Title
A Sterile Alpha Motif Domain Network Involved in Kidney Development

Author
Leettola, Catherine

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Sterile Alpha Motