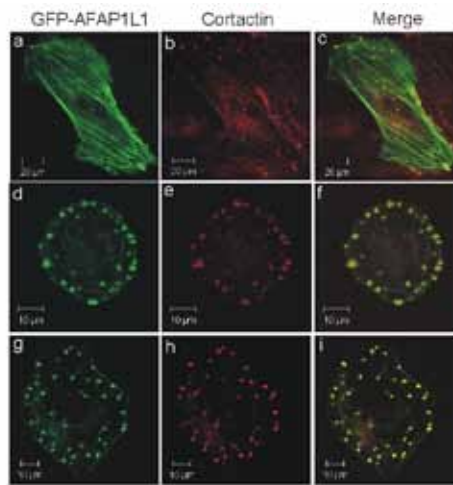
Supplemental Figure 2: Colocalization of GFP-AFAP1L1 and cortactin

A7r5 cells transiently expressing GFP-AFAP1L1 were plated onto fibronectin-coated coverslips and immunolabeled for cortactin (Millipore). Representative images of cells with well-formed stress fibers (panels a-c), podosomes around the periphery of the cell (panels d-f) or podosomes throughout the cell (panels g-i) are presented showing colocalization of GFP-AFAP1L1 and cortactin in podosomes.

Supplemental Figure 3
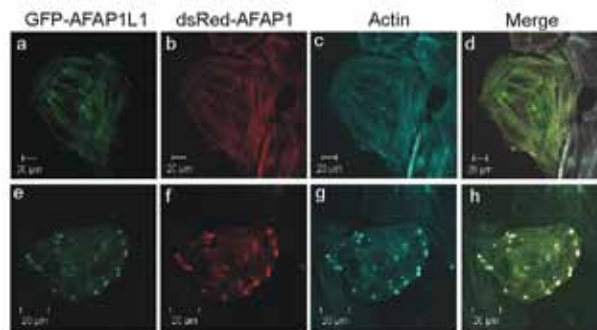
Supplemental Figure 3: Overexpression of dsRed-AFAP1 and GFP-AFAP1L1 in A7r5 cells

A7r5 cells transiently expressing dsRed-AFAP1 and GFP-AFAP1L1 were plated onto fibronectin coated coverslips. Actin was visualized by BODIPY-650/665 (Invitrogen). Representative cells with well formed stress fibers (panels a-d) and podosome formation (panels e-h) show colocalization of dsRed-AFAP1 and GFP-AFAP1L1 on both stress fibers and in podosomes.

CHAPTER 3

AFAP1L1 Additional Data

I. A panomics screen using the AFAP1L1 SH3 binding motif identified potential AFAP1L1 binding partners

A TranSignal SH3 Domain Array 1 of recombinant binding sites of SH3 domain containing proteins spotted onto a membrane was obtained from Panomics (Affymetrix) to probe for potential SH3 domain binding partners of the AFAP1L1 SH3 binding motif (Figure 1A). A biotinylated peptide of the AFAP1L1 binding motif (biotin-ADLPPPLPNKPPE) was synthesized by EZBiolab Custom Peptide Service so as to probe the TranSignal SH3 Domain Array 1 for binding partners. 1mg of biotin-conjugated peptide was brought up in 1ml of sterile water so as to produce a peptide solution of 1mg/ml. Peptide solution was diluted to 0.1µg/µl and 15µl (1.5µg total) was incubated with 15µl of 1.25µg/µl streptavidin-horseradish peroxidase (HRP) conjugate for 30 minutes by rocking at 4°C. Streptavidin was used to bind to the biotin linker of the AFAP1L1 peptide so as to create an HRP-linked AFAP1L1 binding motif peptide. The peptide solution was then placed into 5ml of 3% bovine serum albumin (BSA) in TBST and kept at 4°C.

The SH3 Domain Array membranes were incubated in 5% powdered milk dissolved in TBST to block overnight at 4°C, washed for 30 minutes with TBST and incubated for two hours in 3% BSA/TBST. The membranes were then incubated with either 10µg or 15µg of peptide solution overnight at 4°C. After washing for 30 minutes in TBST, the membranes were developed using Pierce ECL Western Blotting Substrate.

Five duplicate SH3 domain spots giving the strongest signal with biotinylated AFAP1L1 SH3 binding motif were considered for potential AFAP1L1 binding (Figure 1B).

Spectrin alpha chain (non-erythrocytic)

Spectrin alpha chain (alpha-II spectrin/SPTAN1) is a member of a family of scaffolding proteins that stabilize the plasma membrane and is involved in various cell processes such as cell growth and differentiation, vesicular trafficking and neurotransmitter release, among others. It has also been shown to be involved in DNA repair (McMahon et al., 2009). Alpha-II spectrin is a rod-like protein of 2,472 amino acids that heterodimerizes with beta-spectrin subunits and contains 23 spectrin repeats, three EF-hand domains and two potential calcium binding domains (2010; Jain et al., 2009). The SH3 domain is located in the central portion of the protein and has been shown to bind both cSrc and low molecular weight protein-tyrosine phosphatase isoform A (Nedrelow et al., 2003; Nicolas et al., 2002).

Spectrin alpha chain (erythrocytic)

Spectrin alpha chain erythrocyte (SPTA1) is a 2,419 amino acid protein that contains 21 spectrin repeats, three EF-hand domains, two potential calcium binding domain and a central SH3 domain (2010; Jain et al., 2009). Like SPTAN1 in non-erythrocytic cells, SPTA1 provides structural support to the plasma membrane of the erythrocyte by heterodimerizing with beta-spectrin subunits and providing a scaffold for other cytoskeletal proteins (Chakrabarti et al., 2006). The SH3 domain of SPTA1 has been shown to have a possible interaction with Fas ligand (Voss et al., 2009).

Avian sarcoma virus CT10 oncogene homolog, domain #2

Avian sarcoma virus CT-10 oncogene homolog (Crk) is a 304 amino acid protein that acts as an adaptor molecule. Crk exists in two isoforms, Crk-I and Crk-II which have distinct roles in the cell. Crk-I contains an SH2 domain and one SH3 domain while Crk-II contains one SH2 domain, two SH3 domains and multiple sites for tyrosine, serine and threonine phosphorylation (2010; Gelkop et al., 2003; Jain et al., 2009). Biotinylated AFAP1L1 binding motif had the ability to interact with the C-terminal SH3 domain of Crk-II. While the N-terminal SH3 domain of Crk-II has multiple known binding partners, the exact binding function of the N-terminal SH3 domain is still unknown although it has been shown to be involved in the activation of Abl kinase (Gelkop et al., 2003; Reichman et al., 2005). The N-terminal SH3 domain of Crk-II was not available on the TranSignal SH3 Domain Array 1.

Cortactin

Cortactin (CTTN) is a 550 amino acid rod-shaped protein involved in signaling processes such as cell adhesion, migration and invasion and plays an important role by stabilizing branching points of the actin cytoskeleton. Cortactin consists of an N-terminal acidic domain, six and one half tandem cortactin repeats of which the fourth repeat is necessary for actin binding, an α-helical domain, a region rich in prolines, tyrosines and serines and a C-terminal SH3 domain. The SH3 domain of cortactin is known to interact with WASp interacting protein (WIP), neural Wiskott-Aldrich syndrome protein (N-WASp), myosin light chain kinase (MLCK), and dynamin-2 (Ammer and Weed, 2008).

Post-synaptic density protein 95

Post-synaptic density protein 95 (PSD-95) is a 724 amino acid scaffolding protein found in the post-synaptic density, a dense area in the postsynaptic region of a neuron rich in proteins involved with the cytoskeleton and cell adhesion such as adaptor proteins, signaling kinases, phosphatases, membrane-bound receptors and G-proteins (Boeckers, 2006). PSD-95 consists of three PDZ domains, an SH3 domain, a guanylate kinase-like domain and multiple sites for serine and tyrosine phosphorylation (2010; Jain et al., 2009). PSD-95 is a major component of the post synaptic density and has multiple roles such as stabilizing membrane protein localization and mediating synaptic plasticity (Han and Kim, 2008). PSD-95 has been shown to interact with the Src family kinases Src, Lyn and Yes, however the SH3 domain of PSD-95 does not appear to have typical SH3 binding ability and it remains to be seen if it is the SH3 domain of PSD-95 that allows interaction with SFK (Kalia and Salter, 2003).

While the exact function of AFAP1L1 is unknown, due to its similarity with family member AFAP1, we hypothesized that AFAP1L1 would also have a role in actin dynamics. It is interesting that the results of the panomics screen resulted in proteins that have direct roles in the regulation of the actin cytoskeleton and this may strengthen the hypothesis of a role for AFAP1L1 with the actin cytoskeleton. Two SH3 binding domains had the strongest reaction with the AFAP1L1 binding motif: cortactin and PSD-95. Although the results from a panomics array can suggest possible protein-protein interactions, the interaction of SH3 domain peptides with the AFAP1L1 SH3 binding motif peptide could produce artificial results due to such forced interactions. Due to its accessibility and previous data results showing colocalization of AFAP1L1 and cortactin

in cells, we chose to determine if AFAP1L1 and cortactin were true binding partners in our AFAP1L1 characterization studies.

II. Postsynaptic Density Protein 95

The postsynaptic density (PSD) is an electron-dense portion of a neuronal synapse in which multiple classes of proteins such as cell-adhesion proteins, cytoskeletal proteins, scaffolding and adaptor proteins, membrane-bound receptors and channels, G-proteins and modulators and signaling molecules can be found. PSDs can be found on the dendritic shaft or at the tip of a dendritic spine and lie just below the membrane of glutamatergic synapses (Boeckers, 2006). Glutamatergic synapses are excitatory synapses that can contain any of five glutamate transporters, GLAST, GLT-1, EAAC1, EAAT4 and EAAT5, which are differentially expressed throughout the central nervous system (Tanaka, 2000). As an example of a glutamatergic synapse, Purkinje cells are the largest neurons in the brain and are found in a single layer between the molecular and granular layers of the cerebellum. While the output from Purkinje cells is inhibitory, the major input to a Purkinje neuron is excitatory. It is thought that a Purkinje cell refines the multitude of excitatory pulses it receives into one major inhibitory stimulus sent to the deep cerebellar nuclei which can control motor output. Purkinje cells are contacted in two ways by excitatory synapses. Parallel fibers are projections of mossy fibers from the brain stem and spinal cord which pass through the granular cell layer and create thousands of weak excitatory synapses onto Purkinje cells. The major input of excitatory stimulus applied to a Purkinje cell is through projections from the brain stem called climbing fibers. Each Purkinje cell is innervated by a single climbing fiber which wraps itself

around the body of the Purkinje cell and thus creates multiple signaling sites. Basket and stellate cells of the molecular layer which surround the Purkinje cell are thought to relay weak inhibitory signals; however the major input into a Purkinje cell is through excitatory synapses. It is at these excitatory glutamatergic synapses on Purkinje cells that four of the five glutamate transporters are found with EAAT4 being found in the highest concentration in dendritic spines (Takayasu et al., 2009).

Correct regulation of glutamate receptors and subsequent signaling in the PSD is essential and is highly regulated by PSD-95/DLG/ZO-1 (PDZ) domain containing proteins. PSD-95 (SAP90) is a member of the membrane-associated guanylate kinase (MAGUK) family and is considered to be the best characterized and prototype MAGUK family member in the PSD (Boeckers, 2006). PSD-95 is an adaptor protein that consists of three PDZ domains, an SH3 domain and a guanylate kinase-like domain (Han and Kim, 2008). It is through these domains that PSD-95 has the ability to bind to other PSD-95 proteins to form a scaffold, traffic proteins and receptors to the synapse and link receptors at the excitatory synapse to the underlying actin cytoskeleton of the dendritic spine so as to regulate synaptic plasticity and signaling (Boeckers, 2006). While some MAGUK proteins are expressed in early development, PSD-95 is expressed predominantly in later stages (Kim and Sheng, 2004).

PSD-95 is positioned at the synaptic membrane and functions to tether *N*-methyl *D*-aspartate (NMDA) receptors and other transmembrane receptors such as ADAM22, neuroligin 1 and synaptic-like adhesion molecule (SALM) to the cellular membrane by interaction with their cytoplasmic tails (Boeckers, 2006; Han and Kim, 2008). It additionally interacts with other intracellular PSD proteins such as A-kinase anchor

protein 79 (AKAP79), spine associated Rap-Gap (SPAR) and synaptic Ras GTPase-activating protein (SynGAP) by virtue of its adaptor domains to regulate their localization to the synaptic membrane (Boeckers, 2006). These proteins link PSD-95 into a signaling complex that ties it to the underlying actin cytoskeleton. PSD-95 has been shown to directly bind the Src family kinases Src, Lyn and Yes which are found in the PSD (Kalia and Salter, 2003). As signaling to NMDA receptors has been shown to play a role in synaptic plasticity, it is possible that PSD-95 may link together signals from the NMDA receptor to Src family kinases and thus result in alterations of the underlying cytoskeleton to affect dendritic spine morphogenesis. The alteration in morphology of dendritic spines regulated by the actin cytoskeleton is crucial for higher brain functions such as learning and memory (Sekino et al., 2007).

While there are many other important proteins in the postsynaptic density, cortactin in particular for its role in actin rearrangement and dendritic spine morphogenesis, PSD-95 plays a central role in linking synaptic signaling to cell morphology (Sekino et al., 2007). Due to the high level of actin dynamics in the postsynaptic density, it is conceivable that AFAP1L1 could play a role in the postsynaptic density as it is a cortactin binding partner and is hypothesized to play a role in actin rearrangement.

III. Differential immunohistochemical staining of AFAP1L1 in human brain

The dentate nucleus is one of four grey matter structures (dentate nucleus, vestibular nucleus, fastigial nucleus and globose/emboliform nucleus) found deep within the white matter of the cerebellum (Squire, 2008). Very little data exists on the functions

of these nuclei due to their small size and inability to be rigorously tested by imaging techniques. The dentate nucleus is the most easily identifiable deep cerebellar nuclei. A recent study determined that activation in the dentate nucleus mainly occurs during times of complex motor, sensorimotor and cognitive tasks such as exploratory movements, procedural memory, emotional and cognitive functions and cognitive tasks. Notably, the dentate nucleus was highly active in puzzle solving, planning tasks and verbal memory (Habas, 2010).

To determine differences in tissue localization between AFAP1 and AFAP1L1, tissues with known expression of AFAP1 were chosen for analysis. AFAP1 and a brain specific isoform, AFAP-120, are both expressed in the embryonic brain of mouse pups while proteins levels are decreased in adult mice except in the olfactory bulb (Clump et al., 2003). While AFAP1 and AFAP1L1 had similar staining patterns in the cerebellum with staining in microvasculature and around granule cells, AFAP1L1 had a distinct pattern of staining that was found extending from the molecular layer, surrounding the Purkinje neurons of the cerebellar cortex, in distinct locations around the granule cells (Figure 2A) and also surrounding the neurons of the dentate nucleus (Figure 2B). This was in contrast to AFAP1 which was not detected in these regions.

Work done by Tabakoff et al. in HXB/BXH recombinant inbred rats showed an association of AFAP1L1 with alcohol preference and consumption (Tabakoff et al., 2009). Rats were first exposed to 10% ethanol and then given a choice between water and 10% ethanol for seven weeks followed by removal of the brain and RNA extraction for microarray analysis. AFAP1L1 showed a significant increase in expression in the brains of two different strains of mice that preferred alcohol over water using two separate array

platforms. Tabakoff et al. hypothesized that AFAP1L1 is associated with the trafficking of the GABA$_A$ receptor by virtue of its association with actin organization, and that AFAP1L1 is located in the postsynaptic density in association with actin filaments and dynamin which supports our hypothesis that AFAP1L1 may be found in this area.

The postsynaptic density, as previously described, is found just below the membrane of the dendritic spine or dendritic shaft. Distinct from AFAP1, AFAP1L1 is found surrounding the Purkinje neurons and neurons of the dentate nucleus while staining appears to be excluded from the cell body. Due to its localization, it can be hypothesized that AFAP1L1 could be found in the postsynaptic densities of these cells. Postsynaptic densities can by isolated from brain tissue of multiple species through isolation with Triton X-100 and surveyed by Coomassie stain and mass spectrometry or western blot (Carlin et al., 1980; Walikonis et al., 2000). Isolation of the postsynaptic density fraction from cerebellar tissue samples followed by SDS-PAGE analysis with AFAP1L1 specific antibodies could provide a means to determine if AFAP1L1 is found in the postsynaptic density.

IV. **Mutating the SH3 binding motif of AFAP1L1 so as to determine specific residues necessary for cortactin binding**

A Src homology 3 (SH3) domain is a conserved protein domain of approximately 50-70 amino acids that preferentially binds to an SH3 binding motif, a poly-proline rich region of a peptide in which there is a conserved PXXP motif. The amino acids surrounding the PXXP motif confer specificity for the binding of particular SH3 domains. Cortactin contains an SH3 domain in the C-terminus of the protein that

resembles that of the Src family kinases and has been shown to bind a variety of proteins including dynamin-2, N-WASp and MLCK among others (Ammer and Weed, 2008; Sparks et al., 1996; Weed and Parsons, 2001). AFAP1L1 has also been shown to interact with a GST fusion protein of the cortactin SH3 domain. We have shown previously with prototype AFAP family member AFAP1 that mutation of a key proline at amino acid 71 to an alanine (P^{71A}) in the N-terminal SH3 binding motif decreases the ability of AFAP1 to interact with and activate cSrc (Guappone and Flynn, 1997). AFAP1L1 contains one SH3 binding motif with an amino acid sequence of DLPP^{115}PLP^{118}NKP^{121}. Prolines at amino acids 115, 118 and 121 fall into the PXXP motif and may affect the binding ability of AFAP1L1 to cortactin. Thus, we sought to determine if mutation of key proline residues in the AFAP1L1 binding motif could abrogate cortactin binding.

The proline residue at amino acid 115 is encoded by a CCA codon. Mutation of the first cytosine in the codon to a guanine (CCA➔GCA) allows for only a single nucleotide alteration which changes the encoded amino acid from a proline to an alanine. This results in a similar non-polar amino acid so as to limit major modifications that would affect the binding and folding of the protein. Primers used to mutate P^{115A} are as follows with the nucleotide mutation highlighted in yellow:

Forward 5' CC GAC CTG CCT GCA CCG CTC CCC 3'

Reverse 3' GGG GAG CGG TGC AGG CAG GTC GG 5'

The proline residue at amino acid 118 is encoded by a CCC codon. Mutation of the first cytosine to a guanine results in a single amino acid from proline to alanine. Primers used to mutate P^{118A} are as follows with the nucleotide mutation highlighted in yellow:

Forward 5' CCT CCA CCG CTC <mark>G</mark>CC AAC AAG CCT C 3'

Reverse 3' G AGG CTT GTT GG<mark>C</mark> GAG CGG TGG AGG 5'

The proline residue at amino acid 121 is encoded by a CCT codon. Mutation of the first cytosine to a guanine results in a single amino acid change from proline to alanine. Primers used to mutate P^{121A} are as follows with the nucleotide mutation highlighted in yellow:

Forward 5' CTC CCC AAC AAG <mark>G</mark>CT CCC CCT GAG G 3'

Reverse 3' C CTC AGG GGG AG<mark>C</mark> CTT GTT GGG GAG 5'

Mutagenesis of P^{115A}, P^{118A}, and P^{121A}, was accomplished by use of the QuikChange Site-Directed Mutagenesis Kit from Agilent (Stratagene). GFP-tagged AFAP1L1 was mutated so as to contain each separate P ➔ A mutation using the following reaction methods:

10X Reaction Buffer	5µl
GFP-AFAP1L1 plasmid	10ng
dNTP mix	1µl
Forward primer	125ng
Reverse primer	125ng
Pfu Ultra DNA Polymerase	1µl
Water	To 50µl

Reaction mixtures were subjected to polymerase chain reaction in the following manner:

Segment	Cycles	Temperature	Time
1	1	95°C	30 seconds
2	16	95°C	30 seconds
		55°C	1 minute
		68°C	7 minutes (1min per kb)

PCR products were incubated at 37°C for one hour with 1μl of restriction enzyme Dpn1 so as to digest parental DNA. Mutated PCR products were then transformed into dH5α bacterial cells by incubation of 5μl of PCR product with 50μl of bacteria on ice for 30 minutes. Bacteria were heat shocked at 42°C for 30 seconds followed by 2 minutes on ice. 1ml of SOC media was added and bacteria were placed on a 37°C shaker for one hour. Bacteria were spun down, 900μl of media were removed and the bacterial pellet was brought up in the remaining 100μl. Bacteria were plated onto LB-agar plates containing kanamycin and grown at 37°C overnight. Resultant colonies were picked and grown overnight in 3ml of LB media containing kanamycin while shaking at 37°C. 1.5ml of bacterial culture was subjected to EasyPrep lysis so as to recover DNA. DNA concentrations were read using a NanoDrop spectrophotometer (ThermoScientific) and 800ng of DNA product was sent to GeneWiz for sequencing. The primers used for sequencing were as follows:

Forward: 5' GTG AAC ACA GCA GAC CTC CAC 3'

Reverse: 3' GCT CAT TGC GTC AGA GTC ATT 5'

Of the three P^{115A} colonies chosen, two had the correct mutation and P^{115A} #1 was chosen for larger DNA prep (Figure 3A). Four P^{118A} colonies were chosen of which two had the correct mutation and P^{118A} #1 was chosen for larger DNA prep (Figure 3B). Four

P^{121A} colonies were chosen for sequencing and two had the correct mutation with P^{121A} #2 being chosen for larger DNA prep (Figure 3C).

GFP-AFAP1L1 SH3 binding motif mutants were tested for their ability to interact with FLAG tagged cortactin. While equal amounts of cortactin and GFP constructs were expressed (Figure 4, top panel), immunoprecipitation against cortactin and immunoblotting against GFP showed that all AFAP1L1 SH3 binding motif mutants were able to interact with cortactin (Figure 4, bottom panel). GFP-AFAP1 was used as a control and, as previously shown, did not immunoprecipitate with cortactin. Reprobing for cortactin showed that equal amounts of cortactin were pulled down in all AFAP1 and AFAP1L1 immunoprecipitations. The SH3 binding motif of AFAP1L1 is $DLPP^{115}PLP^{118}NKP^{121}$. While P^{115}, P^{118} and P^{121} were chosen due to their localization within a PXXP motif, it is possible that other prolines in the SH3 binding motif may play a role in binding. To further elucidate what amino acids in the SH3 binding motif, if any, are responsible for cortactin binding, amino acids P^{114}, P^{116} or L^{117} could be mutated and tested for their ability to immunoprecipitate cortactin. Also, the entire AFAP1L1 binding motif could be deleted to determine if another region in AFAP1L1 is responsible for cortactin binding.

V. Antibody epitopes and binding

There are currently a limited number of antibodies available for the study of AFAP1L1. A polyclonal antibody created by ProSci Incorporated (1L1-CT) detects an epitope in the C-terminus of AFAP1L1 that corresponds to amino acids 714-727. Two additional polyclonal antibodies from Sigma Aldrich, Ab1 and Ab2, also recognize

AFAP1L1. Ab1 (Sigma C-term) was created using an immunogen corresponding to amino acids 525-659 near the N-terminal portion of AFAP1L1. Ab2 (Sigma N-term) was created using an immunogen corresponding to amino acids 21-159 in the N-terminal portion of AFAP1L1 (Figure 5) (Larkin et al., 2007). The immunogen used to create Sigma N-term encompasses a unique region in the N-terminus of AFAP1L1, a proposed SH2 binding motif and a proposed SH3 binding motif while the immunogen used for Sigma C-term encompasses a unique region in AFAP1L1 after the PH2 domain, the putative leucine zipper and the putative actin binding domain. The epitope recognized by 1L1-CT corresponds to a unique region in the C-terminus of AFAP1L1.

All three antibodies recognize AFAP1L1 at an approximate M_r of 115kDa via western blot in a similar manner in a panel of cell lines with the band corresponding to AFAP1L1 being confirmed through siRNA knockdown. However, each antibody also picks up additional background bands. Antibody 1L1-CT detects a band corresponding to AFAP1L1 and also detects a strong band directly above AFAP1L1 with a M_r of approximately 120kDa in all cell lines tested. 1L1-CT also detects two lower bands with approximate M_r of 90kDa and 75kDa. Sigma C-term detects AFAP1L1, a band directly below with an approximate M_r of 110kDa, a doublet with an approximate M_r of 120kDa and a single band with an approximate M_r of 60kDa. Sigma N-term detects a single band directly above the band corresponding to AFAP1L1 and various lower bands with approximate M_r of 75kDa and 65kDa (Figure 6).

Due to similarity in the AFAP family, it is possible that commercial AFAP1L1 antibodies could cross-react between other AFAP family members . Antibody 1L1-CT was specifically chosen to be created by ProSci Incorporated because of its unique

epitope found in AFAP1L1. At the time of antibody creation, amino acids 714-727 of AFAP1L1 did not show any other possible protein interactions by virtue of a BLAST search against the epitope. 1L1-CT shows some identity and strong similarity to AFAP1 although the epitope is not fully conserved. The epitope is less conserved in AFAP1L2 and therefore is predicted to specifically identify AFAP1L1. Interestingly, the epitope for antibody 1L1-CT is more similar with an N-terminal sequence rather than a C-terminal sequence in AFAP1L2 (Figure 7). A BLAST search identified that Sigma N-term may have the ability to detect all three AFAP family members due to the immunogen covering a conserved SH2 binding motif and SH3 binding motif area in these proteins (Figure 8). Additionally, a BLAST search determined that Sigma C-term antibody may also have the ability to detect all three AFAP family members, again due to the high level of similarity in the protein sequence used to create the antibody (Figure 9).

There are various splice variants of the AFAP family members that may also contribute to the banding pattern seen with AFAP1L1 antibodies. AFAP1 has a similar neural-specific isoform in which an additional 258 base pairs are inserted between amino acids 510 and 511 which is detected at an approximate M_r of 120kDa (Flynn et al., 1995) (Figure 10). AFAP1L1 has a smaller isoform of 725 amino acids predicted by the Ensembl database in which exon 18 has been deleted. While the canonical sequence of AFAP1L1 is predicted to have a molecular weight of approximately 82kDa, the band corresponding to AFAP1L1 is detected at an approximate M_r of 115kDa, most likely due to overall charge on the protein. This is similar to AFAP1 which is predicted to have a molecular weight of 81kDa and is detected at an approximate M_r of 110kDa and AFAP1L2 which is predicted to have a molecular weight of 91kDa and is detected at an

approximate M_r of 130kDa. Exon 18 of AFAP1L1 encodes 43 amino acids; therefore it can be hypothesized that this isoform of AFAP1L1 will run lower than AFAP1L1 although its exact location cannot be identified based solely upon sequence due to the change in overall charge on the protein. Two additional isoforms of each transcript are also predicted by the Ensembl dataset in which the sequence of Exon 1 appears to be altered (Figure 11). AFAP1L2 has four predicted isoforms by the Ensembl database. Canonical AFAP1L2 contains 818 amino acids while a predicted isoform of 814 amino acids lacks four amino acids in the C-terminal unique region of AFAP1L2. A predicted protein coding isoform of 841 amino acid lacks a canonical start codon and contains an additional 28 amino acids inserted between Exons 5 and 6 while lacking the four amino acids in the C-terminal unique region that are also missing in the 814 amino acid isoform. A short 312 amino acid form of AFAP1L2 is predicted by the Ensembl database that contains 18 additional amino acids in Exon 2 and ends at Exon 5 of canonical AFAP1L2, thus encoding only the predicted SH2 and SH3 binding motifs with their surrounding unique sequences (Figure 12).

In addition to other members and splice variants of the AFAP family, AFAP1L1 antibodies may have the ability to interact with other proteins. When created, the epitope for 1L1-CT antibody was run through a human protein BLAST search and returned only AFAP1L1 as an exact protein sequence match. Since the creation of antibody 1L1-CT, an additional human protein which is identical to AFAP1L1 in the C-terminal 444 amino acids and therefore shares the 1L1-CT epitope was discovered. This unnamed protein (Accession BAG64383) contains 625 amino acids and mRNA was isolated from human thymus tissue (Maruyama et al., 2009). This protein is identical to human lysosomal

protein transmembrane 5 in the N-terminal portion of the protein and identical to AFAP1L1 in the C-terminal portion (Figure 13). Apart from protein sequence, there is currently no known functional data for this protein.

Due to the large sequence used to create Sigma N-term and Sigma C-term antibodies, the exact epitope to which the antibodies bind is unknown. It is conceivable that Sigma N- and C-term antibodies may have the ability to bind other proteins outside of the AFAP family.

Although 1L1-CT antibody detects multiple bands in western blot, it specifically identified one single band during immunoprecipitation. Due to the denaturing characteristics of SDS-PAGE, 1L1-CT antibody may have the ability to detect other proteins in their denatured state when looking at whole cell lysates when in fact these epitopes may not be accessible in an intact cell. Immunoprecipitation by 1L1-CT antibody would detect proteins in their native conformation before denaturing conditions and therefore can provide validity to the fact that 1L1-CT antibody can specifically detect AFAP1L1. Currently, immunoprecipitation with Sigma antibodies has not been carried out.

VI. siRNA knockdown of AFAP1L1

Short interfering RNA (siRNA) is a useful tool to transiently decrease the expression of a specific protein through posttranscriptional gene silencing by specific degradation of messenger RNA (mRNA). As a cellular process, long dsRNA is processed through the protein Dicer into duplexes of short nucleotide sequences of approximately 21 base pairs and loaded as single antisense RNA strands into RNA-induced silencing

(RISC) complexes. The single stranded antisense strand of RNA then binds to its complementary sequence in mRNA and causes degradation through cleavage, thus regulating protein expression within the cell. siRNA can be synthetically produced and introduced so as to exploit the cell's machinery and decrease the expression of a protein of choice (Dorsett and Tuschl, 2004; Whitehead et al., 2009).

siRNA knockdown of AFAP1L1 has many benefits. Knockdown of AFAP1L1 expression is useful in western blotting to determine the M_r of AFAP1L1 as siRNA knockdown should deplete AFAP1L1 expression and thus remove the AFAP1L1 signal. Expression of untagged pcDNA3.1-AFAP1L1 in 293T cells resulted in a strong band at an approximate M_r of 115kDa which corresponds to a band seen endogenously in multiple cell lysates. This M_r 115kDa band was not seen after treatment with siRNA specific to AFAP1L1 and was thus confirmed as the M_r of AFAP1L1.

As AFAP1L1 expression has also been shown to play a role in motility structures, siRNA against AFAP1L1 has implications for determining the role of AFAP1L1 in invadosome formation. Preliminary studies show decreased podosome formation in cells lacking AFAP1L1 and further studies are underway to confirm these results.

REFERNCES

(2010). The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Res *38*, D142-148.

Ammer, A.G., and Weed, S.A. (2008). Cortactin branches out: roles in regulating protrusive actin dynamics. Cell Motil Cytoskeleton *65*, 687-707.

Boeckers, T.M. (2006). The postsynaptic density. Cell Tissue Res *326*, 409-422.

Carlin, R.K., Grab, D.J., Cohen, R.S., and Siekevitz, P. (1980). Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. J Cell Biol *86*, 831-845.

Chakrabarti, A., Kelkar, D.A., and Chattopadhyay, A. (2006). Spectrin organization and dynamics: new insights. Biosci Rep *26*, 369-386.

Clump, D.A., Clem, R., Qian, Y., Guappone-Koay, A., Berrebi, A.S., and Flynn, D.C. (2003). Protein expression levels of the Src activating protein AFAP are developmentally regulated in brain. J Neurobiol *54*, 473-485.

Dorsett, Y., and Tuschl, T. (2004). siRNAs: applications in functional genomics and potential as therapeutics. Nat Rev Drug Discov *3*, 318-329.

Flynn, D.C., Koay, T.C., Humphries, C.G., and Guappone, A.C. (1995). AFAP-120. A variant form of the Src SH2/SH3-binding partner AFAP-110 is detected in brain and contains a novel internal sequence which binds to a 67-kDa protein. J Biol Chem *270*, 3894-3899.

Gelkop, S., Babichev, Y., Kalifa, R., Tamir, A., and Isakov, N. (2003). Involvement of crk adapter proteins in regulation of lymphoid cell functions. Immunol Res *28*, 79-91.

Guappone, A.C., and Flynn, D.C. (1997). The integrity of the SH3 binding motif of AFAP-110 is required to facilitate tyrosine phosphorylation by, and stable complex formation with, Src. Mol Cell Biochem *175*, 243-252.

Habas, C. (2010). Functional imaging of the deep cerebellar nuclei: a review. Cerebellum *9*, 22-28.

Han, K., and Kim, E. (2008). Synaptic adhesion molecules and PSD-95. Prog Neurobiol *84*, 263-283.

Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P., and Gasteiger, E. (2009). Infrastructure for the life sciences: design and implementation of the UniProt website. BMC Bioinformatics *10*, 136.

Kalia, L.V., and Salter, M.W. (2003). Interactions between Src family protein tyrosine kinases and PSD-95. Neuropharmacology *45*, 720-728.

Kim, E., and Sheng, M. (2004). PDZ domain proteins of synapses. Nat Rev Neurosci *5*, 771-781.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Maruyama, Y., Wakamatsu, A., Kawamura, Y., Kimura, K., Yamamoto, J., Nishikawa, T., Kisu, Y., Sugano, S., Goshima, N., Isogai, T.*, et al. (2009). Human Gene and Protein Database (HGPD): a novel database presenting a large quantity of experiment-based results in human proteomics. Nucleic Acids Res *37*, D762-766.

McMahon, L.W., Zhang, P., Sridharan, D.M., Lefferts, J.A., and Lambert, M.W. (2009). Knockdown of alphaII spectrin in normal human cells by siRNA leads to chromosomal instability and decreased DNA interstrand cross-link repair. Biochem Biophys Res Commun *381*, 288-293.

Nedrelow, J.H., Cianci, C.D., and Morrow, J.S. (2003). c-Src binds alpha II spectrin's Src homology 3 (SH3) domain and blocks calpain susceptibility by phosphorylating Tyr1176. J Biol Chem *278*, 7735-7741.

Nicolas, G., Fournier, C.M., Galand, C., Malbert-Colas, L., Bournier, O., Kroviarski, Y., Bourgeois, M., Camonis, J.H., Dhermy, D., Grandchamp, B., *et al.* (2002). Tyrosine phosphorylation regulates alpha II spectrin cleavage by calpain. Mol Cell Biol *22*, 3527-3536.

Reichman, C., Singh, K., Liu, Y., Singh, S., Li, H., Fajardo, J.E., Fiser, A., and Birge, R.B. (2005). Transactivation of Abl by the Crk II adapter protein requires a PNAY sequence in the Crk C-terminal SH3 domain. Oncogene *24*, 8187-8199.

Sekino, Y., Kojima, N., and Shirao, T. (2007). Role of actin cytoskeleton in dendritic spine morphogenesis. Neurochem Int *51*, 92-104.

Sparks, A.B., Rider, J.E., Hoffman, N.G., Fowlkes, D.M., Quillam, L.A., and Kay, B.K. (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. Proc Natl Acad Sci U S A *93*, 1540-1544.

Squire, L.B., D.; Bloom, F.; Du Lac, S.; Ghosh A.;Spitzer, N. (2008). Fundamental Neuroscience Third Edition. Elsevier, Inc, Massachusettes.

Tabakoff, B., Saba, L., Printz, M., Flodman, P., Hodgkinson, C., Goldman, D., Koob, G., Richardson, H.N., Kechris, K., Bell, R.L., *et al.* (2009). Genetical genomic determinants of alcohol consumption in rats and humans. BMC Biol *7*, 70.

Takayasu, Y., Iino, M., Takatsuru, Y., Tanaka, K., and Ozawa, S. (2009). Functions of glutamate transporters in cerebellar Purkinje cell synapses. Acta Physiol (Oxf) *197*, 1-12.

Tanaka, K. (2000). Functions of glutamate transporters in the brain. Neurosci Res *37*, 15-19.

Voss, M., Lettau, M., and Janssen, O. (2009). Identification of SH3 domain interaction partners of human FasL (CD178) by phage display screening. BMC Immunol *10*, 53.

Walikonis, R.S., Jensen, O.N., Mann, M., Provance, D.W., Jr., Mercer, J.A., and Kennedy, M.B. (2000). Identification of proteins in the postsynaptic density fraction by mass spectrometry. J Neurosci *20*, 4069-4080.

Weed, S.A., and Parsons, J.T. (2001). Cortactin: coupling membrane dynamics to cortical actin assembly. Oncogene *20*, 6418-6434.

Whitehead, K.A., Langer, R., and Anderson, D.G. (2009). Knocking down barriers: advances in siRNA delivery. Nat Rev Drug Discov *8*, 129-138.

Figure 1A

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
A	Amphiphysin		LCK		SPCN		Cortactin		MLPK3		Yes1		Abl2		SJHUA		Itk		CRK-D2		pos	
B	Dlg2		EMP55		FGR		SLK		Nebulin		c-Src		FYB-D1		Hck		VAV2-D2		NOF2-D1		pos	
C	VAV-D1		NCK1-D3		Y124		PEXD		BTK		RasGAP		PSD95		Tim		HS1		Stam		pos	
D	BLK		Abl		PLC-γ		Riz		PI3-β		ITSN-D1		ITSN-D2		TXK		GST				pos	
E	pos		pos		pos		pos		pos		pos		pos		pos		pos		pos		pos	

SH3 domain list • SH3 Domain Array I

POSITION	DOMAIN	FULL NAME
A1, 2	Amphiphysin	Amphiphysin
B1, 2	Dlg2	Discs large homolog 2
C1, 2	VAV-D1	VAV proto-oncogene, SH3 domain #1
D1, 2	BLK	Beta-lymphocyte specific protein tyrosine kinase
A3, 4	LCK	Human T-lymphocyte specific protein tyrosine kinasep56
LCK		
B3, 4	EMP55	55 kDa erythrocyte membrane protein
C3, 4	NCK1-D3	Cytoplasmic protein NCK1, SH3 domain #3
D3, 4	Abl	Abelson tyrosine kinase
A5, 6	SPCN	Spectrin alpha chain (non-erythrocytic)
B5, 6	FGR	Cellular Gardner-Rasheed feline sarcoma virusprotein
C5, 6	Y124	PAK-interacting exchange factor beta
D5, 6	PLCg	Phospholipase C gamma-1
A7, 8	Cortactin	Cortactin
B7, 8	SLK	Proto-oncogene tyrosine protein kinase FYN
C7, 8	PEXD	Peroxisomal membrane protein PEX13
D7, 8	Riz	Retinoblastoma protein-interacting zinc-finger
A9, 10	MLPK3	Mixed-lineage kinase 3
B9, 10	Nebulin	Nebulin
C9, 10	BTK	Bruton Tyrosine Kinase
D9, 10	PI3b	Phosphoinositide-3-kinase regulatory beta subunit
A11, 12	Yes1	Yamaguchi sarcoma virus oncogene homolog 1
B11, 12	c-Src*	Cellular Rous Sarcoma viral oncogene homolog
C11, 12	RasGAP	Ras GTPase-activating protein 1
D11, 12	ITSN-D1	Intersectin, SH3 Domain #1
A13, 14	Abl2 **	Abelson-related protein; Arg
B13, 14	FYB-D1	Fyn binding protein, SH3 domain #1
C13, 14	PSD95	Presynaptic density protein 95
D13, 14	ITSN-D2	Intersectin 1, SH3 Domain #2
A15, 16	SJHUA	Spectrin alpha chain, erythrocyte
B15, 16	Hck	Hemopoietic cell kinase
C15, 16	Tim	Rho guanine nucleotide exchange factor (GEF) 5
D15, 16	TXK	Tyrosine-protein kinase TXK
A17, 18	Itk	Interleukin-2-inducible T-cell kinase
B17, 18	VAV2-D2	Vav2 oncogene product, SH3 Domain #2
C17, 18	HS1	Hematopoietic specific protein 1
D17, 18	control	negative control
A19, 20	CRK-D2	Avian sarcoma virus CT10 oncogene homolog, domain #2
B19, 20	NOF2-D1	Neurite outgrowth factor or neutrophil cytosol factor 2 domain #1
C19, 20	Stam	Signal transducing adaptor molecule

Figure 1B
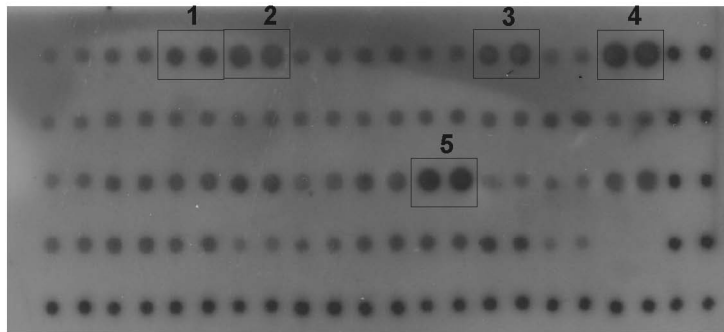
Figure 1. A Panomics TranSignal SH3 Domain Array I predicted AFAP1L1 SH3 domain binding partners

(A) The SH3 domain layout of a TranSignal SH3 Domain Array I from Panomics, Inc. was shown as a guide to determine possible SH3 domain binding partners of a biotinylated AFAP1L1 SH3 binding motif peptide.

(B) Five duplicate SH3 domain spots were chosen as potential sites for AFAP1L1 binding based upon their chemiluminescence. 1 = Spectrin alpha chain (non-erythrocytic), 2 = Cortactin, 3 = Spectrin alpha chain (erythrocytic), 4 = Avian sarcoma virus CT10 oncogene homolog domain #2, and 5 = Post synaptic density protein 95.

Figure 2A
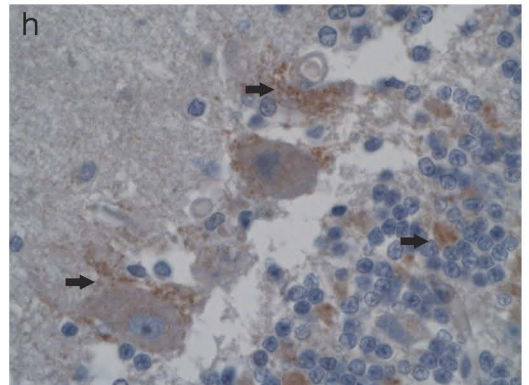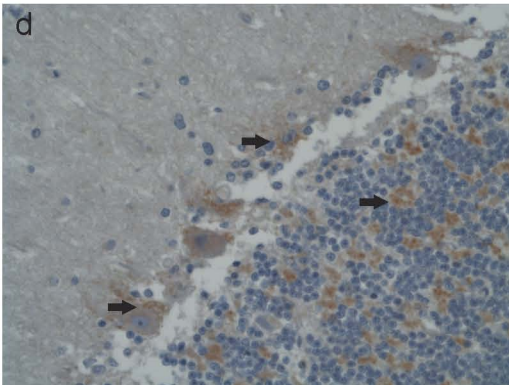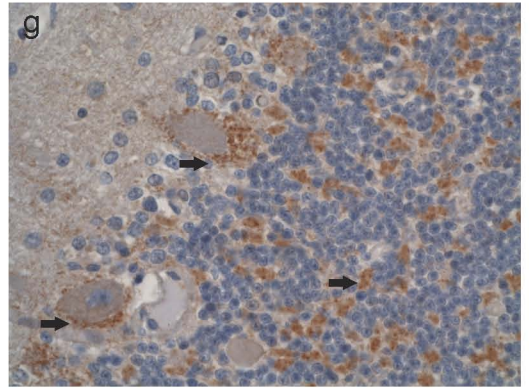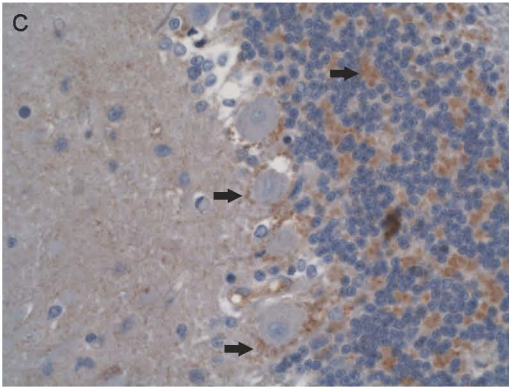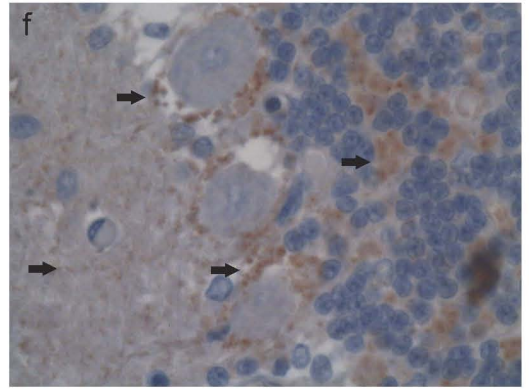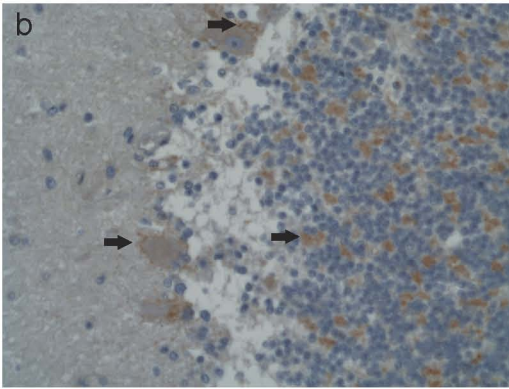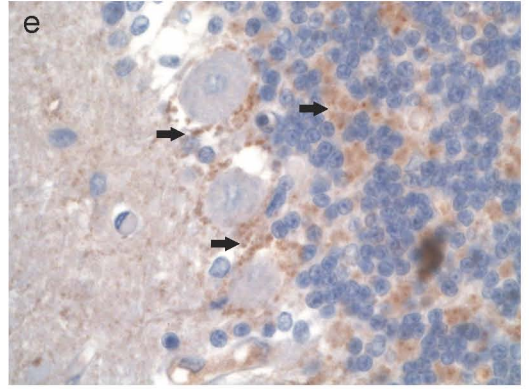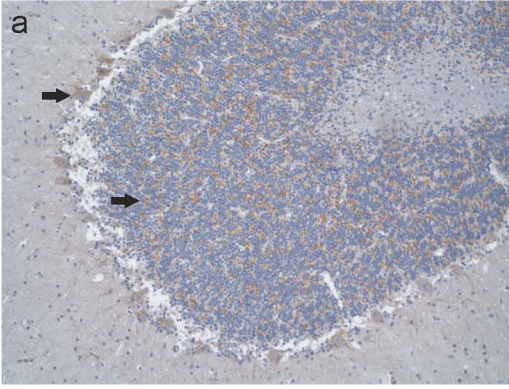
Figure 2B
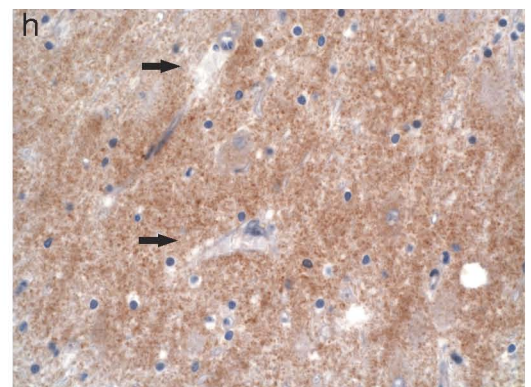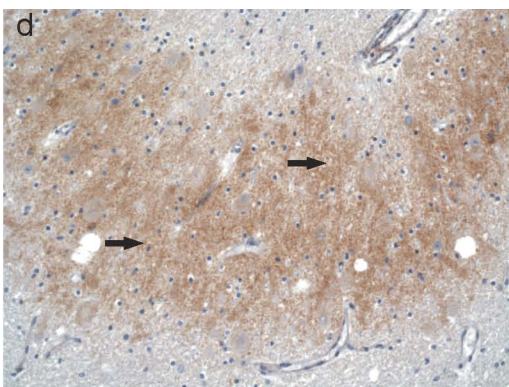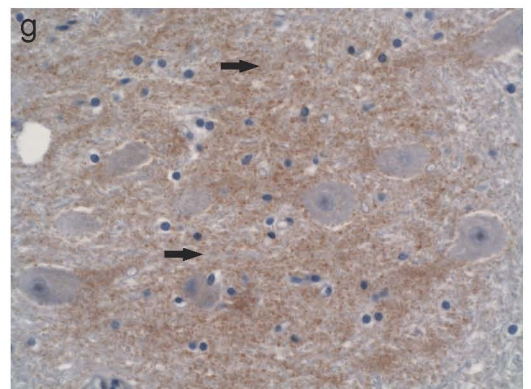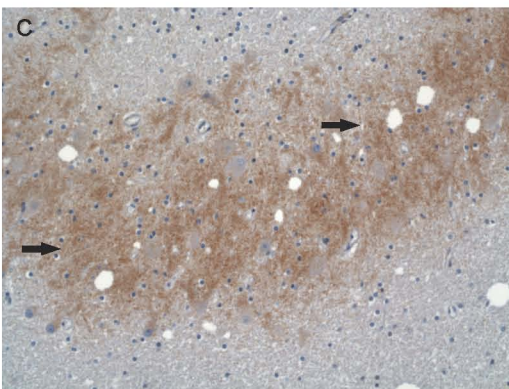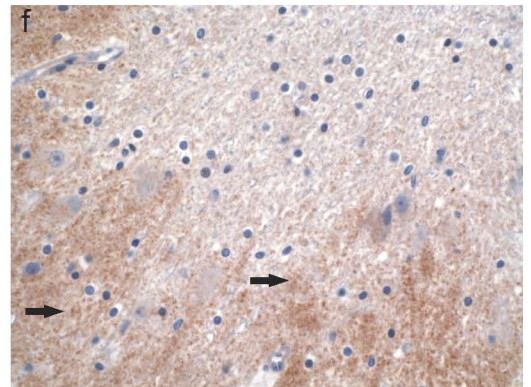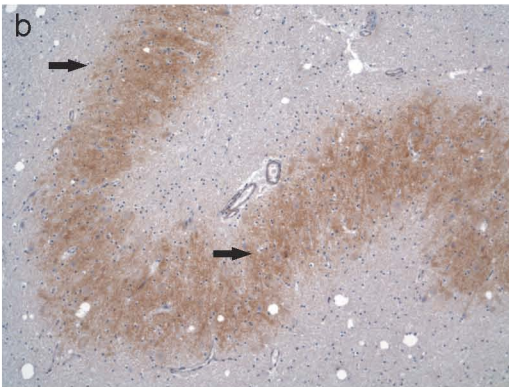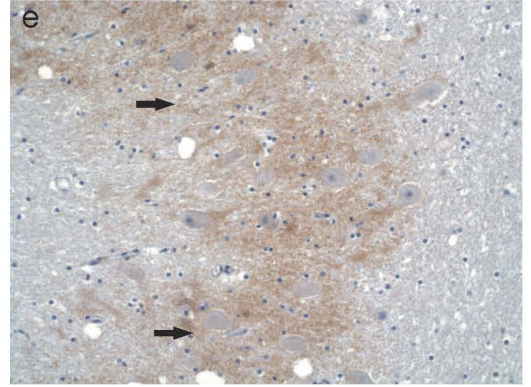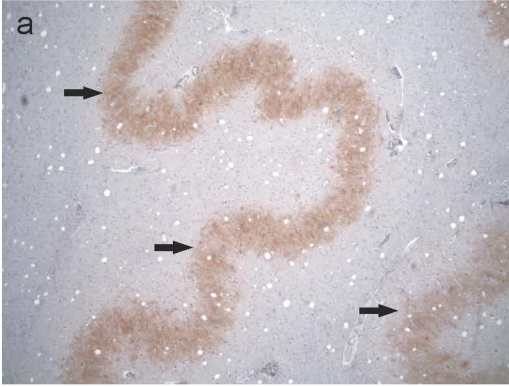
Figure 2. Immunohistochemical staining of AFAP1L1 in human cerebellum and dentate nucleus

(A) AFAP1L1 immunohistochemical signal in human cerebellum shows distinct staining extending from the molecular layer (panels e-f), surrounding the Purkinje neurons (panels a-h) and in distinct locations around the granule cells of the granular layer (panels g-h).

(B) AFAP1L1 immunohistochemical signal in the human dentate nucleus shows distinct signal around the cells of the dentate nucleus.

Figure 3

A.
```
AFAP1L1   CTAGCCAAGAGCCCACGCCTGAGAAACGCGGCCGACCTGCCTCCACCGCTCCCCAACAAG 360
P115A.1   CTAGCCAAGAGCCCACGCCTGAGAAACGCGGCCGACCTGCCTGCACCGCTCCCCAACAAG 162
          ********************************************* ****************
```

B.
```
AFAP1L1   CTAGCCAAGAGCCCACGCCTGAGAAACGCGGCCGACCTGCCTCCACCGCTCCCCAACAAG 360
P118A.1   CTAGCCAAGAGCCCACGCCTGAGAAACGCGGCCGACCTGCCTCCACCGCTCGCCAACAAG 163
          ************************************************** *******
```

C.
```
AFAP1L1   CCTCCCCCTGAGGACTACTATGAAGAGGCCCTTCCTCTGGGACCCGGCAAGTCGCCTGAG 420
P121A.2   GCTCCCCCTGAGGACTACTATGAAGAGGCCCTTCCTCTGGGACCCGGCAAGTCGCCTGAG 224
          ***********************************************************
```

Figure 3. Mutation of the SH3 binding motif of AFAP1L1

Key prolines in the SH3 binding motif of AFAP1L1 were mutated to alanines so as to test the ability of the alanine mutation to abrogate AFAP1L1 binding to cortactin. (A) P[115A], (B) P[118A], and (C) P[121A], were mutated to alanine by site-directed mutagenesis.

Figure 4
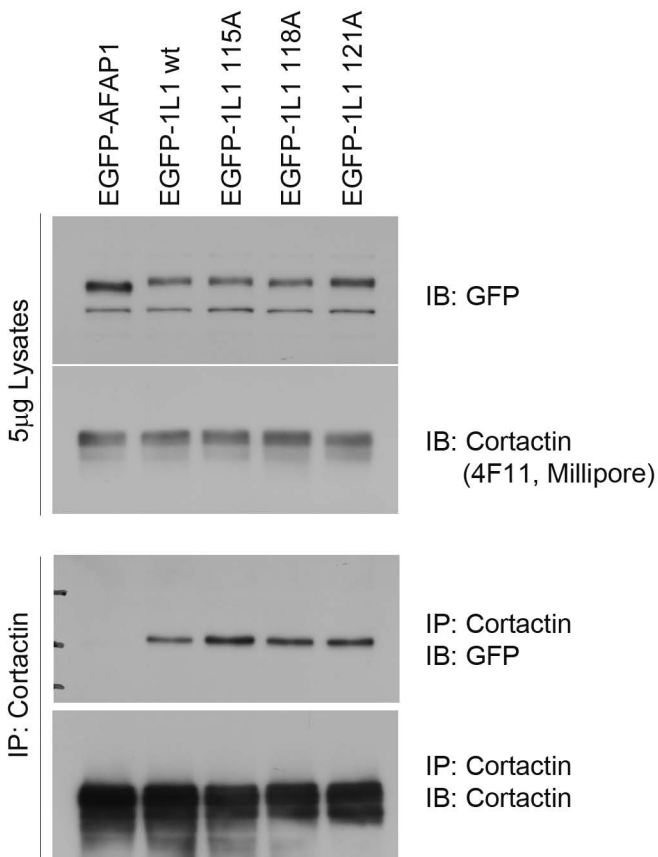
Figure 4. Immunoprecipitation of cortactin with AFAP1L1 SH3 binding motif mutants

GFP-AFAP1 and GFP-AFAP1L1, as well as three AFAP1L1 SH3 binding motif mutants GFP-AFAP1L1^{P115A}, GFP-AFAP1L1^{P118A} and GFP-AFAP1L1^{P121A}, were overexpressed with FLAG-cortactin and immunoprecipitated against cortactin. Lysates were resolved by SDS-PAGE and immunoblotted for GFP. Reprobe against the GFP tag confirmed equal immunoprecipitation of all constructs (bottom panel). 5µg of each lysate was resolved by SDS-PAGE and immunoblotted for both GFP and cortactin to confirm that equal amounts of constructs were expressed in each lysate (top panel).

Figure 5

```
AFAP1L1    MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
SigmaNterm -------------------LDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 40
                              ****************************************

AFAP1L1    DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 120
SigmaNterm DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 100
           ************************************************************

AFAP1L1    PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 180
SigmaNterm PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYY-------------------- 139
           **************************************

AFAP1L1    DSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAKQL 240

AFAP1L1    TVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLALQS 300

AFAP1L1    REQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNC 360

AFAP1L1    STLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVP 420

AFAP1L1    CCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAF 480

AFAP1L1    RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 540
SigmaCterm -----------------------------------------VDVETLTSIVSAGRNS 16
                                                    ***************

AFAP1L1    FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 600
SigmaCterm FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 76
           ************************************************************

AFAP1L1    HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 660
SigmaCterm HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPG- 135
           **********************************************************

AFAP1L1    KLKALERAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSGET 720
1L1-CT     --------------------------------------------------KPKSGET 7
                                                             *******

AFAP1L1    ANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAKEWEMKKT 768
1L1-CT     ANKPQNS----------------------------------------- 14
           *******
```

Figure 5. AFAP1L1 antibodies aligned with AFAP1L1 amino acid sequence

Sigma N-term, Sigma C-term and 1L1-CT antibodies were aligned to full length AFAP1L1 amino acid sequence using ClustalW2 analysis (27) so as to determine where in the sequence antibodies bound and what conserved binding motifs these antibodies may cover. The proposed SH3 domain is highlighted in green, proposed SH2 domains are highlighted in pink, the PH1 and PH2 domains are highlighted in light blue, the substrate domain is highlighted in yellow, the proposed leucine zipper is highlighted in red and the proposed actin binding domain in highlighted in dark blue.

Figure 6
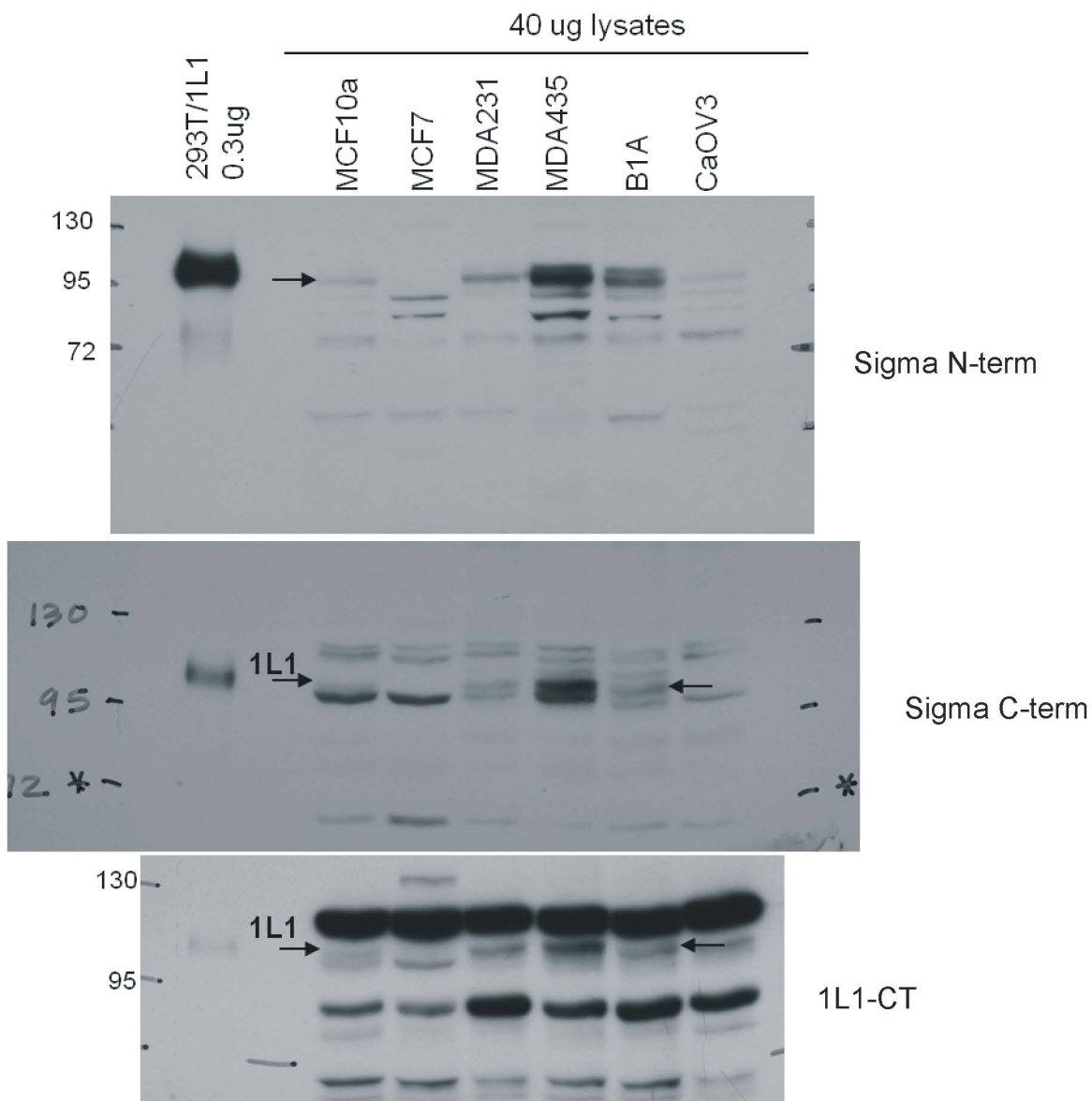
Figure 6. AFAP1L1 antibody specificity

A panel of cell lines was tested against all three AFAP1L1 antibodies to test for specificity. 293T cells overexpressing an untagged pcDNA3.1 *afap1l1* construct were used as a control to determine AFAP1L1 M_r. Cell lines tested were breast lines MCF10a, MCF7, MDA-MB-231, MDA-MB-435, B1A (a knockdown of AFAP1 in MDA-MB-231 cells), and CaOv3, an ovarian cancer cell line. Sigma N-term (A), Sigma C-term (B) and 1L1-CT (C) all recognize AFAP1L1 to differing extents and detect various background bands.

Figure 7

```
AFAP1L1          KPKSGETANKPQNS 727
1L1-CT           KPKSGETANKPQNS 14
Consensus        KPKSGETANKPQNS


AFAP1            EPKSGTSSPQSPVFRHRTLENSPISSCDTSDTEGPVPVNSAA 683
1L1-CT           KPKSGETA------------NKPQNS---------------- 14
Consensus        *KPKS **            N*P *S



AFAP1L2          APEEQGLLPNGEPSQHSSAPQKSLPDLPPPKMIPERKQLAIPKTESP 120
1L1-CT           KPK------SGETANK---PQNS----------------------- 14
Consensus         P*       *GE****   PQ*S
```

Figure 7. 1L1-CT binding across AFAP family members

The amino acid sequence against which antibody 1L1-CT was raised was compared

across AFAP family members. Identical amino acid sequence is shown by its one letter

abbreviation in the consensus sequence while similar amino acid sequence is represented

by *.

Figure 8

```
AFAP1L1      MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
Sigma.N-term -------------------LDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 40
Consensus                       LDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA

AFAP1L1      DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 120
Sigma.N-term DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 100
Consensus    DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK

AFAP1L1      PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 180
Sigma.N-term PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYY-------------------- 139
Consensus    PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYY


AFAP1        MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQETANSLP 60
Sigma.N-term --------------LDHEYLSDTTLEKKMAVASILQSLQ--------------------- 25
Consensus                  LDHEYL**T* EKK ****IL  *Q

AFAP1        APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
Sigma.N-term --------------------PLPAKEVSYLYVNTADLHSGPSFVESLFEEFDCDLS-- 61
Consensus                        PLP*****  Y * A    S   *  E * ***D*D

AFAP1        SYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 180
Sigma.N-term --DLRDMPEDDGE--------PSKGASPELAKSPRLR---------------------- 88
Consensus      *  D  E*DG*         PS* AS *L*K****

AFAP1        LLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 240
Sigma.N-term ------NAADLP------------------------------------------------ 94
Conensus            *  * P

AFAP1        LKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVVENGITTCNGKE 300
Sigma.N-term -----------------PPLPNKPPPEDYYEEALP------LGPGKSPEYISSHNG-- 127
Consensus                     PP P**P  *   E* L*        G G   * I** NG

AFAP1        QVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVLSNS 360
Sigma.N-term ----CSPSHSIVDGYY-------------------------------------------- 139
Consensus        **S*S **G


AFAP1L2      MERYKALEQLLTELDDFLKILDQENLSSTALVKKSCLAELLRLYTKSSSSDEEYIYMNKV 60
Sigma.N-term -------------------LDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 40
Consensus                       LD*E LS*T*L KK **A**L*      **** *Y*Y*N**

AFAP1L2      TINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPPPKMIPERKQLAIPKTESP 120
Sigma.N-term DLHS-------------------------------------------------------- 44
Consensus     ***

AFAP1L2      EGYYEEAEPYDTSLNEDGEAVSSSYESYDEEDGSKGKSAPYQWPSPEAGIELMRDARICA 180
Sigma.N-term ------GPSFVESLFEEFDCDLSDLRDMPEDDGE---------PSKGASPELAKSPRLR- 88
Consensus          * **   SL E* **  S* **  E*EG*         PS  A* EL ***R*

AFAP1L2      FLWRKKWLGQWAKQLCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLK 240
Sigma.N-term ------------------------------------------------------------
Consensus

AFAP1L2      ITPMNADVIVLGLQSKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIAEK 300
Sigma.N-term ----NAADLPPPLPNKPPPEDYYEEALPLG-------------------PGKSPEYISS 124
Consensus         NA  *   L *K  *E** *    **                   **P*  **

AFAP1L2      YLSASEYGSSVDGHPEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSSYL 360
Sigma.N-term HNGCSPSHSIVDGYY-------------------------------------------- 139
Consensus     * **S   S VDG*
```

Figure 8. Sigma N-term binding across AFAP family members

The amino acid sequence against which Sigma N-term antibody was raised was compared across AFAP family members. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 9

```
AFAP1L1      RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 540
Sigma.C-term ---------------------------------------------VDVETLTSIVSAGRNS 16
Consensus                                                 VDVETLTSIVSAGRNS

AFAP1L1      FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 600
Sigma.C-term FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 76
Consensus    FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCDEKSHRVDPQVKVKR

AFAP1L1      HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 660
Sigma.C-term HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPG- 135
Consensus    HASSANQPKQGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPG
```

```
AFAP1        LEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQTAKQTFCFMNRRVISA 480
Sigma.C-term --------------------------------VDVETLTSIVSAGRNSFLYAR------ 21
Consensus                               *DVE  *S********F * *

AFAP1        NPYLGGTSNGYAHPSGTALHYDDVPCINGSLKGKKPPVASNGVTGKGKTLSSQPKKADPA 540
Sigma.C-term -----SCQNQWPEPR----VYDDVPYEKMQDEEPERPTGAQ-VKRHASSCSEKSHRVDPQ 71
Consensus         * *N ***P      YDDVP  * * *  * P**** V* **** S******DP

AFAP1        AVVKRTGSNAAQYKYGKNRVEADAKRLQTKEEELLKRKEALRNRLAQLRKERKDLRAAIE 600
Sigma.C-term VKVKRHASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIR 131
Consensus    * VKR *S*A QYLYGKNR*E DA*R   ***E*L K*KE**R**L  LR*E***L* AI*

AFAP1        VNAGRKPQAILEEKLKQLEEECRQKEAERVSLELELTEVKESLKKALAGGVTLGLAIEPK 660
Sigma.C-term SSPG-------------------------------------------------------- 135
Consensus      **G
```

```
AFAP1L2      GEELAKLEAKSSEEMGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVSAAKNSLLLMQ 480
Sigma.C-term -------------------------------------VDVETLTSIVSAGRNSFLYAR 21
Consensus                                          VD** ***IVSA**NS*L  *

AFAP1L2      RKFSEPNTYIDGLPSQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNERESDRVYLDL 540
Sigma.C-term ---SCQNQWPE-------PRVYDDVPYEKMQ-------------DEEPER--------- 48
Consensus       S  N * *         **YDDV  ***               D*  ER

AFAP1L2      TPVKSFLHGPSSAQAQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCPENLGEQQLES 600
Sigma.C-term ---------PTGAQVKRHAS--SCSEKS-----------------------------HR 67
Consensus             P**AQ**  **  SC ***                            *

AFAP1L2      LEPEDPSLRITTVKIQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVTSAEIKLGKNR 660
Sigma.C-term VDP--------QVKVKRHAS--------------------------SANQYKYGKNR 90
Consensus    **P        VK** * *                          ** * K GKNR

AFAP1L2      TEAEVKRYTEEKERLEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLASLEQKLKEID 720
Sigma.C-term AEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPG-------------- 135
Consensus    *E ***RY  EKE*LEK*KE IR *L  LR*EKRELKE** ****
```

Figure 9. Sigma C-term binding across AFAP family members

The amino acid sequence against which Sigma C-term antibody was raised was compared across AFAP family members. Identical amino acid sequence is shown by its one letter abbreviation in the consensus sequence while similar amino acid sequence is represented by *.

Figure 10

```
AFAP1      MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQETANSLP 60
AFAP120    MEELIVELRLFLELLDHEYLTSTVREKKAVITNILLRIQSSKGFDVKDHAQKQETANSLP 60
           ************************************************************

AFAP1      APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
AFAP120    APPQMPLPEIPQPWLPPDSGPPPLPTSSLPEGYYEEAVPLSPGKAPEYITSNYDSDAMSS 120
           ************************************************************

AFAP1      SYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 180
AFAP120    SYESYDEEEEDGKGKKTRHQWPSEEASMDLVKDAKICAFLLRKKRFGQWTKLLCVIKDTK 180
           ************************************************************

AFAP1      LLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 240
AFAP120    LLCYKSSKDQQPQMELPLQGCNITYIPKDSKKKKHELKITQQGTDPLVLAVQSKEQAEQW 240
           ************************************************************

AFAP1      LKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVVENGITTCNGKE 300
AFAP120    LKVIKEAYSGCSGPVDSECPPPPSSPVHKAELEKKLSSERPSSDGEGVVENGITTCNGKE 300
           ************************************************************

AFAP1      QVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVLSNS 360
AFAP120    QVKRKKSSKSEAKGTVSKVTGKKITKIISLGKKKPSTDEQTSSAEEDVPTCGYLNVLSNS 360
           ************************************************************

AFAP1      RWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSKHPLTFRLLRNGQEVAV 420
AFAP120    RWRERWCRVKDNKLIFHKDRTDLKTHIVSIPLRGCEVIPGLDSKHPLTFRLLRNGQEVAV 420
           ************************************************************

AFAP1      LEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQTAKQTFCFMNRRVISA 480
AFAP120    LEASSSEDMGRWIGILLAETGSSTDPEALHYDYIDVEMSASVIQTAKQTFCFMNRRVISA 480
           ************************************************************

AFAP1      NPYLGGTSNGYAHPSGTALHYDDVPCING------------------------------ 509
AFAP120    NPYLGGTSNGYAHPSGTALHYDDVPCINGSWEPEDGFPASCSRGLGEEVLYDNAGLYDNL 540
           *****************************

AFAP1      --------------------------------------------------SLKGKKP 516
AFAP120    PPPHIFARYSPADRKASRLSADKLSSNHYKYPASAQSVTNTSSVGRASLGLNSQLKGKKP 600
                                                             :*******

AFAP1      PVASNGVTGKGKTLSSQPKKADPAAVVKRTGSNAAQYKYGKNRVEADAKRLQTKEEELLK 576
AFAP120    PVASNGVTGKGKTLSSQPKKADPAAVVKRTGSNAAQYKYGKNRVEADAKRLQTKEEELLK 660
           ************************************************************

AFAP1      RKEALRNRLAQLRKERKDLRAAIEVNAGRKPQAILEEKLKQLEEECRQKEAERVSLELEL 636
AFAP120    RKEALRNRLAQLRKERKDLRAAIEVNAGRKPQAILEEKLKQLEEECRQKEAERVSLELEL 720
           ************************************************************

AFAP1      TEVKESLKKALAGGVTLGLAIEPKSGTSSPQSPVFRHRTLENSPISSCDTSDTEGPVPVN 696
AFAP120    TEVKESLKKALAGGVTLGLAIEPKSGTSSPQSPVFRHRTLENSPISSCDTSDTEGPVPVN 780
           ************************************************************

AFAP1      SAAVLKKSQAAPGSSPCRGHVLRKAKEWELKNGT 730
AFAP120    SAAVLKKSQAAPGSSPCRGHVLRKAKEWELKNGT 814
           *********************************
```

Figure 10. AFAP1 isoforms

AFAP1 canonical sequence was compared to a second neural specific AFAP1 isoform, AFAP120 using ClustalW2 analysis (27). Identical sequence is represented by * while similar sequence is represented by : .

Figure 11

```
AFAP1L1       MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
AFAP1L1.767   -ERGKMLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 59
AFAP1L1.725   MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 60
AFAP1L1.724   -ERGKMLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTA 59
              :**::*******************************************************


AFAP1L1       DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 120
AFAP1L1.767   DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 119
AFAP1L1.725   DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 120
AFAP1L1.724   DLHSGPSFVESLFEEFDCDLSDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNK 119
              ************************************************************


AFAP1L1       PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 180
AFAP1L1.767   PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 179
AFAP1L1.725   PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 180
AFAP1L1.724   PPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYEDADSSYPATRVNGELKSSYN 179
              ************************************************************


AFAP1L1       DSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAKQL 240
AFAP1L1.767   DSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAKQL 239
AFAP1L1.725   DSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAKQL 240
AFAP1L1.724   DSDAMSSSYESYDEEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRFGQWAKQL 239
              ************************************************************


AFAP1L1       TVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLALQS 300
AFAP1L1.767   TVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLALQS 299
AFAP1L1.725   TVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLALQS 300
AFAP1L1.724   TVIREDQLLCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKRHELRFTQGATEVLVLALQS 299
              ************************************************************


AFAP1L1       REQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNC 360
AFAP1L1.767   REQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNC 359
AFAP1L1.725   REQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNC 360
AFAP1L1.724   REQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNC 359
              ************************************************************


AFAP1L1       STLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVP 420
AFAP1L1.767   STLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVP 419
AFAP1L1.725   STLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVP 420
AFAP1L1.724   STLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVP 419
              ************************************************************


AFAP1L1       CCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAF 480
AFAP1L1.767   CCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAF 479
AFAP1L1.725   CCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAF 480
AFAP1L1.724   CCGYLNVLVNQGWKERWCRLKCNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAF 479
              ************************************************************


AFAP1L1       RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 540
AFAP1L1.767   RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 539
AFAP1L1.725   RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 540
AFAP1L1.724   RILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALHYDYVDVETLTSIVSAGRNS 539
              ************************************************************


AFAP1L1       FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 600
AFAP1L1.767   FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 599
AFAP1L1.725   FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 600
AFAP1L1.724   FLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKR 599
              ************************************************************
```

```
AFAP1L1         HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 660
AFAP1L1.767     HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 659
AFAP1L1.725     HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 660
AFAP1L1.724     HASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGA 659
                ************************************************************

AFAP1L1         KLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSGET 720
AFAP1L1.767     KLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSGET 719
AFAP1L1.725     KLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSG-- 718
AFAP1L1.724     KLKALEEAVATLEAQCRAKEERRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSG-- 717
                *********************************************************

AFAP1L1         ANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAKEWEMKKT 768
AFAP1L1.767     ANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAKEWEMKKT 767
AFAP1L1.725     ----------------------------------------EWEMKKT 725
AFAP1L1.724     ----------------------------------------EWEMKKT 724
                                                        *******
```

Figure 11. AFAP1L1 isoforms

Canonical AFAP1L1 amino acid sequence was compared using ClustalW2 analysis (27)

to three other AFAP1L1 isoforms. AFAP1L1 containing 767 amino acids has an altered

Exon 1, AFAP1L1 containing 725 amino acids does not contain Exon 18 and AFAP1L1

containing 724 amino acids does not contain Exon 18 and has an altered Exon 1. Identical

sequence is represented by * while similar sequence is represented by :.

Figure 12

```
AFAP1L2       MERYKA-----------------LEQLLTELDDFLKILDQENLSSTALVKKSCLAELLR 42
AFAP1L2.814   MERYKA-----------------LEQLLTELDDFLKILDQENLSSTALVKKSCLAELLR 42
AFAP1L2.841   -QRVRA-----------------LEQLLTELDDFLKILDQENLSSTALVKKSCLAELLR 41
AFAP1L2.312   MERYKAQGCCCLVVQRRIRQVSASLEQLLTELDDFLKILDQENLSSTALVKKSCLAELLR 60
               :* :*              ********************************

AFAP1L2       LYTKSSSSDEEYIYMNKVTINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPP 102
AFAP1L2.814   LYTKSSSSDEEYIYMNKVTINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPP 102
AFAP1L2.841   LYTKSSSSDEEYIYMNKVTINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPP 101
AFAP1L2.312   LYTKSSSSDEEYIYMNKVTINKQQNAESQGKAPEEQGLLPNGEPSQHSSAPQKSLPDLPP 120
              ***********************************************************

AFAP1L2       PKMIPERKQLAIPKTESPEGYYEEAEPYDTSLN------------------------- 135
AFAP1L2.814   PKMIPERKQLAIPKTESPEGYYEEAEPYDTSLN------------------------- 135
AFAP1L2.841   PKMIPERKQLAIPKTESPEGYYEEAEPYDTSLNGHSGGFLPTGVPRWVQVPEGVIYATIT 161
AFAP1L2.312   PKMIPERKQLAIPKTESPEGYYEEAEPYDTSLN------------------------- 153
              ********************************

AFAP1L2       -EDGEAVSSSYESYDEEDGSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQWAKQ 194
AFAP1L2.814   -EDGEAVSSSYESYDEEDGSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQWAKQ 194
AFAP1L2.841   LEDGEAVSSSYESYDEEDGSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQWAKQ 221
AFAP1L2.312   -EDGEAVSSSYESYDEEDGSKGKSAPYQWPSPEAGIELMRDARICAFLWRKKWLGQWAKQ 212
               *********************************************************

AFAP1L2       LCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVLGLQ 254
AFAP1L2.814   LCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVLGLQ 254
AFAP1L2.841   LCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVLGLQ 281
AFAP1L2.312   LCVIKDNRLLCYKSSKDHSPQLDVNLLGSSVIHKEKQVRKKEHKLKITPMNADVIVLGLQ 272
              ***********************************************************

AFAP1L2       SKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIAEKYLSASEYGSSVDGH 314
AFAP1L2.814   SKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIAEKYLSASEYGSSVDGH 314
AFAP1L2.841   SKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQKPDIAEKYLSASEYGSSVDGH 341
AFAP1L2.312   SKDQAEQWLRVIQEVSGLPSEGASEGNQYTPDAQRFNCQK------------------- 312
              ****************************************

AFAP1L2       PEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSSYLNVLVNSQWKSRWCS 374
AFAP1L2.814   PEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSSYLNVLVNSQWKSRWCS 374
AFAP1L2.841   PEVPETKDVKKKCSAGLKLSNLMNLGRKKSTSLEPVERSLETSSYLNVLVNSQWKSRWCS 401
AFAP1L2.312   -----------------------------------------------------------

AFAP1L2       VRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHLYSFRILHKGEELAKLEAKSSEE 434
AFAP1L2.814   VRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHLYSFRILHKGEELAKLEAKSSEE 434
AFAP1L2.841   VRDNHLHFYQDRNRSKVAQQPLSLVGCEVVPDPSPDHLYSFRILHKGEELAKLEAKSSEE 461
AFAP1L2.312   -----------------------------------------------------------

AFAP1L2       MGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVSAAKNSLLLMQRKFSEPNTYIDGLP 494
AFAP1L2.814   MGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVSAAKNSLLLMQRKFSEPNTYIDGLP 494
AFAP1L2.841   MGHWLGLLLSESGSKTDPEEFTYDYVDADRVSCIVSAAKNSLLLMQRKFSEPNTYIDGLP 521
AFAP1L2.312   -----------------------------------------------------------

AFAP1L2       SQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNERESDRVYLDLTPVKSFLHGPSSAQ 554
AFAP1L2.814   SQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNERESDRVYLDLTPVKSFLHGPSSAQ 554
AFAP1L2.841   SQDRQEELYDDVDLSELTAAVEPTEEATPVADDPNERESDRVYLDLTPVKSFLHGPSSAQ 581
AFAP1L2.312   -----------------------------------------------------------
```

```
AFAP1L2        AQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCPENLGEQQLESLEPEDPSLRITTVK 614
AFAP1L2.814    AQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCPENLGEQQLESLEPEDPSLRITTVK 614
AFAP1L2.841    AQASSPTLSCLDNATEALPADSGPGPTPDEPCIKCPENLGEQQLESLEPEDPSLRITTVK 641
AFAP1L2.312    ------------------------------------------------------------


AFAP1L2        IQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVTSAEIKLGKNRTEAEVKRYTEEKER 674
AFAP1L2.814    IQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVTSAEIKLGKNRTEAEVKRYTEEKER 674
AFAP1L2.841    IQTEQQRISFPPSCPDAVVATPPGASPPVKDRLRVTSAEIKLGKNRTEAEVKRYTEEKER 701
AFAP1L2.312    ------------------------------------------------------------


AFAP1L2        LEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLASLEQKLKEIDEECRGEESRRVDLE 734
AFAP1L2.814    LEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLASLEQKLKEIDEECRGEESRRVDLE 734
AFAP1L2.841    LEKKKEEIRGHLAQLRKEKRELKETLLKCTDKEVLASLEQKLKEIDEECRGEESRRVDLE 761
AFAP1L2.312    ------------------------------------------------------------


AFAP1L2        LSIMEVKDNLKKAEAGPVTLGTTVDTTHLENVSPRPKAVTPASAPDCTPVNSATTLKNRP 794
AFAP1L2.814    LSIMEVKDNLKKAEAGPVTLGTTVDTTHLEN----PKAVTPASAPDCTPVNSATTLKNRP 790
AFAP1L2.841    LSIMEVKDNLKKAEAGPVTLGTTVDTTHLEN----PKAVTPASAPDCTPVNSATTLKNRP 817
AFAP1L2.312    ------------------------------------------------------------


AFAP1L2        LSVVVTGKGTVLQKAKEWEKKGAS 818
AFAP1L2.814    LSVVVTGKGTVLQKAKEWEKKGAS 814
AFAP1L2.841    LSVVVTGKGTVLQKAKEWEKKGAS 841
AFAP1L2.312    -----------------------
```

Figure 12. AFAP1L2 isoforms

Canonical AFAP1L2 amino acid sequence was compared by ClustalW2 analysis (27) to three other AFAP1L2 predicted isoforms. AFAP1L2 containing 814 amino acids lacks four amino acids in the C-terminal region. AFAP1L2 containing 841 amino acids has a undefined start codon, lacks the four amino acids in the C-terminal region as in AFAP1L2 containing 814 amino acids and has an additional 28 amino acids inserted between Exons 5 and 6. AFAP1L2 containing 312 amino acids contains 18 additional amino acids in Exon 2 and ends at Exon 5 of canonical AFAP1L2. Identical amino acids are represented by * while similar amino acids are represented by :.

Figure 13

```
BAG64383   ----------------------------------------------------------------------------------
AFAP1L1    MDRGQVLEQLLPELTGLLSLLDHEYLSDTTLEKKMAVASILQSLQPLPAKEVSYLYVNTADLHSGPSFVESLFEEFDCDL  80
LAPTM5     ----------------------------------------------------------------------------------

BAG64383   ----------MDPRLSTVRQTCCCFNVRIATTALAIYHVIMSVLLFIEHSVEVAHGKA--------SCKLSQM---GYLR  59
AFAP1L1    SDLRDMPEDDGEPSKGASPELAKSPRLRNAADLPPPLPNKPPPEDYYEEALPLGPGKSPEYISSHNGCSPSHSIVDGYYE 160
LAPTM5     ----------MDPRLSTVRQTCCCFNVRIATTALAIYHVIMSVLLFIEHSVEVAHGKA--------SCKLSQM--GYLR  59

BAG64383   IADLISSFLLITMLFIISLSLLIGVVKNREKY-------LLPFLSLQ----------IMDYLLCLLTL----L-------  111
AFAP1L1    DADSSYPATRVNGELKSSYNDSDAMSSSYESYDEEEEGKSPQPRHQWPSEEASMHLVRECRICAFLLRKKRF-------  233
LAPTM5     IADLISSFLLITMLFIISLSLLIGVVKNREKY-------LLPFLSLQ----------IMDYLLCLLTLLGSYIELPAYLK 122

BAG64383   ---------------------------GSYIELPTYLNFKSMNHMNYLPSQEDMPH----------------NQFIK  145
AFAP1L1    ---------------------------GQWAKQLTVIREDQL--LCYKSSKDRQPHLRLALDTCSIIYVPKDSRHKR  281
LAPTM5     LASRSRASSSKFPLMTLQLLDFCLSILTLCSSYMEVPTYLNFKSMNHMNYLPSQEDMPH----------------NQFIK 186

BAG64383   MMIIFSIAFITVLIFKVYMFKCVWRCYRLIKCM-------NSVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNCS 218
AFAP1L1    HELRFTQGATEVLVLALQSREQAEEWLKVIREVSKPVGGAEGVEVPRSPVLLCKLDLDKRLSQEKQTSDSDSVGVGDNCS 361
LAPTM5     MMIIFSIAFITVLIFKVYMFKCVWRCYRLIKCM-------NSVEEKRNSKMLQKVVLP---SYEEALSPAGPV----SCS 252

BAG64383   TLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVPCCGYLNVLVNQGWKERWCRLK 298
AFAP1L1    TLGRRETCDHGKGKKSSLAELKGSMSRAAGRKITRIIGFSKKKTLADDLQTSSTEEEVPCCGYLNVLVNQGWKERWCRLK 441
LAPTM5     QLGQRFRLSHSSGL--SIIQPNNAW----------FISISSD-TCLDDWPF------------------------- 290

BAG64383   CNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAFRILRNRQEVAILEASCSEDMGRWLGLLLVGMGSRVTPEALH 378
AFAP1L1    CNTLYFHKDHMDLRTHVNAIALQGCEVAPGFGPRHPFAFRILRNRQEVAILEASCSEDMGRWLGLLLVEMGSRVTPEALH 521
LAPTM5     ---------------------------------------------------------------------------------

BAG64383   YDYVDVETLTSIVSAGRNSFLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKRH 458
AFAP1L1    YDYVDVETLTSIVSAGRNSFLYARSCQNQWPEPRVYDDVPYEKMQDEEPERPTGAQVKRHASSCSEKSHRVDPQVKVKRH 601
LAPTM5     ---------------------------------------------------------------------------------

BAG64383   ASSANQYKYGKNRAEEDARKYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGAKLKALEEAVATLEAQCRAKEE 538
AFAP1L1    ASSANQYKYGKNRAEEDARRYLVEKEKLEKEKETIRTELIALRQEKRELKEAIRSSPGAKLKALEEAVATLEAQCRAKEE 681
LAPTM5     ---------------------------------------------------------------------------------

BAG64383   RRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSGETANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAK 618
AFAP1L1    RRIDLELKLVAVKERLQQSLAGGPALGLSVSSKPKSGETANKPQNSVPEQPLPVNCVSELRKRSPSIVASNQGRVLQKAK 761
LAPTM5     ---------------------------------------------------------------------------------

BAG64383   EWEMKKT  625
AFAP1L1    EWEMKKT  768
LAPTM5     -------
```

Figure 13. AFAP1L1 and BAG64383

A protein identical to lysosomal protein transmembrane 5 (LAPTM5) in its N-terminal 181 amino acids and identical to AFAP1L1 in its C-terminal 444 amino acids was aligned to each protein by ClustalW2 analysis (27). BAG64383 sequence identical to LAPTM5 is highlighted in yellow, sequence identical to AFAP1L1 is highlighted in green, sequence shared between all three proteins is highlighted in pink and sequence unique to BAG64383 is highlighted in light blue.

CHAPTER 4

UNIQUE METHODS

I. Mutation and formation of full length AFAP1L1 from separate plasmids

For purposes of creating a plasmid to express tagged and untagged full length nucleotide sequence of AFAP1L1, two vectors were purchased from Thermo Scientific OpenBiosystems, Huntsville, Alabama. The first, a pCMV-SPORT6 vector, contained coding sequence for the first 340 amino acids of *afap1l1*. The second, a pINCY vector, contained coding sequence for amino acids 274 to 768. An overlap region of 67 amino acids existed between the two sequences, encompassing amino acids 274 to 340 (Figure 1). To piece together the 5' and 3' halves of *afap1l1*, we identified a restriction enzyme site in the overlap region that could be used to digest *afap1l1* part 1 from pCMV-SPORT6 and *afap1l1* part 2 from pINCY which would create sticky ends which could be ligated together to form full length *afap1l1*. Using the New England BioLabs NEBcutter program to find restriction enzyme sites, we determined that a BstYI restriction enzyme site (AGATCC) in the overlap region could be mutated to a BglII restriction enzyme site (AGATCT) by mutating one base pair of amino acid 329, TCC to TCT. This would create a restriction site unique to each plasmid so that multiple cut sites would not be found and also would not change the amino acid encoded as TCC and TCT both code for the amino acid serine.

Mutagenesis of amino acid 329 was accomplished by using the QuikChange Multi-Site Directed Mutagenesis Kit from Stratagene. Forward and reverse primers were synthesized (Forward 5'GAGGTCCCCAGATCTCCAGTCCTCCTGTG3' and Reverse 5'CACAGGAGGACTGGAGATCTGGGGACCTC3') and the following reaction mixture was used:

10X Buffer	2.5µl
dH₂0	16µl
pSPORT6 or pINCY template	2.5µl (50ng total)
Forward primer	1µl (100ng total)
Reverse primer	1µl (100ng total)
dNTPs	1µl
QuikChange Enzyme	1µl

The reaction mixture was incubated in the following PCR protocol:

Segment	Cycle	Temperature	Time
1	1	95°C	1 minute
2	30	95°C	1min
		55°C	1min
		65°C	9min
3	Hold	4°C	Hold

The PCR product was incubated at 37°C for 1 hour with addition of restriction enzyme Dpn1 to degrade parental DNA and the final product was transformed into dH5α bacterial cells by incubation of 1µg PCR product with 50µl dH5α for 30 minutes followed by 30 seconds of heat shock at 42°C and incubation on ice for 2 minutes. 1ml of SOC media was added and the bacterial mixture was shaken at 37°C for one hour in a heated shaker. After 1 hour, bacteria were spun down at 5000rpm, 900µl of media was removed, bacteria were resuspended in the remaining 100µl of SOC media and plated onto ampicillin plates. Control transformations with no plasmid added were also conducted.

Multiple colonies from each mutated pCMV-SPORT6 and pINCY were picked and grown overnight in 3ml of LB+Ampicillin shaking at 37°C. 1.5ml of overnight culture was centrifuged at 12,000rpm for 30 seconds and the bacterial pellet was lysed in 50µl of EasyPrep lysis buffer (10mM TrisHCl pH8.0, 1mM EDTA, 15% sucrose, 2mg/ml lysozyme, 0.2mg/ml RNase A and 0.1mg/ml BSA). Lysates were incubated at room temperature for 5 minutes, boiled at 100°C for 90 seconds, chilled on ice for 2 minutes and then centrifuged 12,000rpm for 20 minutes at room temperature. 5µl of each DNA prep was digested with restriction enzymes in the following manner to determine if they contained their respective *afap1l1* insert. pCMV-SPORT6 vectors were cut with Not1 and Sal1 to cut out *afap1l1* part 1 and pINCY vectors were cut with EcoRI and Not1 to check for *afap1l1* part 2.

pCMV-SPORT6	pINCY
5µl DNA	5µl DNA
2µl dH20	2µl dH$_2$O
2µl Buffer H	2µl Buffer H
0.5µl Not1	0.5µl EcoRI
0.5µl Sal1	0.5µl Not1

Digests were placed at 37°C for two hours and run on a 1% Seakem agarose gel containing ethidium bromide to determine their banding pattern. If the *afap1l1* insert was present, pCMV-SPORT6 should have produced two bands, 4396bp and 1314bp. The pINCY vector should create two bands with a molecular weight (MW) of 4080bp and 2607bp. Five colonies from pCMV-SPORT6 and three colonies from pINCY provided the correct banding patterns. 200µl of pCMV-SPORT6 colonies 2 and 7 and pINCY

colonies 2, 7 and 11 were grown overnight in 500ml LB+Ampicillin shaking at 37°C to perform a midi prep of DNA so as to send for sequencing.

Colonies 2 and 7 of pCMV-SPORT6 and colony 11 of pINCY showed mutation of amino acid 329 from a TCC codon to TCT while keeping the amino acid as a serine (Figure 2). To ligate together the two halves of *afap1l1*, it was decided to insert each mutated portion of *afap1l1* in to a pBlueScriptII KS+ vector (Stratagene) followed by digestion of the vector with the engineered BglII site and a unique ScaI site and ligation so as to bring together the two halves of KS+ and full length *afa1l1*. The restriction sites HindIII and EcoRI in the KS+ vector were chosen because pCMV-SPORT6 contained a HindIII site which could be used to cut out and ligate *afap1l1* part 1. An EcoRI site was engineered through site-directed mutagenesis (forward 5'CGAATTCCTTGTCCAGGTCCAACTT3', reverse 5'AAGTTGGACCTGGACAAGGAATTCG3') at the end of the *afap1l1* part 1 fragment so as to insert the *afap1l1* sequencing into the KS+ vector. Additionally, pINCY contained an EcoRI site at the end of the *afap1l1* part 2 sequence and an engineered HindIII site (forward 5'ATAAGCTTTGCCCAAGGACAGCCGGCA3', reverse 5'TGCCGGCTGTCCTTGGGCAAAGCTTAT3') was created in the beginning of the *afap1l1* part 2 sequence so as to ligate to the KS+ vector. PCR mixture and parameters were as follows:

pCMV-SPORT6	pINCY
40 μl dH$_2$0	40 μl dH$_2$0
5μl Pfu Buffer	5μl Pfu Buffer
1μl dNTPs(10mM)	1μl dNTPs (10mM)
1μl pCMV-SPORT6 (100ng/ml)	1μl pINCY (100ng/ml)
1μl pCMV-SPORT6 EcoRI forward primer	1μl pINCY HindIII forward primer
1μl pCMV-SPORT 6 reverse primer	1μl pINCY HindIII reverse primer
1μl Pfu Polymerase (2.5U)	1μl Pfu Polymerase (2.5U)

Segment	Cycle	Temperature	Time
1	1	95°	2min
2	30	95°	30sec
		Primer Tm-5°	30sec
		72°	1min per kb
3	1	72°	10min
4	Hold	4°	Hold

PCR products were cleaned using the Invitrogen PCR Purification Kit and digested with their respective restriction enzymes to produce sticky ends compatible for ligation. 30μl of PCR product were digested using 4μl Buffer B, 4μl H$_2$O, 3μl HindIII and 3μl EcoRI by incubation at 37°C for two hours. KS+ vector was also cut using HindIII and EcoRI restriction enzymes for two hours and phosphatased for an additional hour. Cut PCR products and vector were run on a 1% Seakem agarose gel, stained by incubation in water containing ethidium bromide for one hour and cut and eluted from the gel using an Invitrogen Quick Gel Extraction Kit. *Afap1l1* part 1 and *afap1l1* part 2 were each ligated separately to KS+ using T4 ligase. 1μl of KS+ plasmid treated with alkaline phosphatase was incubated with 21μl dH2O, 3μl T4 Buffer, 1μl T4 Ligase and 4μl of either *afap1l1* part 1 or *afap1l1* part 2 overnight at 16°C. 2.5μl of each ligation product was incubated with 60μl dH5α bacterial cells on ice for 30 minutes, heat shocked for 30

seconds and placed back onto ice for 2 minutes. 1ml of SOC media was added and bacteria were grown while shaking at 37°C for one hour. Bacteria were spun down, 900µl of SOC media removed and bacteria brought up in the remaining 100µl SOC media and plated onto LB plates containing ampicillin. Colonies formed from ligations were grown overnight in 3ml of LB + Ampicillin. 1.5ml of overnight culture was subjected to EasyPrep lysis and 5µl used for digestion with EcoRI and HindIII restriction enzymes to insure that the KS+ vector contained the respective *afap1l1* insert. If the digest did drop out an EcoRI/HindIII band, this bacterial lysate was also subjected to digestion with BglII to insure that the insert was indeed mutated *afap1l1*. Once KS+ vectors containing the correct *afap1l1* part 1 and *afap1l1* part 2 inserts were determined, 5µg of each of these vectors was cut using the engineered BglII site in the overlap region of AFAP1L1 and a unique Sca1 site in the KS+ vector in the following manner:

AFAP1L1 part 1	AFAP1L1 part 2
3µl part 1 (5µg)	4.5µl part 2 (5µg)
4µl dH20	2.5µl dH$_2$O
1µl Buffer H	1µl Buffer H
1µl BglII	1µl BglII
1µl Sca1	1µl Sca1

Additionally, the KS+/*afap1l1* part 1 was chosen to be phosphatased for one hour after digestion. Digests were run on a 1% Seakem agarose gel and stained for one hour with water containing ethidium bromide. The cut pCMV-SPORT6 vector should produce two bands, a 2790bp fragment that contains the *afap1l1* part 1 insert and an 1158bp fragment. The cut pINCY vector should produce two bands, a 2479bp fragment that

contains *afap1l1* part 2 and a 1794bp fragment. Fragments containing their respective inserts were eluted from the gel using the Invitrogen Quick Gel Extraction Kit. The KS+/*afap1l1* part 1 and KS+/*afap1l1* part 2 fragments were ligated together using T4 ligase overnight at 16°C and bacteria were transformed and plated onto LB+Ampicillin plates. Colonies were picked, grown overnight in 3ml LB+Ampicillin and subjected to EasyPrep lysis. 5μl of lysate was digested using EcoRI and HindIII to look for a fragment consistent with full length *afap1l1*. A colony containing the correct banding pattern consistent with full length *afap1l1* was chosen to grow up in 500ml of LB+Ampicillin overnight for midi prep. Vector KS+ containing full length *afap1l1* was subjected to digestion with EcoRI and HindIII and full length *afap1l1* was inserted into a pEGFP vector so as to utilize GFP tagged AFAP1L1. Additionally, the full length *afap1l1* was also inserted into a pcDNA3.1+ vector so as to utilize an untagged AFAP1L1. Sequencing of pEGFP-AFAP1L1 confirmed that full length *afap1l1* was expressed in the pEGFP vector.

II. Creation and purification of a rabbit polyclonal antibody against AFAP1L1: 1L1-CT

During early work with AFAP1L1, no commercial antibodies were available for use in molecular biology techniques. ProSci Incorporated of Poway, CA was commissioned by our lab to create an antibody against AFAP1L1. ProSci Inc. analyzed full length AFAP1L1 sequence and provided three possible antibody epitopes based upon regions of high hydrophilicity (H), antigenicity (A) and surface probability (SP) scored from 1-10, with 10 being the highest.

	Residue 82 DLRDMPEDDGEPSK	Residue 198 EGKSPQPRHQWPSE	Residue 714 KPKSGETANKPQNS
Turns	1	1	2
H	8	8	8
A	8	8	10
SP	8	8	7

A BLAST search was performed on each epitope to determine suitability for a unique antibody that would not react with other proteins, including other AFAP family members. As of August 21, 2007, Residue 82 epitope was discarded due to similarity with the ubiquitin-protein ligase Tom1. Residue 198 epitope showed homology with human and mouse forms of AFAP1L1. Residue 714 epitope, which is near the C-terminus, showed homology with the human form of AFAP1L1 and was chosen due to its lack of similarity with other proteins and also as a recommendation from ProSci Inc. that N- and C- termini epitopes tend to perform the best for western blotting and can perform better than higher ranked internal sequences. An immunogen of NH2-K^{714}PKSGETANKPQNS^{727}-OH was synthesized by ProSci Incorporated and used to inject two separate rabbits, PAS12219 and PAS12220. 5ml of pre-immune serum was collected from each rabbit and rabbits were immunized with synthesized immunogen at Week 0 using 200mg/rabbit in Complete Freud's Adjuvant. Rabbits were additionally immunized with 100mg/rabbit of immunogen in Incomplete Freud's Adjuvant at Weeks 2, 4 and 6. The first production bleed was collected at Week 5 with a 2nd bleed at Week 7, a 3rd bleed at Week 8 and a final bleed within five days of the 3rd bleed. Serum was shipped from ProSci Incorporated after each bleed and stored at -80°C.

Beginning with PAS12219, rabbit serum was purified using an Ultralink Iodoacetyl Gel peptide column from Pierce with 5mg immunogen C-

KPKSGETANKPQNS cross-linked to the gel. The antibody column was prepared by first capping the end of a column tube and adding 3ml of Ultralink Iodoacetyl Gel bead slurry to the provided column so as to end with 1.5ml of gel bed as this would hold a total of 6mg of peptide. The gel bed was allowed to settle for 30 minutes and the bottom cap was removed to allow gel bed liquid to flow out while re-capping the column as the liquid meniscus reached the top of the gel bed so as to avoid the potential for air pockets. Five gel bed volumes (7.5ml) of Coupling Buffer (50mM Tris, 5mM EDTA, pH8.46) were added to the column and allowed to flow through until liquid neared the top of the gel bed. 5mg of AFAP1L1 peptide was brought up in 1.5ml of Coupling Buffer, saving a small portion for before binding optical density (OD) reading, and added to the column. The column was secured tight with parafilm and rotated at room temperature for 15 minutes. The column was allowed to settle for 30 minutes and remaining buffer was allowed to flow out while keeping the flow through so as to compare to the before binding OD reading to determine how much protein was bound. The OD reading at wavelength 280 for peptide mixture before binding was 0.156. The OD reading for the flowthrough after binding was 0.050. It was determined that 32% of the protein came through in the flowthrough and 68% of the peptide remained bound to the column. The peptide column was washed with 3 gel bed volumes (4.5ml) of Coupling Buffer and allowed to drain until liquid reached the top of the gel bed. The column was then incubated with 1 gel bed volume (1.5ml) of Coupling Buffer containing 50mM L-Cysteine, mixed and rotated for 15 minutes at room temperature so as to block any other peptide binding sites in the peptide column. The column was allowed to settle for 30 minutes. With Couling Buffer with L-Cysteine still in the column, a column disc was

floated and soaked in the column so as to remove air bubbles from the disc. Using the provided cylinder, the disc was pushed down into the column to within 1mm of the gel bed without touching the gel bed. The column was then washed with 5 gel bed volumes (7.5ml) of 1M NaCl and then 5 gel bed volumes (7.5ml) of PBS with 0.02% NaN$_3$. 1ml of PBS with 0.02% NaN$_3$ was added to the column for preservation and it was stored upright at 4°C when not in use.

To pass serum over the column for the purpose of purifying antibody, the column and all solutions were allowed to come to room temperature so as to avoid introducing air bubbles into the column. Serum to be used (PAS12219 Bleed 1 for example) was thawed on ice and filtered through a 0.45μm filter using a syringe so as to remove any buildup that would stop flow of the column. Once to room temperature, the column was washed with 5 gel bed volumes (7.5ml) of PBS with no azide. After washing, the serum was passed over the column at 4°C in 2ml aliquots, running the entire amount of serum over the column three times. Returning to room temperature, the column was washed with 5 gel bed volumes (7.5ml) of PBS with no azide. To collect fractions, 750ul of Elution Buffer (100mM glycine, pH 2.3-3.0) was passed over the column. In the tube to collect fractions, 1/10 of the volume (75μl) of 1M Tris pH 7.5 was added so as to neutralize the low pH flowthrough. The OD of each fraction was read at 280nm as each fraction was collected so as to determine when peptide antibody came out of the column. For PAS12219 Bleed 1, 10 fractions were collected, resulting in ODs of

Fraction	OD 280nm
1	.360
2	.327
3	.264
4	1.610
5	1.168
6	.413
7	.413
8	.057
9	.057
10	.051

When graphed, this showed that the antibody eluted from the column in fractions 4-7. After antibody was successfully eluted from the column, the column was washed with 5 gel bed volumes (7.5ml) of PBS following by 5 gel bed volumes (7.5ml) with PBS containing 0.02% NaN_3 and stored at 4°C with 1ml PBS containing NaN_3 until the next antibody purification would be needed.

In addition to determining the OD reading for each fraction collected, an 11% acrylamide gel was run and stained with Coomassie stain so as to visualize any heavy and light chains present in the fractions. Although the OD reading can tell that peptide is present, visualization of the heavy and light chains is necessary to show that the peptide is antibody. Once confirmed, antibody can be used in various applications such as western blotting, immunohistochemistry, immunoprecipitation and other various molecular biology techniques.

Figure 1

```
ATG GAC CGA GGC CAG GTG CTG GAG CAG CTG CTC CCA GAG CTC ACC GGG CTG CTC AGC CTC CTG GAC CAC GAG
 M   D   R   G   Q   V   L   E   Q   L   L   P   E   L   T   G   L   L   S   L   L   D   H   E

TAC CTC AGC GAT ACC ACC CTG GAA AAG AAG ATG GCC GTG GCC TCC ATC CTG CAG AGC CTG CAG CCC CTT CCA
 Y   L   S   D   T   T   L   E   K   K   M   A   V   A   S   I   L   Q   S   L   Q   P   L   P

GCA AAG GAG GTC TCC TAC CTG TAT GTG AAC ACA GCA GAC CTC CAC TCG GGG CCC AGC TTC GTG GAA TCC CTC
 A   K   E   V   S   Y   L   Y   V   N   T   A   D   L   H   S   G   P   S   F   V   E   S   L

TTT GAA GAA TTT GAC TGT GAC CTG AGT GAC CTT CGG GAC ATG CCA GAG GAT GAT GGG GAG CCC AGC AAA GGA
 F   E   E   F   D   C   D   L   S   D   L   R   D   M   P   E   D   D   G   E   P   S   K   G

GCC AGC CCT GAG CTA GCC AAG AGC CCA CGC CTG AGA AAC GCG GCC GAC CTG CCT CCA CCG CTC CCC AAC AAG
 A   S   P   E   L   A   K   S   P   R   L   R   N   A   A   D   L   P   P   P   L   P   N   K

CCT CCC CCT GAG GAC TAC TAT GAA GAG GCC CTT CCT CTG GGA CCC AAG TCG CCT GAG TAC ATC AGC TCC
 P   P   P   E   D   Y   Y   E   E   A   L   P   L   G   P   K   S   P   E   Y   I   S   S

CAC AAT GGC TGC AGC CCC TCA CAC TCG ATT GTG GAT GGC TAC TAT GAG GAC GCA GAC AGC AGC TAC CCT GCA
 H   N   G   C   S   P   S   H   S   I   V   D   G   Y   Y   E   D   A   D   S   S   Y   P   A

ACC AGG GTG AAC GGC GAG CTT AAG AGC TCC TAT AAT GAC TCT GAC GCA ATG AGC AGC TCC TAT GAG TCC TAC
 T   R   V   N   G   E   L   K   S   S   Y   N   D   S   D   A   M   S   S   S   Y   E   S   Y

GAT GAA GAG GAG GAG GAA GGG AAG AGC CCG CAG CCC CGA CAC CAG TGG CCC TCA GAG GAG GCC TCC ATG CAC
 D   E   E   E   E   G   K   S   P   Q   P   R   H   Q   W   P   S   E   E   A   S   M   H

CTG GTG AGG GAA TGC AGG ATA TGT GCC TTC CTG CTG CGG AAA AAG CGT TTC GGG CAG TGG GCC AAG CAG CTG
 L   V   R   E   C   R   I   C   A   F   L   L   R   K   K   R   F   G   Q   W   A   K   Q   L

ACG GTC ATC AGG GAG GAC CAG CTC CTG TGT TAC AAA AGC TCC AAG GAT CGG CAG CCA CAT CTG AGG TTG GCA
 T   V   I   R   E   D   Q   L   L   C   Y   K   S   S   K   D   R   Q   P   H   L   R   L   A

CTG GAT ACC TGC AGC ATC ATC TAC GTG CCC AAG GAC AGC CGG CAC AAG AGG CAC GAG CTG CGT TTC ACC CAG
 L   D   T   C   S   I   I   Y   V   P   K   D   S   R   H   K   R   H   E   L   R   F   T   Q

GGG GCT ACC GAG GTC TTG GTG CTG GCA CTG CAG AGC CGA GAG CAG GCC GAG GAG TGG CTG AAG GTC ATC CGA
 G   A   T   E   V   L   V   L   A   L   Q   S   R   E   Q   A   E   E   W   L   K   V   I   R

GAA GTG AGC AAG CCA GTT GGG GGA GCT GAG GGA GTG GTC CCC AGA TCC CCA GTC CTC CTG TGC AAG TTG
 E   V   S   K   P   V   G   G   A   E   G   V   E   V   P   R   S   P   V   L   L   C   K   L

GAC CTG GAC AAG AGG CTG TCC CAA GAG AAG CAG ACC TCA GAT TCT GAC AGC GTG GGT GTG GGT GAC AAC TGT
 D   L   D   K   R   L   S   Q   E   K   Q   T   S   D   S   D   S   V   G   V   G   D   N   C

TCT ACC CTT GGC CGC CGG GAG ACC TGT GAT CAC GGC AAA GGG AAG AAG AGC AGC CTG GCA GAA CTG AAG GGC
 S   T   L   G   R   R   E   T   C   D   H   G   K   G   K   K   S   S   L   A   E   L   K   G

TCA ATG AGC AGG GCT GCG GGC CGC AAG ATC ACC CGT ATC ATT GGC TTC TCC AAG AAG AAG ACA CTG GCC GAT
 S   M   S   R   A   A   G   R   K   I   T   R   I   I   G   F   S   K   K   K   T   L   A   D

GAC CTG CAG ACG TCC TCC ACC GAG GAG GAG GTT CCC TGC TGT GGC TAC CTG AAC GTG CTG GTG AAC CAG GGC
 D   L   Q   T   S   S   T   E   E   E   V   P   C   C   G   Y   L   N   V   L   V   N   Q   G

TGG AAG GAA CGC TGG TGC CGC CTG AAG TGC AAC ACT CTG TAT TTC CAC AAG GAT CAC ATG GAC CTG CGA ACC
 W   K   E   R   W   C   R   L   K   C   N   T   L   Y   F   H   K   D   H   M   D   L   R   T

CAT GTG AAC GCC ATC GCC CTG CAA GGC TGT GAG GTG GCC CCG GGC TTT GGG CCC CGA CAC CCA TTT GCC TTC
 H   V   N   A   I   A   L   Q   G   C   E   V   A   P   G   F   G   P   R   H   P   F   A   F

AGG ATC CTG CGC AAC CGG CAG GAG GTG GCC ATC TTG GAG GCA AGC TGT TCA GAG GAC ATG GGT CGC TGG CTC
 R   I   L   R   N   R   Q   E   V   A   I   L   E   A   S   C   S   E   D   M   G   R   W   L

GGG CTG CTG CTG GTG GAG ATG GGC TCC AGA GTC ACT CCG GAG GCG CTG CAC TAT GAC TAC GTG GAT GTG GAG
 G   L   L   L   V   E   M   G   S   R   V   T   P   E   A   L   H   Y   D   Y   V   D   V   E

ACC TTA ACC AGC ATC GTC AGT GCT GGG CGC AAC TCC TTC CTA TAT GCA AGA TCC TGC CAG AAT CAG TGG CCT
 T   L   T   S   I   V   S   A   G   R   N   S   F   L   Y   A   R   S   C   Q   N   Q   W   P

GAG CCC CGA GTC TAT GAT GAT GTT CCT TAT GAA AAG ATG CAG GAC GAG GAG CCC GAG CGC CCC ACA GGG GCC
 E   P   R   V   Y   D   D   V   P   Y   E   K   M   Q   D   E   E   P   E   R   P   T   G   A

CAG GTG AAG CGT CAC GCC TCC TCC TGC AGT GAG AAG TCC CAT CGT GTG GAC CCG CAG GTC AAA GTC AAA CGC
 Q   V   K   R   H   A   S   S   C   S   E   K   S   H   R   V   D   P   Q   V   K   V   K   R

CAC GCC TCC AGT GCC AAT CAA TAC AAG TAT GGC AAG AAC CGA GCC GAG GAG GAT GCC CGG AGG TAC TTG GTA
 H   A   S   S   A   N   Q   Y   K   Y   G   K   N   R   A   E   E   D   A   R   R   Y   L   V

GAA AAA GAG AAG CTG AAG GAG AAA GAG AAG GAG ACG ATT CGG ACA GAG CTG ATA GCA CTG AGA CAG GAG AAG AGG
 E   K   E   K   L   K   E   K   E   K   E   T   I   R   T   E   L   I   A   L   R   Q   E   K   R

GAA CTG AAG GAA GCC ATT CGG AGC AGC CCA GGA GCA AAA TTA AAG GCT CTG GAA GAA GCC GTG GCC ACC CTG
 E   L   K   E   A   I   R   S   S   P   G   A   K   L   K   A   L   E   E   A   V   A   T   L

GAA GCT CAG TGT CGG GCA AAG GAG GAG CGC CGG ATT GAC CTG GAG CTG AAG CTG GTG GCT GTG AAG GAG CGC
 E   A   Q   C   R   A   K   E   E   R   R   I   D   L   E   L   K   L   V   A   V   K   E   R

TTG CAG CAG TCC CTG GCA GGA GGG CCA GCC CTG GGC CTC TCC GTG AGC AGC AAG CCC AAG AGT GGG GAA ACT
 L   Q   Q   S   L   A   G   G   P   A   L   G   L   S   V   S   S   K   P   K   S   G   E   T

GCA AAT AAA CCC CAG AAC AGC GTT CCA GAG CAA CCT CTC CCT GTC AAC TGT GTT TCT GAG CTG AGG AAG AGG
 A   N   K   P   Q   N   S   V   P   E   Q   P   L   P   V   N   C   V   S   E   L   R   K   R

AGC CCA TCC ATC GTA GCC TCC AAC CAA GGA AGG GTG CTA CAG AAA GCC AAG GAA TGG GAA ATG AAG AAG ACC
 S   P   S   I   V   A   S   N   Q   G   R   V   L   Q   K   A   K   E   W   E   M   K   K   T

TAG
STOP
```

Figure 1: Amino acid sequence of AFAP1L1 with highlighted overlap

The pCMV-SPORT6 vector contained coding sequence for amino acids 1-340 of *afap1l1*. The pINCY vector contained coding sequence for amino acids 274-768 of *afap1l1*. An overlap region (highlighted) between the two sequences encompassed 67 amino acids from amino acid 274 to 340. A BstYI restriction site in the overlap region was changed to a unique BglII restriction site by mutation of amino acid 329, changing coding sequence from TCC to TCT to create the BglII site while keeping the amino acid a serine.

Figure 2

A.

```
pSPORT6.mut.check.    CGAGAGCAGGCCGAGGAGTGGCTGAAGGTCATCCGAGAAGTGAGCAAGCCAGTTGGGGGA 151

AFAP1L1               CGAGAGCAGGCCGAGGAGTGGCTGAAGGTCATCCGAGAAGTGAGCAAGCCAGTTGGGGGA 960

                      ************************************************************


pSPORT6.mut.check.    GCTGAGGGAGTGGAGGTCCCCAGATCTCCAGTCCTCCTGTGCAAGTTGGACCTGGACAAG 211

AFAP1L1               GCTGAGGGAGTGGAGGTCCCCAGATCCCCAGTCCTCCTGTGCAAGTTGGACCTGGACAAG 1020

                      ***********************  *******************************
```

B.

```
pINCY.mut.check.      CGAGAGCAGGCCGAGGAGTGGCTGAAGGTCATCCGAGAAGTGAGCAAGCCAGTTGGGGGA 118

AFAP1L1               CGAGAGCAGGCCGAGGAGTGGCTGAAGGTCATCCGAGAAGTGAGCAAGCCAGTTGGGGGA 960

                      ************************************************************


pINCY.mut.check.      GCTGAGGGAGTGGAGGTCCCCAGATCTCCAGTCCTCCTGTGCAAGTTGGACCTGGACAAG 178

AFAP1L1               GCTGAGGGAGTGGAGGTCCCCAGATCCCCAGTCCTCCTGTGCAAGTTGGACCTGGACAAG 1020

                      ***********************  *******************************
```

Figure 2. Mutation of amino acid 329 in *afap1l1*

Afap1l1 part 1 in the pCMV-SPORT6 vector (A) and *afap1l1* part 2 in the pINCY vector
(B) were successfully mutated from TCC to TCT (highlighted) so as to introduce a
unique BglII restriction enzyme site while keeping the amino acid a serine.

CHAPTER 5

General Discussion

The Actin Filament Associated Proteins represent a family of adaptor proteins that have similar overall structure while each displays a signature function. All three family members share modular domain structure with at least one N-terminal SH3 binding motif, at least one N-terminal SH2 binding motif, two PH domains in the central portion of the protein which surround a region rich in serine and threonine residues, at least one SH2 binding motif in the C-terminal region and a helical region at the C-terminus. In AFAP1, this region is known to contain a helical leucine zipper motif and actin binding domain while AFAP1L1 also contains a putative leucine zipper and actin binding domain. AFAP1L2 contains a coiled-coil in this region and is hypothesized to bind actin through a more C-terminal region (Baisden et al., 2001; Lodyga et al., 2010; Xu et al., 2007). It is the sequence similarity of the PH domains that define the AFAP proteins as a family. Although PH domains contain a characteristic helix-loop-helix structure, the amino acids within PH domains vary widely and little conservation is seen among proteins (DiNitto and Lambright, 2006). This is not true, however, for the AFAP proteins as they share 44% identity (81% similarity) in the PH1 domains and 40% identity (73% similarity) in their PH2 domains. The AFAP proteins may be evolutionarily related by two chromosomal duplications that are hypothesized to have taken place during the evolution of invertebrates to vertebrates (Dehal and Boore, 2005; Ohno, 1970). Conservation of overall structure through evolution may explain the high level of similarity between the PH domains of the different family members. Although similar in their modular domain, each AFAP family member has its own unique function that is characterized by its individual amino acid sequence. While AFAP1 has been intensively studied for its cellular function and binding abilities and AFAP1L2 has

currently been a topic of study for its role in the PI3K/Akt pathway, this dissertation served to define the cellular role of AFAP1L1 (Baisden et al., 2001; Lodyga et al., 2009).

So named for their ability to interact with actin filaments, the AFAP proteins may play pivotal roles in regulation of the actin cytoskeleton. Cytoskeletal rearrangement is indispensable to a cell and is responsible for a number of processes such as motility in embryogenesis and wound healing, immune cell adherence and movement, pathogen infection and invasion. Each AFAP family member has been shown to interact with actin (Flynn et al., 1993; Lodyga et al., 2010). AFAP1 plays a role in cell contractility by virtue of its ability to bind actin filaments and multimerize, thus bundling actin filaments together. The loss of AFAP1 results in a loss of stress fiber integrity (Dorfleutner et al., 2007). AFAP1L2 translocates from the cytosol to the lamellipodium and interacts with branched F-actin networks in a Rac-dependent manner (Lodyga et al., 2010). AFAP1L1 is shown to bind actin filaments and may play a similar role to AFAP1 in actin cross-linking due to conservation of a putative leucine zipper. This may give AFAP1L1 the ability to bind other AFAP molecules and thus also play a role in actin filament cross-linking and stress fiber formation.

While each AFAP family member encodes at least one SH3 binding motif, they vary in their ability to bind different proteins. The core amino acids of an SH3 binding motif, a PXXP motif, are surrounded by various amino acids that convey specificity for the binding of SH3 domains. The SH3 domains of AFAP1 and AFAP1L2 have been shown to interact with cSrc while the SH3 domain of AFAP1L1 has been shown to interact with cortactin (Guappone and Flynn, 1997; Xu et al., 2007). This cortactin

interaction is unique to AFAP1L1 and may play an important role in cell motility and invasion.

Cortactin is an adaptor protein that is necessary for stabilization of actin branching in growing actin filaments and also for the formation of invadosomes. In its inactive state, cortactin displays diffuse cytoplasmic staining around the perinuclear region. Upon cellular stimulation with growth factors, cortactin moves from its perinuclear region to active sites of cytoskeletal branching where it binds to the sides of growing actin filaments and interacts with the Arp2/3 complex to stabilize branching points (Ammer and Weed, 2008). Cortactin is also found in cellular adhesions such as invadosomes and is necessary for their formation and function (Ayala et al., 2008). High levels of cortactin indicate poor prognosis in tumors as these tend to be more aggressive and have an increased ability to metastasize (Weaver, 2008). Although cellular levels of AFAP1L1 are relatively low in many cell lines, the cancer cell line MDA-MB-435 expresses high levels of AFAP1L1. When comparing the endogenous localizations of AFAP1L1 and cortactin in unstimulated MDA-MB-435 cells, AFAP1L1 is found to decorate actin filaments and have some cytoplasmic staining in the perinuclear region while cortactin is found in its expected perinuclear location. When active cSrc was transfected so as to create a transformed phenotype in these cells, both AFAP1L1 and cortactin moved to invadopodia. Cortactin is a substrate for cSrc phosphorylation on key tyrosine residues in the helical region and, upon their phosphorylation, is hypothesized to undergo a conformation change under which cortactin may bind actin filaments more efficiently and also open up binding sites for other proteins (Ammer and Weed, 2008). It is through the phosphorylation of cortactin by active cSrc and subsequently activated

kinases which may have allowed for AFAP1L1 binding to the SH3 domain of cortactin through its SH3 binding motif. AFAP1L1 and cortactin were then able to localize to invadopodia and, by virtue of their binding motifs, act as scaffolds for other invadosome-related proteins.

Interestingly, it was found that upon the overexpression of GFP-AFAP1L1, A7r5 smooth muscle cells expressing high levels of GFP-AFAP1L1 had the ability to spontaneously dissolve their stress fibers and form podosomes in which both GFP-AFAP1L1 and cortactin colocalized. This is similar to the expression of AFAP1 as AFAP1 is found to decorate actin filaments on an endogenous level. In unstimulated cells, the AFAP1 binding partner cSrc is found in an inactive state on vesicles in the perinuclear region. AFAP1 moves to the perinuclear region where it binds to and activates cSrc which then moves as a complex to sites of actin rearrangement (Walker et al., 2007). High levels of exogenous GFP-AFAP1 have the ability to induce podosome formation without cell stimulation in a subset of cells. This was attributed to the culture conditions of the cells in which endogenous levels of PKCα may become activated, thus activating the high levels of AFAP1 and in turn activating cSrc and podosome formation (Dorfleutner et al., 2008). This may also be the case for AFAP1L1; however it is interesting that while many cells overexpresse GFP-AFAP1L1, cells expressing low levels of GFP-AFAP1L1 showed highly formed stress fibers while only those with the highest levels of GFP-AFAP1L1 expression formed spontaneous podosomes. Perhaps there is an intrinsic ability of AFAP1L1 to induce podosome formation. This ability may allow AFAP1L1 to move to cortactin and ferry it to invadosomes, much like the activity of AFAP1 and cSrc.

Despite the fact that the mechanism for the relocation of AFAP1L1 to invadosomes remains unclear, AFAP1L1 is a newly described invadosome protein, adding to the large array of cytoskeletal and signaling proteins found in these complexes. Invadosomes have both physiological and pathological roles. The ability of a cell to cross tissue barriers is necessary in the case of immune cells but is exploited by transformed cancer cells which use the process for metastasis (Linder, 2009). While podosomes have been extensively studied in a number of different cell types, they have recently been shown to play a role in acetylcholine receptor clustering on the skeletal muscle side of the neuromuscular junction and this is their first description for a role in neuronal synapses (Proszynski et al., 2009). These podosomes are necessary for remodeling of the postsynaptic membrane and are an integral part of its maturity. The neuromuscular junction is quite large when compared to typical synapses between neurons, and these smaller synapses are not large enough to contain conventional podosomes. However, neuronal synapses do share a similar maturation pathway with the neuromuscular junction and it is hypothesized that podosome-like complexes may be found within these synapses as they contain very similar proteins that are involved in actin reorganization (Proszynski et al., 2009).

The dendrite of a receiving neuron undergoes constant actin cytoskeletal reorganization, known as synaptic plasticity, to strengthen or weaken its receiving signal. Dendritic spines are small protrusions from the dendrite which receive signals from the presynaptic neuron and have the ability to change their shape or number quickly in response to receiving stimuli. Dendritic spines cluster neurotransmitter receptors to an area at the tip of the spine known as the postsynaptic density which houses a diverse

group of adaptor proteins that link transmembrane neurotransmitter receptors to the underlying actin cytoskeleton so as to transmit the receiving signal (Boeckers, 2006). Cortactin is one of the major constituents of the postsynaptic density as its role in cytoskeletal rearrangement is paramount to the function of the dendritic spine (Hering and Sheng, 2003; Sekino et al., 2007).

Immunohistochemical detection shows a differential expression of AFAP1L1 from AFAP1 in the adult human brain. AFAP1L1 has a unique localization around the Purkinje neurons and granule cells of the cerebellum and is also found surrounding the neurons of the dentate nucleus. Although we cannot identify the exact location of AFAP1L1 through this immunohistochemical signal, it is possible that these immunoreactive sites represent the clustering of actin associated proteins at the postsynaptic density in which AFAP1L1 may play a role. The ability of AFAP1L1 to localize with actin filaments, bind to cortactin and be involved in podosome formation may play a role in the maturation of neuronal synapses by neurotransmitter clustering and up and down regulation of dendritic spines.

Work by Tabakoff et al. supports the hypothesis of a role for AFAP1L1 in the postsynaptic density. AFAP1L1 was shown to be upregulated in the brains of rats that became addicted to alcohol and is considered a candidate gene that influences dependency versus non-dependency. Tabakoff et al. hypothesize that AFAP1L1 plays a role in receptor trafficking by virtue of its actin filament association and propose its location in the neuronal synapse to be at the postsynaptic density in association with F-actin and other cytoskeletal associated proteins (Tabakoff et al., 2009).

The Oncomine database predicts high expression levels of AFAP1L1 in cancers of the nervous system, particular neuroblastoma and glioblastoma. Neuroblastoma occurs most commonly in children from bundles of immature nerve fibers and has a fairly good prognosis while glioblastoma arises from glial cells and is the most common and most aggressive primary brain tumor in humans (Maris, 2010; Miller and Perry, 2007). We have shown that AFAP1L1 can be detected by immunohistochemistry in glial cells. AFAP1 has also been shown to play a role in glioblastoma as it is highly expressed in last-stage glioblastoma cell lines. This expression overlaps with the expression of the Src family kinases cSrc, Fyn and Lyn. AFAP1 is poised in glioblastoma to relay signals through SFK that that can promote proliferation and invasion of the tumor (Clump, 2008). AFAP1L1 may play a similar role and its expression in glial tumors may enhance the aggressiveness of the tumor.

The Oncomine database also predicts high levels of AFAP1L1 in melanoma. Melanoma arises from melanin-producing melanocytes in the skin and is considered a very serious and aggressive form of cancer (Berwick, 2006). Cortactin is known to be highly increased in metastatic melanoma and its aberrant expression may have a role in tumor progession (Xu et al., 2010). While we have not tested for levels of AFAP1L1 in melanoma in our studies, our observations of high levels of AFAP1L1 in MDA-MB-435 cells provides as intriguing idea that AFAP1L1 may play a role in melanoma as well. MDA-MB-435 cells were originally isolated in 1978 from a pleural effusion of a breast cancer patient (Cailleau et al., 1978). Considered a highly aggressive and metastatic form of breast cancer, the MDA-MB-435 cell line was used in many studies to determine the signaling pathways of breast cancer cells until it was noticed by Ross et al. that MDA-

MB-435 cells did not cluster with other breast cancer cell lines based upon their mRNA expression patterns and instead clustered among various melanoma cell lines (Ross et al., 2000). Many studies have taken place since this discovery to determine the exact origin of MDA-MB-435 cells. Theories range from misidentification of the cell line to contamination with and eventual succession by the M14 melanoma cell line. Additionally, it is hypothesized that the MDA-MB-435 cell line was created from an undiagnosed metastatic melanoma lesion and not a breast lesion from the beginning. Multiple studies have shown that MDA-MB-435 cells do not express genes typically found in breast cancer and instead express those involved in melanoma. Through extensive gene expression profiling, SNP profiling and karyotyping, MDA-MB-435 cells are considered to be more melanocytic in nature than related to breast cells. Regardless of their origin, AFAP1L1 expression levels are high in MDA-MB-435 cells. If it is true that these cells are indeed melanocytic or have taken on a melanocytic genotype, it is intriguing that AFAP1L1 and cortactin are highly expressed in conjuction with each other in cancer.

AFAP1L1 represents a newly identified AFAP family member that may have arisen from AFAP1 through gene duplication and plays a role in the regulation of the actin cytoskeleton. Through its unique interaction with cortactin, AFAP1L1 may play a role in motility and invasion as witnessed by its ability to localize to invadosomes and induce spontaneous podosomes upon high levels of overexpression. The distinct expression of AFAP1L1 in the area surrounding the Purkinje neurons and granule cells of the cerebellum as well as the neurons of the dentate nucleus implies that AFAP1L1 may play a role in these areas, particularly in the postsynaptic density which is a site of high

actin dynamics. As podosomes have recently been described in the neuromuscular junction, it is possible that similar complexes occur in neuronal synapses at the postsynaptic density and AFAP1L1 is a regulator of these complexes to govern receptor trafficking and clustering. The unique staining of AFAP1L1 in the brain is also supported by the fact that AFAP1L1 is predicted to be high in cancers of the central nervous system, particularly glioblastoma, as we find high levels of AFAP1L1 in glial cells. This interaction with cortactin and ability to alter the actin cytoskeleton may allow AFAP1L1 to play a role in normal cellular physiological processes involving actin dynamics and to also a play a role in the metastatic ability of cancer.

REFERENCES

Ammer, A.G., and Weed, S.A. (2008). Cortactin branches out: roles in regulating protrusive actin dynamics. Cell Motil Cytoskeleton *65*, 687-707.

Ayala, I., Baldassarre, M., Giacchetti, G., Caldieri, G., Tete, S., Luini, A., and Buccione, R. (2008). Multiple regulatory inputs converge on cortactin to control invadopodia biogenesis and extracellular matrix degradation. J Cell Sci *121*, 369-378.

Baisden, J.M., Qian, Y., Zot, H.M., and Flynn, D.C. (2001). The actin filament-associated protein AFAP-110 is an adaptor protein that modulates changes in actin filament integrity. Oncogene *20*, 6435-6447.

Berwick, M. (2006). Pathways to the development of melanoma: a complex issue. J Invest Dermatol *126*, 1932-1933.

Boeckers, T.M. (2006). The postsynaptic density. Cell Tissue Res *326*, 409-422.

Cailleau, R., Olive, M., and Cruciger, Q.V. (1978). Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. In Vitro *14*, 911-915.

Clump, D.A. (2008). A Genetic Variant of the Adaptor Protein, AFAP-110, Efficiently Activates cSrc, Resulting in Podosome Formation.

Dehal, P., and Boore, J.L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol *3*, e314.

DiNitto, J.P., and Lambright, D.G. (2006). Membrane and juxtamembrane targeting by PH and PTB domains. Biochim Biophys Acta *1761*, 850-867.

Dorfleutner, A., Cho, Y., Vincent, D., Cunnick, J., Lin, H., Weed, S.A., Stehlik, C., and Flynn, D.C. (2008). Phosphorylation of AFAP-110 affects podosome lifespan in A7r5 cells. J Cell Sci *121*, 2394-2405.

Dorfleutner, A., Stehlik, C., Zhang, J., Gallick, G.E., and Flynn, D.C. (2007). AFAP-110 is required for actin stress fiber formation and cell adhesion in MDA-MB-231 breast cancer cells. J Cell Physiol *213*, 740-749.

Flynn, D.C., Leu, T.H., Reynolds, A.B., and Parsons, J.T. (1993). Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. Mol Cell Biol *13*, 7892-7900.

Guappone, A.C., and Flynn, D.C. (1997). The integrity of the SH3 binding motif of AFAP-110 is required to facilitate tyrosine phosphorylation by, and stable complex formation with, Src. Mol Cell Biochem *175*, 243-252.

Hering, H., and Sheng, M. (2003). Activity-dependent redistribution and essential role of cortactin in dendritic spine morphogenesis. J Neurosci *23*, 11759-11769.

Linder, S. (2009). Invadosomes at a glance. J Cell Sci *122*, 3009-3013.

Lodyga, M., Bai, X.H., Kapus, A., and Liu, M. (2010). Adaptor protein XB130 is a Rac-controlled component of lamellipodia that regulates cell motility and invasion. J Cell Sci *123*, 4156-4169.

Lodyga, M., De Falco, V., Bai, X.H., Kapus, A., Melillo, R.M., Santoro, M., and Liu, M. (2009). XB130, a tissue-specific adaptor protein that couples the RET/PTC oncogenic kinase to PI 3-kinase pathway. Oncogene *28*, 937-949.

Maris, J.M. (2010). Recent advances in neuroblastoma. N Engl J Med *362*, 2202-2211.

Miller, C.R., and Perry, A. (2007). Glioblastoma. Arch Pathol Lab Med *131*, 397-406.

Ohno, S. (1970). Evolution by Gene Duplication. London, Allen and Unwin New York, Spring-Verlag.

Proszynski, T.J., Gingras, J., Valdez, G., Krzewski, K., and Sanes, J.R. (2009). Podosomes are present in a postsynaptic apparatus and participate in its maturation. Proc Natl Acad Sci U S A *106*, 18373-18378.

Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltham, M.*, et al.* (2000). Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet *24*, 227-235.

Sekino, Y., Kojima, N., and Shirao, T. (2007). Role of actin cytoskeleton in dendritic spine morphogenesis. Neurochem Int *51*, 92-104.

Tabakoff, B., Saba, L., Printz, M., Flodman, P., Hodgkinson, C., Goldman, D., Koob, G., Richardson, H.N., Kechris, K., Bell, R.L.*, et al.* (2009). Genetical genomic determinants of alcohol consumption in rats and humans. BMC Biol *7*, 70.

Walker, V.G., Ammer, A., Cao, Z., Clump, A.C., Jiang, B.H., Kelley, L.C., Weed, S.A., Zot, H., and Flynn, D.C. (2007). PI3K activation is required for PMA-directed activation of cSrc by AFAP-110. Am J Physiol Cell Physiol *293*, C119-132.

Weaver, A.M. (2008). Cortactin in tumor invasiveness. Cancer Lett *265*, 157-166.

Xu, J., Bai, X.H., Lodyga, M., Han, B., Xiao, H., Keshavjee, S., Hu, J., Zhang, H., Yang, B.B., and Liu, M. (2007). XB130, a novel adaptor protein for signal transduction. J Biol Chem *282*, 16401-16412.

Xu, X.Z., Garcia, M.V., Li, T.Y., Khor, L.Y., Gajapathy, R.S., Spittle, C., Weed, S., Lessin, S.R., and Wu, H. (2010). Cytoskeleton alterations in melanoma: aberrant expression of cortactin, an actin-binding adapter protein, correlates with melanocytic tumor progression. Mod Pathol *23*, 187-196.

Appendix A

A polymorphic variant of AFAP-110 enhances cSrc activity

David A. Clump MD/PhD[1+], Jing Jie Yu PhD[2], YoungJin Cho PhD[1#], Rui Gao MS[3], John Jett BA[4], Henry Zot PhD[5], Jess M. Cunnick PhD[6#], Brandi Snyder BS[2], Anne C. Clump BA[1], Melissa Shockey BA[1], Peter Gannett PhD[4], James E. Coad MD[7], Robert Shurina PhD[8], W. Douglas Figg PhD[3], Eddie Reed MD[9*] and Daniel C. Flynn PhD[1*#]

[1] The Mary Babb Randolph Cancer Center and the Department of Microbiology, Immunology and Cell Biology, West Virginia University, Morgantown, WV 26505-9300

[2] The Mary Babb Randolph Cancer Center and the Department of Biochemistry, West Virginia University, Morgantown, WV 26505-9300

[3] Molecular Pharmacology Section, Medical Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

[4] Department of Basic Pharmaceutical Sciences, West Virginia University, Morgantown, WV 26506-9530

[5] Department of Biology, University of West Georgia, Carrolton, GA 30118

[6] The Mary Babb Randolph Cancer Center and the Department of Pathology, West Virginia University, Morgantown, WV 26505-9300

[7] Department of Pathology, West Virginia University, Morgantown, WV 26505-9300

[8] Department of Biology, Wheeling Jesuit University, Wheeling, WV 26003

[9] Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, CoCHP, The Centers for Disease Control and Prevention, Atlanta, Georgia 30341

[+] Current address: University of Pittsburgh Medical Center, Radiation Oncology, 5230 Centre Avenue, Pittsburgh, Pa 15231

[#] Current address: The Commonwealth Medical College, 501 Madison Ave., Scranton, PA 18510

ABSTRACT

Enhanced expression and activity of cSrc is associated with ovarian cancer progression. Generally, cSrc does not contain activating mutations; rather its activity is increased in response to signals that effect a conformational change that releases its autoinhibition. In this report we analyzed ovarian cancer tissues for expression of a cSrc-activating protein, AFAP-110. AFAP-110 activates cSrc through a direct interaction that releases it from its autoinhibited conformation. Immunohistochemical analysis revealed a concomitant increase of AFAP-110 and cSrc in ovarian cancer tissues. An analysis of the AFAP-110 coding sequence revealed the presence of a nonsynonymous, single nucleotide polymorphism (SNP) that resulted in a change of Ser403 to Cys403. In cells that express enhanced levels of cSrc, AFAP-110^{403C} directed the activation of cSrc and the formation of podosomes independently of input signals, in contrast to wild-type AFAP-110. We therefore propose that under conditions of cSrc over-expression, the polymorphic variant of AFAP-110 promotes cSrc activation. Further, these data indicate a mechanism by which an inherited genetic variation could influence ovarian cancer progression and be used to predict the response to targeted therapy.

INTRODUCTION

Ovarian cancer, the most lethal gynecological malignancy, is characterized by tumor disruption of the ovarian capsule and dissemination and seeding of the pelvic and abdominal cavities (Naora and Montell, 2005). A combination of unreliable screening techniques, unspecific symptoms, and chemotherapy resistance results in 15,000 mortalities per year in the United States (Jemal et al., 2007). BRCA1 and BRCA2 are relevant for the disease and mutations of these genes are found in approximately 15% of ovarian cancer cases (Pal et al., 2005; Risch et al., 2006). However, the majority of cases consist of inconspicuous associations between inherited susceptibility and the environment. These associations may be explained by haplotype mapping studies, which predict that single nucleotide polymorphisms (SNPs) are not inherited independently, but instead associate with one another, as well as with environmental stimuli, producing the disease (Goldstein and Cavalleri, 2005). While SNPs that influence drug metabolism and cancer-related symptoms are described (Reyes-Gibby et al., 2008), little is known about genetic variants that modulate tumorigenesis. Identification of these genes may enhance our understanding of the progression of neoplasms such as ovarian cancer. In addition, these polymorphisms may serve as biomarkers that predict susceptibility to cancer or response to therapy.

One protein contributing to ovarian cancer progression is cSrc. This tyrosine kinase is over-expressed and activated in ovarian cancer cell lines and ovarian tumors (Wiener et al., 2003). cSrc promotes motility and invasion, the alteration of adhesion, and epithelial-mesenchymal transition (Frame, 2004; Yeatman, 2004). In addition, cSrc contributes to chemotherapy resistance, as inhibiting cSrc restores sensitivity to paclitaxel

(Chen et al., 2005; George et al., 2005). cSrc activation does not correlate with intrinsic mutations or SNPs; rather, signals from growth factors in the tumor microenvironment or intracellular activators of cSrc direct cSrc activation. A few cSrc activators have genetic variations that potentially modulate cSrc activity (Chen et al., 2005; George et al., 2005), and these may eventually serve as biomarkers useful for identifying the tumors most likely to respond to cSrc inhibition.

One cSrc activator, the actin-filament associated protein of 110 kDa (AFAP-110) is encoded by a polymorphic gene. The NCBI dsSNP database identifies a nonsynonymous C1210G coding substitution in exon 9 that predicts a serine to cysteine change at amino acid 403 (AFAP-110^{403C}) (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=60312). AFAP-110, via its intrinsic multimerization and a carboxy-terminal actin-binding domain, promotes actin filament cross-linking (Baisden et al., 2001; Qian et al., 2004). Additionally, AFAP-110 relays signals from PKCα that activate cSrc (Gatesman et al., 2004; Qian et al., 2002). These functions are autoinhibited by an intermolecular interaction between the carboxy-terminal leucine-zipper motif and an amino-terminal pleckstrin homology domain (PH1) (Qian et al., 2002; Qian et al., 2004). Upon experimental deletion of the leucine zipper domain (AFAP-110$^{\Delta Lzip}$) or upon PKCα activation, AFAP-110 is uninhibited and facilitates cSrc activation (Qian et al., 2002; Qian et al., 2004). This correlates with trafficking of activated cSrc to the cell membrane and the formation of the actin-rich invasive structures, podosomes (Gatesman et al., 2004; Walker et al., 2007).

To determine if AFAP-110 is positioned to activate cSrc in ovarian cancer, we performed immunohistochemical analysis on ovarian cancer tissues. In doing so, we

demonstrated that AFAP-110 exhibited a concomitant increase in expression with cSrc. Using PCR analysis, we discovered that a polymorphism of AFAP-110 was expressed in 1/4th of the population. This polymorphic variant, AFAP-110^{403C}, activated cSrc and triggered the formation of podosomes, suggesting that this variant of AFAP-110 may contribute to the progression of ovarian cancer.

MATERIALS AND METHODS

Reagents

The rabbit anti-human cortactin polyclonal antibody was purchased from Abcam (Cambridge, MA, USA). The AFAP-110 antibody F1 was previously characterized (Qian et al., 1999). The mouse anti-avian Src monoclonal antibody (clone EC10) was obtained from Upstate Biotechnology (Lake Placid, NY, USA). The rabbit anti-human cSrc monoclonal antibody (clone EG107) was from Novus Biologicals (Littleton, CO, USA). The rabbit anti-phospho-Src (Tyr416) monoclonal antibody (clone 100F9) was purchased from Cell Signaling (Danvers, MA) and the rabbit anti-phospho-Src (Tyr418) polyclonal antibody was from Biosource (Camarillo, CA, USA). Alexa Fluor secondary antibodies and fluorescently-labeled phalloidins were purchased from Molecular Probes. TRITC-phalloidin was purchased from Sigma (St Louis, MO, USA).

Cell culture

Mouse embryo fibroblasts (MEF), MEFs devoid of Src, Yes, and Fyn (SYF) and SYF cells re-expressing cSrc (SYF-cSrc) were obtained from the ATCC (Rockville, MD, USA). Cell lines were cultured in high glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 2 mM glutamine, 100 U/mL penicillin and 100 µg/mL streptomycin.

Study subjects and tissue samples

280 normal tissues and 124 serous papillary ovarian carcinomas in stage 3 or stage 4 were obtained from the Cooperative Human Tissue Network, Pediatric Division,

Children's Hospital, Columbus, OH, USA and from the WVU Pathology Department. Samples were collected prior to drug treatment and snap-frozen at -80°C until RNA/DNA extraction was performed. All specimens were diagnosed and classified by pathologists.

cDNA preparation and RT-PCR

From the studied specimens, total cellular RNA was isolated and purified by hot phenol/chloroform extraction. Purified RNAs were precipitated and dissolved in DEPC-treated water. Through reverse transcription, using the SuperScript Preamplification System, cDNAs were generated with oligo-dT primers from 5 µg of total RNA per sample (Reverse Transcription System, Promega, Madison, WI, USA). Exon 9 of the AFAP-110 gene which contains the SNP for 403C was amplified by the polymerase chain reaction (PCR). The primer set used for amplification contained the sequence of the 6 intron bases and the 20 exon bases that flank each end of exon 9 (CCGCAGGCTATCTGAACGTGCTCTCC   and TCCTACCTCCAATACTGCAACCTCCT). The PCR conditions were 95°C for 3 minutes flowed by 40 cycles of 95°C for 30 seconds, 58°C for 30 seconds and 72°C for 45 seconds; and then 1 cycle of 72°C for 10 minutes.

DNA sequencing and genotyping

PCR products were separated on agarose gel and purified using the QIAQuick Gel Extraction Kit (Qiagen, Valencia, CA, USA). The purified fragments were sequenced to identify AFAP-110 mutations using the CEQ 8000 Genetic Analysis System with

GenomeLab DTCS-Quick Start Kit (Beckman Coulter, Fullerton, CA, USA) and then using ABI Prism DNA Sequencer with BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) to confirm identified mutations. In addition, sequence variants were confirmed in duplicate independent PCR amplifications and sequencing reactions to insure that the mutations were not a result of PCR artifact.

Immunohistochemical methods

Immunohistochemistry was performed on 50 serous papillary ovarian carcinomas that were paraffin-embedded and cut into 5 μm-thick sections and mounted on positive-charge coated slides. Tissue sections were dried overnight in a 45º C oven, then deparaffinized, rehydrated, and subjected to heat-induced epitope retrieval for two hours in 1 mM citrate buffer (pH 6.00) in an 80º C water bath. Endogenous peroxide activity was blocked with 3% hydrogen peroxide and was followed by treatment with a serum-free protein blocker to block nonspecific binding. Following each step of the immunoreaction except the protein blocker, sections were rinsed in Tris-buffered saline with Tween. Tissues were incubated for two hours with anti-AFAP-110 antibody (F1, 6 μg/mL) at a dilution of 1:10 or with anti-cSrc (EG107, diluted 1:50) in 10% normal horse serum. Negative controls (i.e., preimmune serum or normal rabbit IgG) were incubated in 10% normal horse serum. Sections were incubated with biotinylated secondary link antibody, followed by treatment with streptavidin peroxidase and stained with diaminobenzidine substrate-chromogen solution. Counterstaining was performed with hematoxylin, followed by a water rinse and bluing solution. Tissues were then dehydrated and coverslipped.

Plasmid constructs

Mutagenesis was performed on human AFAP-110 to generate AFAP-110[403C]. AFAP-110 and AFAP-110[403C] were cloned into pEGFP-C3 (Clontech, Mountain View, CA, USA) as previously described (Qian et al., 2002). The pGEX-6P-1 vector from Amersham Pharmacia Biotech was used to create fusion proteins expressing the PH2 domain of AFAP as previously described (Qian et al., 2004). The pGexX-6P-1 PH2 [403C] construct was created by PCR cloning the PH2[403C] fragment from the pEGFP vector with BamHI and EcoRI ends and subsequent cloning into the pGEX-6P-1 vector. Human AFAP-110 wild-type and AFAP110[403C] cDNA were PCR amplified and ligated into pFLAG CMV vector using EcoRI and XhoI sites to generate Flag-tagged AFAP-110.

Immunofluorescence

Transfection of either GFP AFAP-110 or GFP AFAP-110[403C] into MEF, SYF, and SYF-cSrc cells for immunofluorescence was carried out using either Lipofectamine PLUS (Invitrogen, Carlsbad, CA, USA) or Nucleofector (Amaxa, Walkersville, MD, USA) according to the manufacturer's specification. Cells were plated on glass coverslips immediately after transfection or allowed to recover for 24 hr after transfection and then plated on glass coverslips coated with 10 μg/ml fibronectin (BD Bioscience, San Jose, CA, USA). Cells were serum starved for 12 hours prior to fixation or left untreated as indicated in the results section. Fixation, permeabilization, and staining procedures, including antibody dilutions, were performed as previously described (Gatesman et al., 2004). Confocal images were acquired with a Zeiss LSM 510 microscope with an average slice thickness of 1 μm. Fluorescence channels were sequentially recorded using

the multi-track recording module. Fluorescence images were obtained with a Zeiss Axiovert 200M microscope. Images were subsequently analyzed with LSM 510 software, Adobe Photoshop, and Image J (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Md, USA, http://rsb.info.nih.gov/ij/, 1997-2007).

Quantification of podosomes

Actin-rich structures at the ventral surface of cells were identified as podosomes if three podosome markers: AFAP-110, cortactin, and phospho-cSrc colocalized. 250 to 300 cells from two independent experiments were analyzed to determine the percentage of cells forming podosomes. The number of podosomes/cell was also counted from at least 100 podosome-positive cells and the distribution of podosomes/cell was plotted. Student's *t*-test with Bonferroni adjustment was used to detect the statistical significance.

Western Blot analysis and cSrc activation assay

Cultures or human ovarian tissue were lysed or homogenized in radioimmunoprecipitation assay buffer (RIPA) (50mM TrisHCl, pH7.4, 150mM NaCl, 2mM EDTA, 1% NP-40) containing leupeptin, aprotinin, sodium vanadate, EGTA, and phenylmethylsulfonyl fluoride (PMSF). Protein concentration was determined by the BCA assay (Pierce) and equal amount of proteins were resolved on SDS-PAGE and transferred to PVDF. SYF cells were transfected with either Flag-AFAP-110 or Flag-AFAP-110 [403C] and cSrc using Lipofectamine PLUS (Invitrogen). 48 hours after transfection, cells were stimulated and lysed in RIPA buffer. The levels of phosphorylated Y416 cSrc and total cSrc were assessed in the subsequent western blot

analysis. Antibody dilutions were performed according to previously described protocols (Gatesman et al., 2004).

Molecular Modeling

The PH domain of SKAP-hom (PDB ID#1U5E), a SKAP55 homologue and a Src-associated adaptor protein, was used as a structural template to create a homology model of the AFAP PH2-WT domain using the homology module of Insight II. SKAP-hom was chosen as the best template based on a BLAST server analysis and fold recognition programs such as PHYRE, 123D+, and FUGUE. A sequence alignment was created using the Clustal W server and manual adjustments were made by integrating secondary structure prediction data for the PH2–WT domain from the PROF server with the known secondary structure of SKAP-hom. The model was then minimized with 1000 steps of steepest descent minimization. The structure was equilibrated for 300ps using explicit water molecules as solvent. Equilibration was completed using the sander command of Amber 8. The model was analyzed to ensure no mis-folded regions existed using the Profiles3D program in Insight II. The sander command of Amber 8 was then used to create a 700ps trajectory of the protein domain in explicit water molecules and a hydrogen bond analysis for PH2^{403S} was completed using the default parameters of the ptraj hbond command in Amber 8. This process was repeated for PH2^{403C}.

Lipid dot-blot

The lipid dot-blot was used to detect interactions between soluble GST proteins and phospholipids immobilized on nitrocellulose. Lipid spotted membranes (PIP strips,

Echelon Biosciences, Salt Lake City, UT, USA) were blocked 1 hour with 0.2% BSA in TBS and incubated overnight with 0.5 mg/ml recombinant GST protein in TBS. Bound GST protein was detected with rabbit anti-GST antibody (Sigma-Aldrich, St. Louis, MO, USA).

Preparation of Lipid Vesicles

Large unilamellar lipid vesicles (LUV) were prepared by the extrusion method using 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and the indicated lipid (Avanti polar Lipids, Alabaster, AL, USA). Phospholipids were combined by molar ratio (10% indicated lipid, 90% POPC or 100% POPC) in chloroform: methanol: water (60: 30: 4), dried with a stream of N_2, and evacuated to remove traces of solvent. The residue was hydrated with Buffer B (50 mM HEPES, pH 7.2, 80 mM KCl, and 3 mM EGTA) to attain 2 mM lipid sheets when resuspended by vortex. Samples were subjected to 10 cycles of freeze thaw and then 10 passages through two layers of 0.10 μm polycarbonate filters under high pressure N_2. The resulting LUV were used directly for sedimentation assays.

Sedimentation assay for PtdIns lipids

To detect pleckstrin homology domain protein associated with membrane phospholipids, a sedimentation assay was used. Samples containing 1.7 mM vesicle lipid and 0.025 mM recombinant GST fusion protein (PH1) were prepared in 150 μl binding buffer. Samples were incubated 60 min and centrifuged with a Beckman Airfuge for 15 min at room temperature. From each centrifuged sample, 16% of the supernatant and 100% of the

pellet were analyzed by SDS-PAGE. Gels were stained with SPYRO orange (Invitrogen

Corp., Carlsbad, CA, USA).

RESULTS

AFAP-110 and cSrc are over-expressed in ovarian cancer

Immunohistochemistry (IHC) determined that AFAP-110 and cSrc were over-expressed in 30/33 and 32/33 ovarian cancer samples, respectively (Supplemental Table 1). In 60% of the samples, AFAP-110 was over-expressed focally (Figure 1A and B) while in 86% of samples cSrc expression was diffuse (Figure 1C and D). AFAP-110 expression in blood vessels is observed in Figure 1B.

AFAP-110 and cSrc were not detected in normal ovaries (Figure 2, panels A-C), although blood vessels exhibited AFAP-110 expression (Figure 2, panel A). AFAP-110 was always expressed with cSrc in well-differentiated tumors (Figure 2, panels D-F). Further, AFAP-110 and cSrc always co-localized in desmoplastic regions (Figure 2, panels G-I). AFAP-110 is over-expressed in undifferentiated tumor specimens (13 specimens analyzed), but co-expression of AFAP-110 and cSrc was observed in only four of these specimens (Figure 2, panels J-L).

A SNP in AFAP-110

Genetic variation within AFAP-110 was examined. Analysis of AFAP-110 cDNA isolated from the ovarian cancer cell lines OvCAR-3 and 2008 revealed a synonymous SNP (G297A, CCG to CCA) not found in the A2780, MCAS or SKOV3 lines. A nonsynonymous SNP (C1210G, TCT to TGT) identified in exon 9 resulted in a serine to cysteine (S403C) substitution and was also found in the OVCAR-3 and 2008 cell lines.

Ovarian tissues were screened to determine the prevalence of this nonsynonymous SNP, C1210G.   C1210G was identified in 19/91 (20.9%) tumor samples and 9/41 (21.2%) tumor-adjacent normal tissues (Supplemental Table 2).   In addition, the 33 samples used for IHC were analyzed and revealed the SNP in 9/33 samples (27.3%).

Tissue not associated with cancer was obtained to determine if the SNP was enriched in ovarian cancer.   C1210G was present in 80/280 normal tissues (28.6%) (Supplemental Table 2).   Thus, AFAP-110[403C] was not enriched in ovarian cancer.

We analyzed ovarian tumor tissues and adjacent normal tissue to confirm that cSrc was over-expressed in tissues that have either wild type AFAP-110 or AFAP-110[403C].   For comparison, we analyzed the intensity of cSrc expression in SYF cells, which have no Src family kinases (Klinghoffer et al., 1999), and SYF-cSrc cells, which over-express cSrc.   Western blotting revealed that cSrc expression in tumors approximated that detected in SYF-cSrc cells (Figure 3A). An antibody that detected both forms of AFAP-110 indicated a reduction of AFAP-110 expression in adjacent normal tissues confirming the over-expression in tumors observed by IHC (Figure 3B). Over-expression of cSrc in ovarian tumor samples was also confirmed.

403C is located within the PH2 domain

The serine to cysteine (S403C) substitution is located in the second PH domain (PH2) of AFAP-110. PH domains are characterized by seven β-strands that are linked by loop regions. These loop regions differ in length and amino acid composition and therefore confer the differential binding of PH domains to lipids (Haslam et al., 1993; Hyvonen et al., 1995; Lemmon et al., 1995; Macias et al., 1994; Mayer et al., 1993;

Musacchio et al., 1993; Yoon et al., 1994). The S403C substitution occurs in a loop region between the 5th and 6th β-strand (Figure 4). In AFAP-110, the hydroxyl R-group of Ser[403] is predicted to interact with water 71% of the time (Figure 4 panel B). However, in AFAP-110[403C], the sulfhydryl R-group of Cys[403] is not predicted to be an efficient hydrogen binding partner (36%, Figure 4 panel D), indicating that the PH[403C] domain may exhibit differential binding specificity secondary to changes in structural flexibility.

PH domains participate in protein-protein interactions. For example, PH domains in pleckstrin and in AFAP-110 (PH1) bind to PKCα (Abrams et al., 1995; Qian et al., 2002). Further, the PH1 domain of AFAP-110 binds to AFAP-110, stabilizing the AFAP-110 multimer (Qian et al., 2004). To determine differential binding between AFAP-110 and AFAP-110[403C], affinity precipitation assays using GST-fusion proteins were performed. GST-PH2 was more efficient than GST-PH2[403C] in pulling down AFAP-110 (Figure 5). CaOV3 cell lysates were used as a negative control, as they have AFAP-110 expression levels that are at or below detection limits (Gatesman et al., 2004). These data indicated that the S403C change may reduce the affinity and therefore the ability of the PH2 domain to bind to AFAP-110. Pull-down assays using GST-PH2 did not reveal other binding partners (data not shown).

PH domains are also known to interact with lipids; however, an analysis of the AFAP-110 PH2 domain demonstrated that it does not contain the conserved basic residues that mediate electrostatic interactions with negatively charged phospholipids (Figure 4, panels B and D). A lipid dot- blot analysis as well as a lipid vesicle sedimentation assay was used to determine if the GST-PH2 fusion protein would bind

phosphoinositides (Supplemental Figure 1A-B and Supplemental Methods). Unlike the

positive control, the PH domain from DAPP1 (Dowler et al., 1999), GST-PH2 did not

bind to phosphoinositides.

Effects of AFAP-110^{403C} on actin filament modulation.

Affinity precipitation data indicated that the PH2 domain mediates self-

association. Earlier work demonstrated that intermolecular interactions that stabilize self-

association had an autoinhibitory effect on AFAP-110 (Qian et al., 2004). As GST-

PH2^{403C} is less-efficient in binding to AFAP-110 than GST-PH2 and destabilization of

the AFAP-110 multimer correlated with an acquired ability to activate cSrc, we sought to

determine if AFAP-110^{403C} had the capacity to activate cSrc.

Co-transfection of Flag-tagged AFAP-110 or Flag-tagged AFAP-110^{403C} with

cSrc into SYF cells confirmed that AFAP-110^{403C} was able to direct cSrc activation in

contrast to wild type AFAP-110 (Figure 6A). The SYF cell lines allowed us to determine

the effect of AFAP-110 on cSrc activity in the absence of other Src family members. This

data indicated that AFAP-110^{403C} can activate cSrc in cells under conditions of dual over-

expression.

To further examine the effect of 403C on the function of AFAP-110, fibroblasts

expressing varying levels of cSrc were used. SYF cells do not express cSrc, MEFs

express a low level, while SYF-cSrc cells express a relatively high level of cSrc (Figure

6B). Neither GFP-AFAP-110 or GFP-AFAP-110^{403C} affect detectable changes in cellular

morphology in SYF (data not shown) nor MEF cells (Figure 7A). Anti-phospho

cSrcY416 which recognizes phosphorylated tyrosine 416 in active cSrc was used to

assess cSrc activity. Active cSrc was undetectable in both SYF and MEF cells. However, in SYF-cSrc cells, GFP-AFAP-110^{403C} directed cSrc activation and the formation of podosomes (Figure 7B). Podosome formation was confirmed based on the co-localization of AFAP-110, actin and cortactin in punctate structures on the ventral surface of the cells (Linder and Aepfelbacher, 2003) (Figure 7C). By quantifying the number of cells exhibiting podosomes and the number of podosomes/cell, we determined that podosome formation was strongly associated with expression of AFAP-110^{403C} in SYF-cSrc cells, and that cells expressing AFAP-110^{403C} had significantly more podosomes/cell than that of cells expressing only endogenous or over-expressed AFAP-110 (Figure 8).

DISCUSSION

Ovarian cancer results from a combination of inherited and acquired genetic alterations as well as from environmental influences. Detection is limited because of inadequate screening and nonspecific symptoms. Since this leads to a delayed diagnosis, ovarian cancer is the most lethal gynecological malignancy. This creates interest in identifying biomarkers that stratify patients into high risk subgroups, as well as potentially guide the development of individualized therapy. The present study demonstrates that AFAP-110[403C] results in activation of cSrc under conditions of over-expression. Therefore, we hypothesize that the presence and expression levels of AFAP-110[403C] may have value in predicting risk and treatment strategies for ovarian cancer.

AFAP-110 functions as an actin filament cross-linking protein and an adaptor protein that relays signals from PKCα that activate cSrc (Chen et al., 1985; Dorfleutner et al., 2007; Kanner et al., 1991; Qian et al., 2002; Qian et al., 2004). Activated cSrc leads to an increase in cell motility and the production of podosomes, which may be precursors to invadopodia (Gatesman et al., 2004; Walker et al., 2007). As podosome formation requires both cSrc activation and dynamic changes in actin filament integrity, AFAP-110 may be uniquely positioned to regulate these two cellular signals. This ability to activate cSrc and contribute to the formation of invasive structures may be relevant for cancer progression (Flynn, 2008).

As AFAP-110 is a cSrc activating protein, this report determined if its expression or its genetic variant AFAP-110[403C] was associated with ovarian cancer. IHC demonstrated that both AFAP-110 and cSrc were over-expressed in ovarian cancer. In general, cSrc expression was diffuse, while AFAP-110 expression was focal.

Interestingly, both AFAP-110 and cSrc were always over-expressed together in well-differentiated tumors and in desmoplastic regions of the tumor. Co-localization in desmoplastic areas may represent dynamic interactions between the host and the invasive tumor. As cSrc activation correlates with acquisition of the invasive phenotype (Summy and Gallick, 2006), we hypothesized that AFAP-110 is positioned to activate cSrc and therefore promote invasion in these discrete areas.

AFAP-110 was scanned for genetic changes in human ovarian cancer cell lines by isolating the cDNA of AFAP-110. A SNP was identified that affected a nonsynonymous coding change at base pair 1210, changing $Ser^{403} \rightarrow Cys^{403}$. Additionally, 124 ovarian tumors were analyzed and determined to contain the SNP in 22% of the samples. Adjacent, normal ovarian tissue as well as that obtained from women with no known history of malignancy revealed a similar SNP profile. Thus, although expression levels of AFAP-110 were elevated in ovarian cancer, the presence of the SNP was not enriched in tumors. Therefore, this study focused on determining if high expression levels of AFAP-110 or AFAP-110[403C] affected differences in cSrc activation or cell morphology.

Serine 403 of AFAP-110 is positioned on the loop between the fifth and sixth β-strands of the PH2 domain. Molecular modeling indicated that the peripherally positioned R-hydroxyl group of Ser^{403} forms hydrogen bonding with H_2O 71% of the time, while the Cys^{403} forms hydrogen bonding with H_2O less efficiently (36%). Hydrogen bonds may stabilize the loops of the PH domain and increase the rigidity in the binding pocket.

There are as many as 258 different proteins that contain PH domains (Lemmon et al., 2002; McPherson et al., 2001). Of these, only 10% are predicted to facilitate phospholipid binding by forming interactions between positively charged Lys or Arg

residues within the PH binding pockets and the negatively charged phospholipids. The PH2 domain of AFAP-110 is unable to bind phospholipid.

PH domains also bind to proteins. Affinity precipitation revealed that similar to the PH1 domain, the PH2 domain also bound AFAP-110. Thus, it is predicted that the PH2 domain may foster either intramolecular or intermolecular interactions, promoting multimerization or stabilization of the multimer. However, GST-PH2^{403C} bound less efficiently to AFAP-110. Loss of AFAP-110 multimer stability correlated with a gain-of-function, including an ability to colocalize and activate cSrc (Gatesman et al., 2004; Qian et al., 2004; Walker et al., 2007). Thus, we sought to determine if AFAP-110^{403C} activates cSrc.

Expression of AFAP-110 or AFAP-110^{403C} in MEF cells, which contain detectable levels of cSrc – estimated around 20,000 molecules per cell (Flynn unpublished data) did not result in cSrc activation or morphological changes that are associated with cSrc activation. Ectopically expressed AFAP-110 or AFAP-110^{403C} in SYF-cSrc cells, which over-express cSrc at levels that are higher than MEFs, was used to determine if AFAP-110^{403C} has a differential capacity to activate cSrc. Compared to AFAP-110, AFAP-110^{403C} was a more efficient activator of cSrc and more efficient in inducing the formation of podosomes. Thus these data indicate that in cells over-expressing cSrc, AFAP-110^{403C} more efficiently activates cSrc compared to wild type AFAP-110. Previous evidence suggested that in chicken embryo fibroblasts transformed by the Rous sarcoma virus only 5% of AFAP-110 is complexed with v-Src (Kanner et al., 1991). Since only a subset of AFAP-110 and Src interact, it is possible that in MEF cells only 5% or less of the cSrc population that is expressed would be engaged with AFAP-

110^{403C} and that this stoichiometry of binding could be below detection. Further, if 5% of cSrc were activated in MEF cells, activation may be insufficient to direct morphological changes characteristic of Src-transformed cells. Thus, under conditions where both cSrc and AFAP-110^{403C} expression is low, as in normal tissues, cSrc activation is unlikely to occur or affect cellular changes even in those cells that inherit AFAP^{403C}. Under conditions of dual over-expression of AFAP-110^{403C} and cSrc, AFAP-110^{403C} may independently activate cSrc and promote tumor progression. Since AFAP-110 and cSrc are over-expressed in the same tumors, AFAP-110 may enhance cSrc activation by receiving input signals that enable cSrc activation or, alternatively, AFAP-110^{403C} may result in a reduced capacity to self-associate resulting in the independent activation of cSrc. Although not enriched in ovarian cancer tumors, the 403C variant may lead to a more aggressive and metastatic disease through its promotion of cSrc activation in those tumors in which it is found. Future studies using ovarian cancer cell lines and additional patient samples should address this issue. These data also indicate a mechanism by which an inherited genetic variation could influence ovarian cancer progression and be used to predict the response to targeted therapy.

ACKNOWLEDGEMENTS

REFERENCES

Abrams, C.S., Zhao, W., Belmonte, E., and Brass, L.F. (1995). Protein kinase C regulates pleckstrin by phosphorylation of sites adjacent to the N-terminal pleckstrin homology domain. J Biol Chem *270*, 23317-23321.

Baisden, J.M., Gatesman, A.S., Cherezova, L., Jiang, B.H., and Flynn, D.C. (2001). The intrinsic ability of AFAP-110 to alter actin filament integrity is linked with its ability to also activate cellular tyrosine kinases. Oncogene *20*, 6607-6616.

Chen, T., Pengetnze, Y., and Taylor, C.C. (2005). Src inhibition enhances paclitaxel cytotoxicity in ovarian cancer cells by caspase-9-independent activation of caspase-3. Mol Cancer Ther *4*, 217-224.

Chen, W.T., Chen, J.M., Parsons, S.J., and Parsons, J.T. (1985). Local degradation of fibronectin at sites of expression of the transforming gene product pp60src. Nature *316*, 156-158.

Dorfleutner, A., Stehlik, C., Zhang, J., Gallick, G.E., and Flynn, D.C. (2007). AFAP-110 is required for actin stress fiber formation and cell adhesion in MDA-MB-231 breast cancer cells. J Cell Physiol *213*, 740-749.

Dowler, S., Currie, R.A., Downes, C.P., and Alessi, D.R. (1999). DAPP1: a dual adaptor for phosphotyrosine and 3-phosphoinositides. Biochem J *342 (Pt 1)*, 7-12.

Flynn, D.C., Cho, YJ., Vincent, D., Cunnick, J. M. (2008). Podosomes and Invadopodia: Related structures

with Common Protein Components that May Promote Breast

Cancer Cellular Invasion. Breast Cancer: Basic and Clinical Research *2*, 17-29.

Frame, M.C. (2004). Newest findings on the oldest oncogene; how activated src does it. J Cell Sci *117*, 989-998.

Gatesman, A., Walker, V.G., Baisden, J.M., Weed, S.A., and Flynn, D.C. (2004). Protein kinase Calpha activates c-Src and induces podosome formation via AFAP-110. Mol Cell Biol *24*, 7578-7597.

George, J.A., Chen, T., and Taylor, C.C. (2005). SRC tyrosine kinase and multidrug resistance protein-1 inhibitions act independently but cooperatively to restore paclitaxel sensitivity to paclitaxel-resistant ovarian cancer cells. Cancer Res *65*, 10381-10388.

Goldstein, D.B., and Cavalleri, G.L. (2005). Genomics: understanding human diversity. Nature *437*, 1241-1242.

Haslam, R.J., Koide, H.B., and Hemmings, B.A. (1993). Pleckstrin domain homology. Nature *363*, 309-310.

Hyvonen, M., Macias, M.J., Nilges, M., Oschkinat, H., Saraste, M., and Wilmanns, M. (1995). Structure of the binding site for inositol phosphates in a PH domain. EMBO J *14*, 4676-4685.

Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., and Thun, M.J. (2007). Cancer statistics, 2007. CA Cancer J Clin *57*, 43-66.

Kanner, S.B., Reynolds, A.B., and Parsons, J.T. (1991). Tyrosine phosphorylation of a 120-kilodalton pp60src substrate upon epidermal growth factor and platelet-derived growth factor receptor stimulation and in polyomavirus middle-T-antigen-transformed cells. Mol Cell Biol *11*, 713-720.

Klinghoffer, R.A., Sachsenmaier, C., Cooper, J.A., and Soriano, P. (1999). Src family kinases are required for integrin but not PDGFR signal transduction. EMBO J *18*, 2459-2471.

Lemmon, M.A., Ferguson, K.M., and Abrams, C.S. (2002). Pleckstrin homology domains and the cytoskeleton. FEBS Lett *513*, 71-76.

Lemmon, M.A., Ferguson, K.M., O'Brien, R., Sigler, P.B., and Schlessinger, J. (1995). Specific and high-affinity binding of inositol phosphates to an isolated pleckstrin homology domain. Proc Natl Acad Sci U S A *92*, 10472-10476.

Linder, S., and Aepfelbacher, M. (2003). Podosomes: adhesion hot-spots of invasive cells. Trends Cell Biol *13*, 376-385.

Macias, M.J., Musacchio, A., Ponstingl, H., Nilges, M., Saraste, M., and Oschkinat, H. (1994). Structure of the pleckstrin homology domain from beta-spectrin. Nature *369*, 675-677.

Mayer, B.J., Ren, R., Clark, K.L., and Baltimore, D. (1993). A putative modular domain present in diverse signaling proteins. Cell *73*, 629-630.

McPherson, J.D., Marra, M., Hillier, L., Waterston, R.H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E.R., Wilson, R.K., *et al.* (2001). A physical map of the human genome. Nature *409*, 934-941.

Musacchio, A., Gibson, T., Rice, P., Thompson, J., and Saraste, M. (1993). The PH domain: a common piece in the structural patchwork of signalling proteins. Trends Biochem Sci *18*, 343-348.

Naora, H., and Montell, D.J. (2005). Ovarian cancer metastasis: integrating insights from disparate model organisms. Nat Rev Cancer *5*, 355-366.

Pal, T., Permuth-Wey, J., Betts, J.A., Krischer, J.P., Fiorica, J., Arango, H., LaPolla, J., Hoffman, M., Martino, M.A., Wakeley, K., *et al.* (2005). BRCA1 and BRCA2 mutations account for a large proportion of ovarian carcinoma cases. Cancer *104*, 2807-2816.

Qian, Y., Baisden, J.M., Cherezova, L., Summy, J.M., Guappone-Koay, A., Shi, X., Mast, T., Pustula, J., Zot, H.G., Mazloum, N., *et al.* (2002). PC phosphorylation increases the ability of AFAP-110 to cross-link actin filaments. Mol Biol Cell *13*, 2311-2322.

Qian, Y., Gatesman, A.S., Baisden, J.M., Zot, H.G., Cherezova, L., Qazi, I., Mazloum, N., Lee, M.Y., Guappone-Koay, A., and Flynn, D.C. (2004). Analysis of the role of the leucine zipper motif in regulating the ability of AFAP-110 to alter actin filament integrity. J Cell Biochem *91*, 602-620.

Qian, Y., Guappone, A.C., Baisden, J.M., Hill, M.W., Summy, J.M., and Flynn, D.C. (1999). Monoclonal antibodies directed against AFAP-110 recognize species-specific and conserved epitopes. Hybridoma *18*, 167-175.

Reyes-Gibby, C.C., Wu, X., Spitz, M., Kurzrock, R., Fisch, M., Bruera, E., and Shete, S. (2008). Molecular epidemiology, cancer-related symptoms, and cytokines pathway. Lancet Oncol *9*, 777-785.

Risch, H.A., McLaughlin, J.R., Cole, D.E., Rosen, B., Bradley, L., Fan, I., Tang, J., Li, S., Zhang, S., Shaw, P.A., *et al.* (2006). Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: a kin-cohort study in Ontario, Canada. J Natl Cancer Inst *98*, 1694-1706.

Summy, J.M., and Gallick, G.E. (2006). Treatment for advanced tumors: SRC reclaims center stage. Clin Cancer Res *12*, 1398-1401.

Walker, V.G., Ammer, A., Cao, Z., Clump, A.C., Jiang, B.H., Kelley, L.C., Weed, S.A., Zot, H., and Flynn, D.C. (2007). PI3K activation is required for PMA-directed activation of cSrc by AFAP-110. Am J Physiol Cell Physiol *293*, C119-132.

Wiener, J.R., Windham, T.C., Estrella, V.C., Parikh, N.U., Thall, P.F., Deavers, M.T., Bast, R.C., Mills, G.B., and Gallick, G.E. (2003). Activated SRC protein tyrosine kinase is overexpressed in late-stage human ovarian cancers. Gynecol Oncol *88*, 73-79.

Yeatman, T.J. (2004). A renaissance for SRC. Nat Rev Cancer *4*, 470-480.

Yoon, H.S., Hajduk, P.J., Petros, A.M., Olejniczak, E.T., Meadows, R.P., and Fesik, S.W. (1994). Solution structure of a pleckstrin-homology domain. Nature *369*, 672-675.

This page intentionally left blank

Figure 1
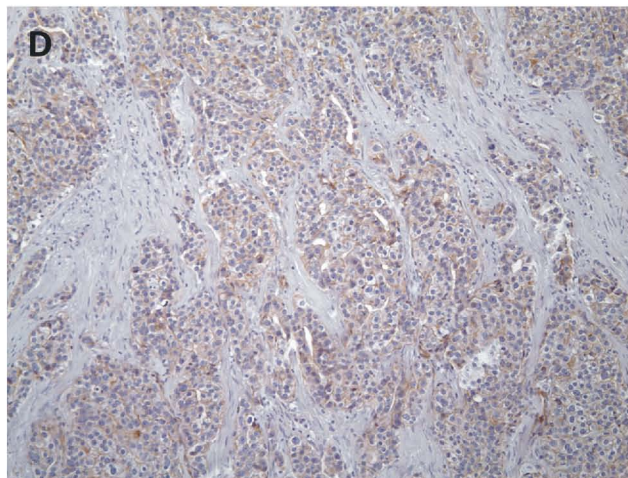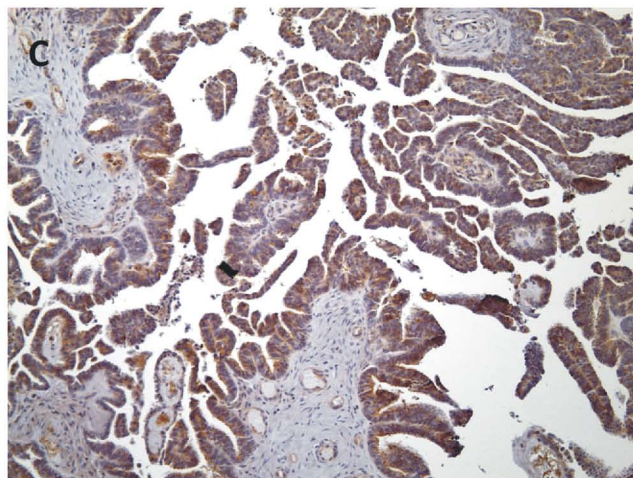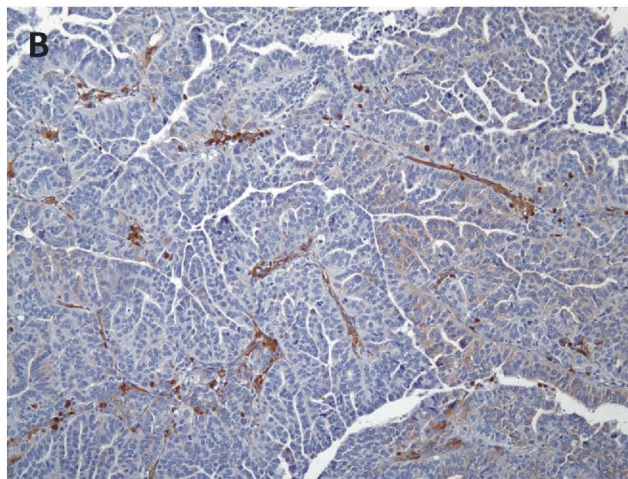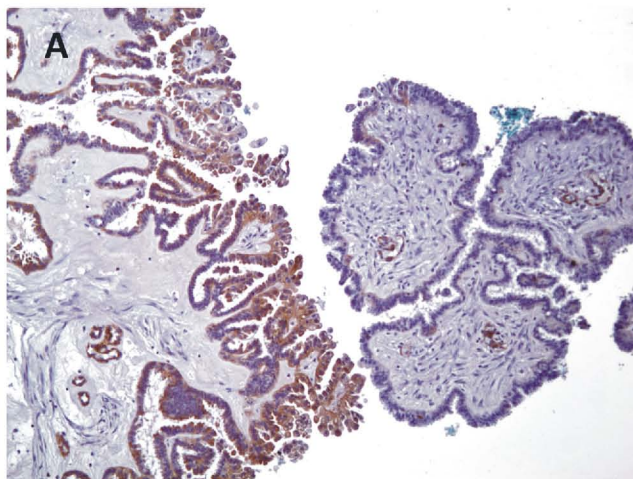
Figure 1. Focal and diffuse expression patterns of AFAP-110 and cSrc

Ovarian cancer tissues were sectioned and immunohistochemistry performed with anti-AFAP-110 (pAb F1) or anti-cSrc antibody (monoclonal antibody EG107 and the intensity of immunolabeling (brown color) qualitatively assessed by a pathologist. Samples that show focal (highly localized) or diffuse immunolabeling that exhibit deep, robust or weak immunostaining are shown for comparison and these types of images represent the assessment of tissue immunostaining shown in Supplemental Table 1. (A) AFAP-110 expressed strongly in focal areas of the tumor, (B) AFAP-110 expressed weakly in focal areas of the tumor, (C) cSrc expressed strongly and diffusely throughout the tumor or (D) cSrc expressed weakly and diffusely throughout the tumor.
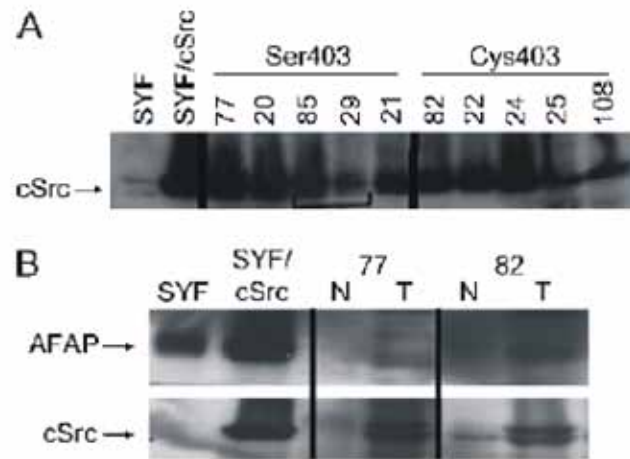
Figure 2
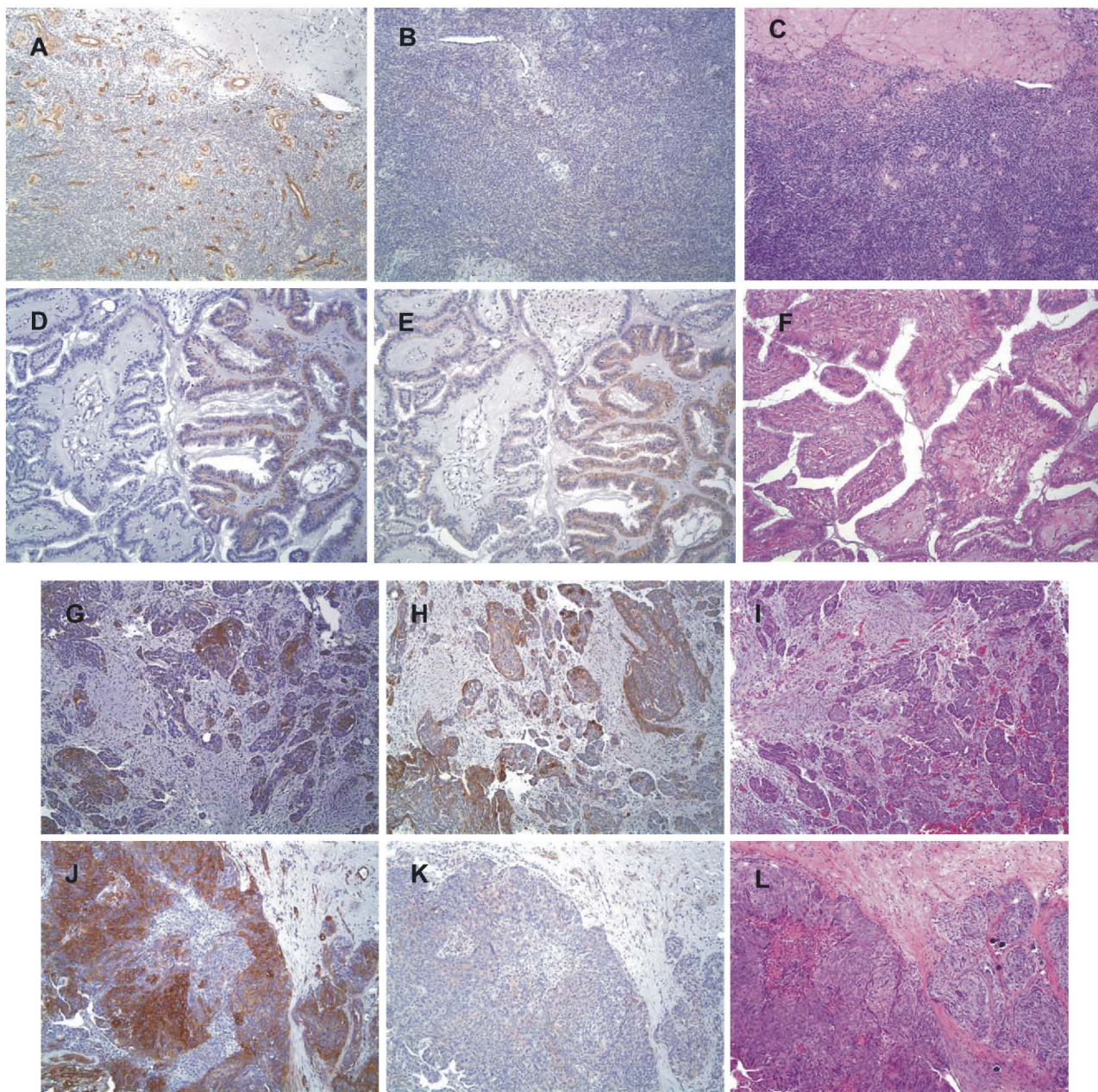
Figure 2. Co-expression of AFAP-110 and cSrc in ovarian tumors

Serial sections of normal human ovary tissues, well differentiated, and undifferentiated human ovary tissues were immunolabeled for AFAP with pAb F1, cSrc antibody EG107, and H&E staining as described under Material and Methods. (A) Normal ovary/AFAP-110; (B) normal ovary/cSrc; (C) normal ovary/H&E; (D) well-differentiated tumor/AFAP-110; (E) well-differentiated tumor/cSrc; (F) well-differentiated tumor/H&E; (G) ovarian tumor/AFAP-110; (H) ovarian tumor/cSrc; (I) ovarian tumor/H&E; (J) undifferentiated tumor/AFAP-110; (K) undifferentiated tumor/cSrc; (L) undifferentiated tumor/H&E. In comparison to normal human ovaries (A-C), AFAP-110 and cSrc protein are over-expressed and colocalized in well-differentiated tumors (D-F), as well as desmoplastic regions of human ovarian tumors (G-I). AFAP-110 is over-expressed in undifferentiated tissue specimens (J-L); however, colocalization with cSrc was variable. All images were captured at a magnification of 100x. As stated in Supplementary table 1, 33 ovarian tumors were analyzed. Of these, 20 were differentiatied tumors and 13 were undifferentiated. Six normal ovaries were analyzed.
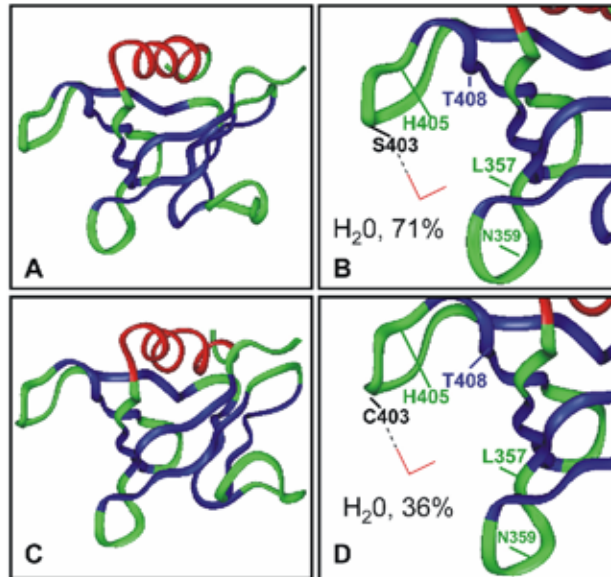
Figure 3

Figure 3. cSrc expression level in ovarian tumors match those levels detected in SYF-cSrc cells

(A) 50 µg of SYFor SYF-cSrc cellular lysates or human ovarian tissue lysates were resolved by 8% SDS-PAGE, transferred to PVDF and western blot analysis performed with anti-cSrc antibodies. Code numbers for de-identified patient samples are shown. Five samples had Ser^{403} encoded on at least one allele, and five samples had Cys^{403} encoded on at least one allele. (B) cSrc and AFAP-110 expression levels are increased in ovarian tumors relative to normal, adjacent tissue. Western blot analysis with antibodies to AFAP-110 or cSrc of two ovarian tumor samples (T) and matching, adjacent control tissues (N) from two patient samples (#77 and #82). Patient #77 has the AFAP-110 Ser^{403} wild-type isoform, while patient # 82 has the AFAP-110 Cys^{403} SNP on at least one allele.
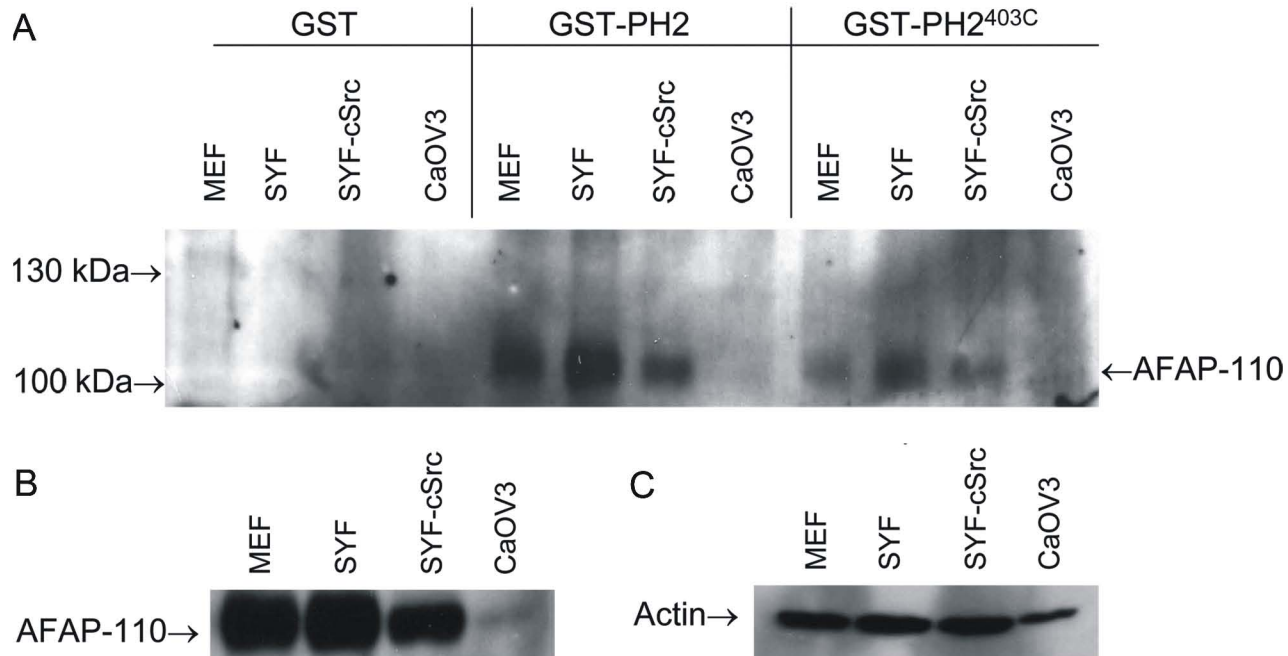
Figure 4

Figure 4. Molecular modeling of the PH2 domain

Homology models for AFAP-110 PH2-WT (A and B) and AFAP-110 PH2^{403C} (C and D) demonstrated that the amino acid change occurs in a loop region between the 5th and 6th β-strand. Performing a hydrogen bond (dashed black lines) analysis for each structure predicted that AFAP-110 PH2-WT binds to water molecules (solid red lines) 71% of the time potentially forming a rigid binding region. In contrast, AFAP-110 PH2^{403C} was predicted to bind to water only 36% of the time. Labeled amino acids occur at coordinates predicted to interact with phospholipid head groups. Intrastrand loops: green; β-strands: blue; α-helix: red.

Figure 5

A

GST | GST-PH2 | GST-PH2^{403C}

MEF | SYF | SYF-cSrc | CaOV3 | MEF | SYF | SYF-cSrc | CaOV3 | MEF | SYF | SYF-cSrc | CaOV3

130 kDa→

100 kDa→

←AFAP-110

B

MEF | SYF | SYF-cSrc | CaOV3

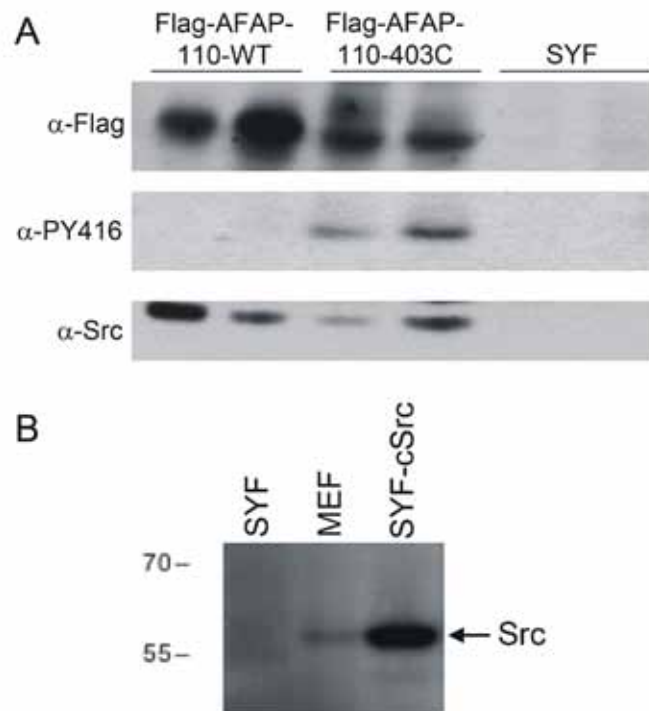AFAP-110→

C

MEF | SYF | SYF-cSrc | CaOV3

Actin→

Figure 5. Affinity precipitation of AFAP-110 with GST-PH2 and GST-PH2^{403C}

GST-affinity precipitation experiment comparing the differential ability of GST-PH2 and GST-PH2^{403C} to bind AFAP-110 in MEF, SYF, SYF-cSrc, and CaOV3 cell lysates (A). Equal quantities of GST-fusion proteins were used to affinity precipitate AFAP-110 from equal amounts of cell lysates. (A) Western blot analysis with pAb F1 indicate that GST-PH2 is more efficient in binding AFAP-110 then GST-PH2^{403C}. CaOV3 cells serve as the negative control in these experiments as this cell line has low to undetectable amount of AFAP-110. Western blots of cell lysates for AFAP-110 (B) and actin loading controls (C) are also shown.
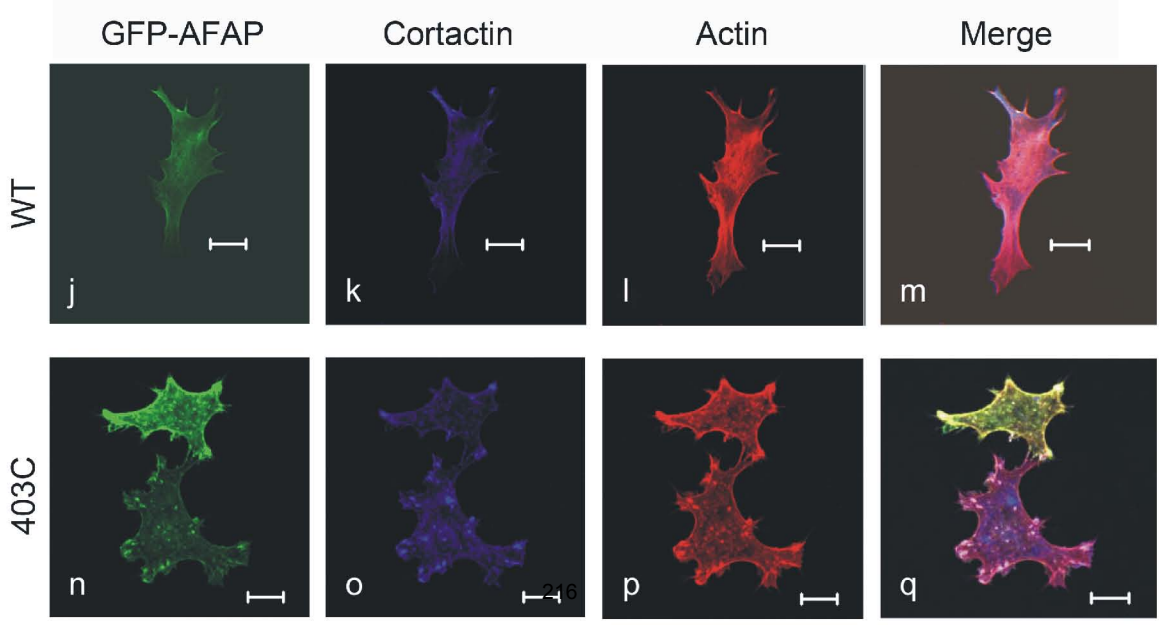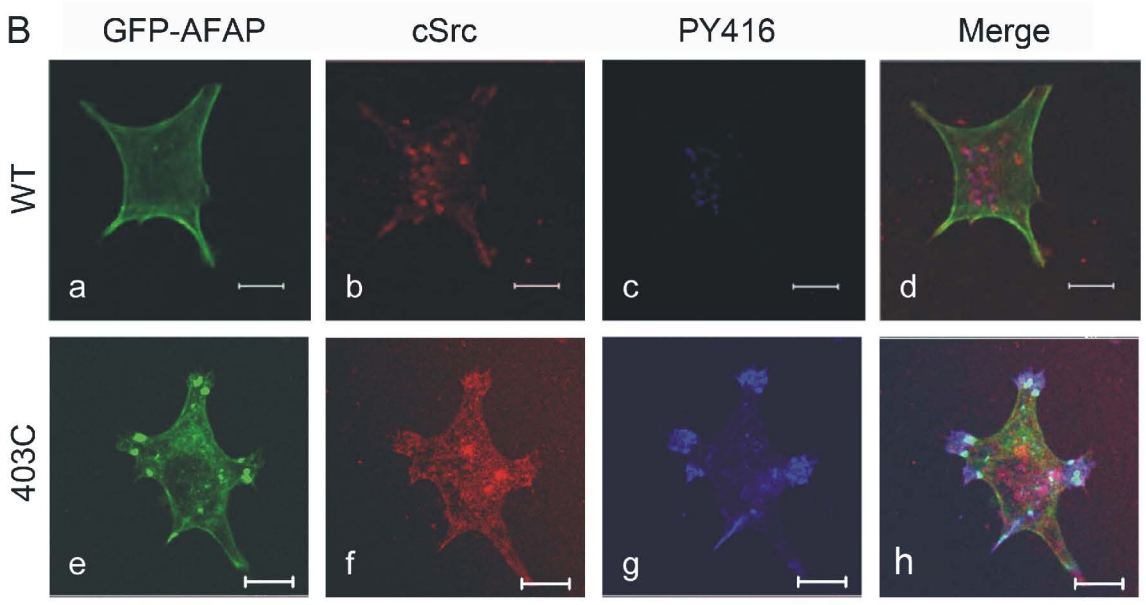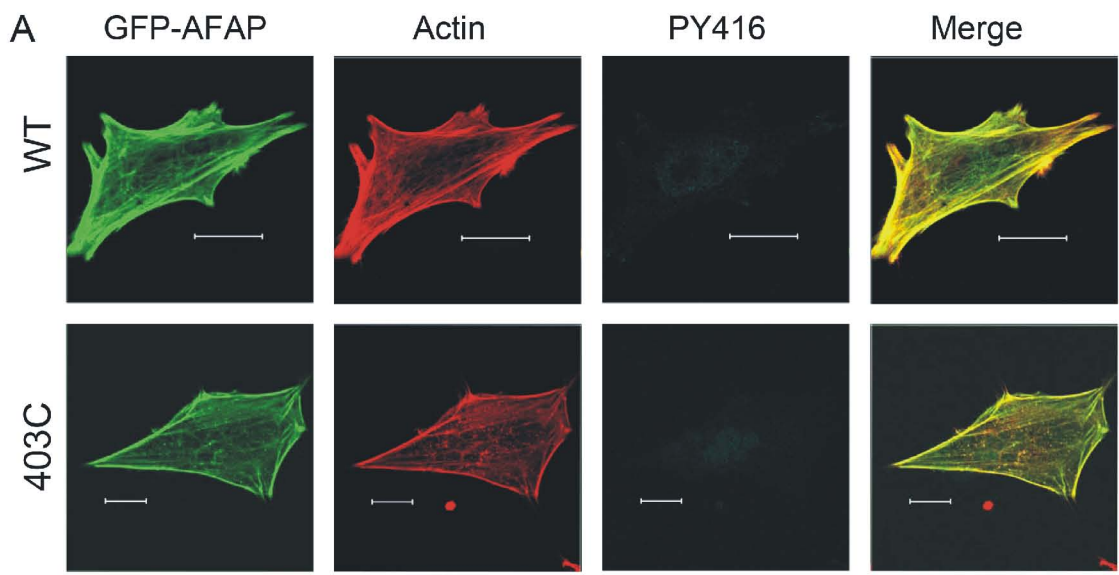
Figure 6

Figure 6. cSrc activity in SYF cells expressing cSrc and AFAP-110[403C]

(A) Flag-tagged AFAP-110 or AFAP-110[403C] was transfected into SYF-cSrc cells and expression levels detected with anti-Flag antibodies. cSrc expression levels and immunoreactivity with anti-pSrc416 antibodies were determined. (B) 50 μg of MEF, SYF or SYF-cSrc cellular lysates were resolved by 8% SDS-PAGE, transferred to PVDF and cSrc detected with anti-cSrc antibodies.

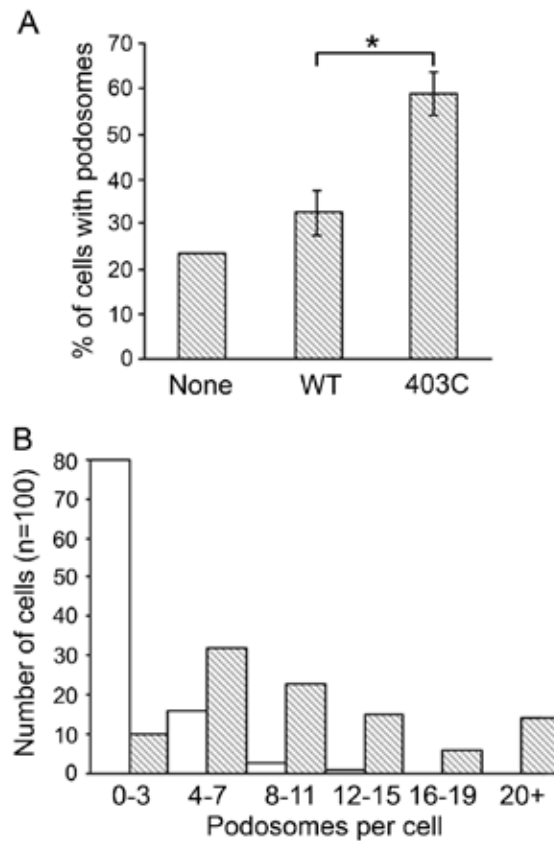**Figure 7**

A

| GFP-AFAP | Actin | PY416 | Merge |

WT

403C

B

| GFP-AFAP | cSrc | PY416 | Merge |

WT

a | b | c | d

403C

e | f | g | h

| GFP-AFAP | Cortactin | Actin | Merge |

WT

j | k | l | m

403C

n | o | p | q

226

Figure 7. cSrc activation and podosome formation in SYF-cSrc cells expressing AFAP-110^{403C}

(A) MEF cells were transfected with GFP-AFAP-110 or GFP-AFAP-110^{403C} and analyzed for activation of endogenous cSrc or changes in actin filament integrity and podosome formation. Bars = 20 μm. (B) SFY-cSrc cells similarly transiently transfected with GFP-AFAP-110 or GFP-AFAP-110^{403C} and immunolabeled with anti-Src antibody (b and f) and phospho-Src family (Y416) antibody (c and g). Unlike wild-type GFP-AFAP-110, expression of GFP-AFAP-110^{403C} resulted in the formation of punctate structures on the ventral surface of the cells enriched for GFP-AFAP-110^{403C} (e) that exhibited an increase in c-Src phosphorylation at the Y416 position (merged image, h). (C) Punctate structures resulting from the expression of GFP-AFAP-110^{403C} were also enriched for actin and cortactin (merged image, q). In contrast, cells expressing wild-type GFP-AFAP-110 maintained actin filaments and did not exhibit the formation of actin-rich podosomes (j-m). Bars: 10μm (panels a-d, e-h, n-q) and 20 μm (panels j-m).

Figure 8

Figure 8. AFAP-110[403C] directed podosome formation in SYF-cSrc cells

(A) Podosomes were counted in the transfected cells and the percentage of cells expressing podosomes was quantified. (*p = 0.015, n=500 cells). (B) The number of podosomes per cell was quantified. The podosome distribution was determineed by comparing AFAP-110 to AFAP-110[403C]. While 80% of the cells transfected with AFAP-110 wild type (empty bars) exhibit between 0-3 podosomes/cell, cells transfected with AFAP-110[403C] (hatched bars) exhibit a broad distribution with greater than 50% exhibiting more than 8 podosomes per cell.

Table W1. Summary of IHC Intensity and Pattern of cSrc and AFAP-110 Expression in 33 Ovarian Tumor Samples.

| Tumor Sample	F/D (AFAP)	S/W (AFAP)	F/D (Src)	S/W (Src)
1	0	0	D	S
2	F	S	D	W
3	F	W	D	W
4	0	0	D	S
5	F	W	D	W
6	F	W	D	S
7	F	W	D	W
8	F	S	D	S
9	D	S	D	S
10	F	S	D	W
11	D	W	D	W
12	D	S	F	S
13	F	S	D	S
14	F	S	0	0
15	F	W	D	W
16	F	S	D	W
17	D	S	F	W
18	F	W	D	S
19	0	0	F	W
20	F	W	D	W
21	D	S	D	W
22	D	S	D	W
23	F	S	D	S
24	D	S	D	W
25	D	S	F	S
26	D	S	D	W
27	F	S	D	W
28	D	S	F	W
29	D	S	D	S
30	F	S	D	W
31	F	W	D	W
32	F	S	D	W
33	D	S	D	W
Negative		3		1
Positive		30	91%	32
D/S		11		9
D/W		1		18
F/S		10		2
F/W		8		3
Diffuse		12	40%	27
Focal		18	60%	5

D, diffuse immunostaining; F, focal immunostaining; S, strong intensity; W, weak intensity.

Table W2. Presence of C1209G Nonsynonymous SNP in Tissue Samples.

Sample	CC (Wild Type)	CG SNP (One Allele)	GG SNP (Both Alleles)	Total SNP (One or Two Alleles)
Tumor (GoG)	72/91 (79.1%)	18/91 (19.8%)	1/91 (1.1%)	19/91 (20.9%)
Tumor (WVU)	24/33 (72.8%)	7/33 (21.2%)	2/33 (6.1%)	9/33 (27.3%)
Tumor (total)	97/124 (78.2%)	24/124 (19.3%)	3/124 (2.4%)	27/124 (21.7%)
Normal tissue	200/280 (71.4%)	73/280 (26.1%)	7/280 (2.5%)	80/280 (28.6%)

Exon 9 was amplified from 124 ovarian cancer tissue samples and 280 normal tissues. Sequence analysis was performed, and the presence of the wild-type sequence (C on both alleles — CC) or the SNP (G on one allele — CG) or on both alleles (GG) was calculated. In tumor-adjacent normal ovarian tissues, the nonsynonymous SNP was identified in 9 (21.2%) of 41 samples obtained from GoG. WVU indicates West Virginia University.

Figure W1
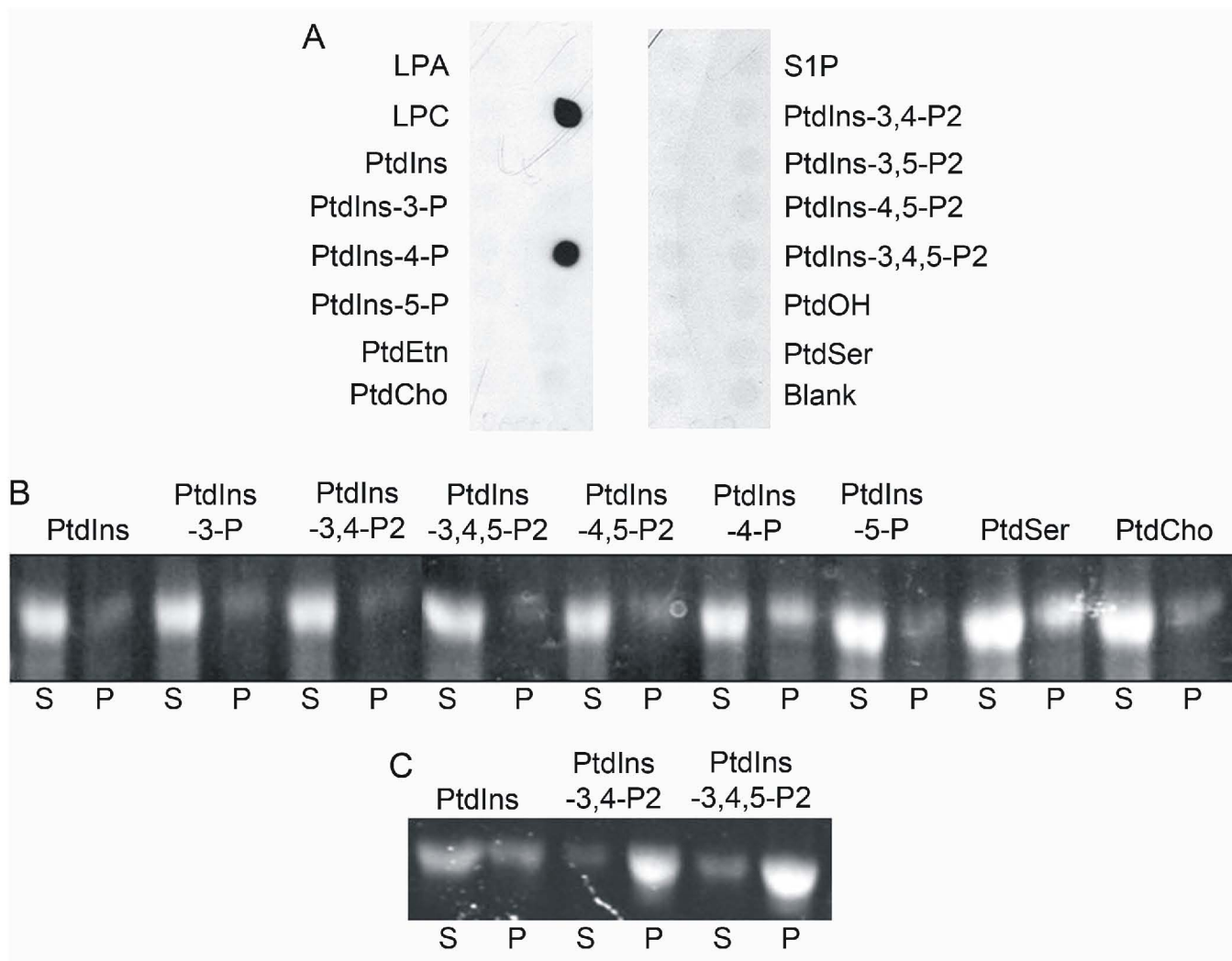
Figure W1. The PH2 domain does not bind to phospholipids

(A) Lipid dot blot analysis was used to examine the ability of GST-AFAP1-PH2 domain to bind immobilized phospholipids. GST-DAPP1-PH was used as a positive control and bound both PtdIns-3-4-P$_2$ and PtdIns-3-4-5-P$_3$, consistent with published data [36]. The GST-AFAP-110-PH2 did not bind any immobilized phosphoinositides tested. (B and C) A lipid vesicle sedimentation assay was performed with GST-PH2 (B) or the positive control GST-DAPP1-PH (C). After SDS-PAGE of the supernatants (S) and the pellets (P), the gels were stained sith SYPRO orange. Data are consistent with lipid dot-blot analysis.

Brandi Nicole Snyder

CONTACT INFORMATION

111 Queen Anne Colony

Morgantown, WV

(724)880-0275

bsnyder@hsc.wvu.edu

EDUCATION

West Virginia University School of Medicine, Morgantown, WV

PhD, Cancer Cell Biology   **2005-2011**

University of Pittsburgh, Pittsburgh, PA

Bachelor of Science, Biology *cum laude* **2001-2005**

RESEARCH EXPERIENCE

West Virginia University   2011-present

Mary Babb Randolph Cancer Center

Morgantown, WV

Post-Doctoral Fellow

Advisor: Dr. Yong Qian, PhD, Dr. Lan Guo, PhD

Studied the effect of multi-walled carbon nanotubes on the epithelial cells of the lung and how these signals affect the associated endothelial cells through the determination of cell proliferation, cell invasion, reactive oxygen species production, angiogenesis and the up- or downregulation of miRNA.

West Virginia University School of Medicine   2005-2011

Morgantown, WV

Graduate Research Assistant

Advisor: Dr. Daniel C. Flynn, PhD

Contributed 100% effort on independent research project to characterize newly identified adaptor protein AFAP1L1. Working towards identification of function and cellular binding partners of AFAP1L1 using techniques such as tissue culture, biochemical analysis, confocal microscopy and RNAi.

University of Pittsburgh Medical Center Eye Center   2004-2005

Pittsburgh, PA

Lab Assistant

Summer Undergraduate Research at Pitt (SURP) participant (2004)

Advisor: Dr. James Funderburgh, PhD

Responsible for daily lab duties such as tissue culture and buffer preparation. Assisted with the isolation of primary keratocytes from bovine corneas. Contributed to the identification of cellular markers for corneal wound healing cells.

University of Pittsburgh School of Medicine   2003-2004

Pittsburgh, PA

Lab Assistant

Advisor: Dr. Gerard Apodaca, PhD

Responsible for daily lab duties such as glassware, buffers, lab reagents and ordering.

Rangos Research Center Pediatric Neuro-Oncology   2003-2004

Pittsburgh, PA

Lab Assistant

Advisor: Dr. Ian F. Pollock, MD

Performed cell viability assays using Trypan Blue exclusion method on pediatric neuronal tumors treated with experimental chemotherapy.

PUBLICATIONS   **Snyder, BN**., Cho, YJ., Qian, Y., Coad, JE., Cunnick, JM., Flynn, DC. **AFAP1L1 is a novel adaptor protein of the AFAP family that interacts with cortactin and localized to invadosomes** European Journal of Cell Biology, Accepted, Nov. 23, 2010

Clump, DA., Yu, JJ., Cho, YJ., Gao, R., Jett, J., Zot, H., Cunnick, JM., **Snyder, BN**., Clump, AC., Shockey, M., Gannett, P., Coad, JE., Shurina, R., Figg, WD., Reed, E., Flynn, DC. **A polymorphic variant of AFAP-110 enhances cSrc activity** Transl. Oncol. 2010 Aug 1;3(4):276-85

PROFESSIONAL   <u>WVU Van Liere Convocation Research Day</u>   2010

DEVELOPMENT   West Virginia University, Morgantown, WV

Poster Presentation: Adaptor Protein AFAP1L1: A

Potential Cortactin Binding Partner

The American Society for Cell Biology 49th Annual Meeting 2009
San Diego, CA
Poster Presentation: Characterization of a Novel Adaptor Protein:
AFAP1L1

WVU Van Liere Convocation Research Day   2009
West Virginia University, Morgantown, WV
Poster Presentation: Characterization of a Novel Adaptor Protein:
AFAP1L1

The American Society for Cell Biology 48th Annual Meeting 2008
San Francisco, CA
Poster Presentation: Identification of a New Member of the AFAP
Family: AFAP1-L1

WVU Van Liere Research Day Convocation   2008
West Virginia University, Morgantown, WV
Poster Presentation: Identification of a New Member of the AFAP
Family: AFAP1-L1

WVU Van Liere Research Day Convocation   2007
West Virginia University, Morgantown, WV
Poster Presentation: A Polymorphic Variant of AFAP-110 is an
Efficient cSrc Activator

American Association for Cancer Research Annual Meeting   2007
Los Angeles, CA
Poster Presentation: A Polymorphic Variant of AFAP-110 is an
Efficient cSrc Activator

Translational Research Cancer Center Conference   2007

Cleveland, OH

Poster Presentation: A Polymorphic Variant of AFAP-110 is an efficient cSrc Activator

HONORS/

MEMBERSHIPS   West Virginia University   2010

Van Liere Research Day Convocation 3rd Place Poster Presentation

Basic Science IV Division

The American Society for Cell Biology   2008-present

Member

West Virginia University   2008

Van Liere Research Day Convocation 2nd Place Poster Presentation

Basic Science IV Division

American Association of Cancer Research   2007-2008

Member

University of Pittsburgh   2004-2005

President – Beta Beta Beta National Biological Honor Society

University of Pittsburgh   2003-2005

Member – Beta Beta Beta National Biological Honor Society

LEADERSHIP/ COMMUNITY SERVICE	American Cancer Society	2008-present
	Relay for Life of Monongalia County West Virginia	
	Planning Committee	
	Public Relations Chair	
	Advocacy Co-Chair	
	American Cancer Society	2008-present
	Relay for Life of Monongalia County West Virginia	
	Team Captain	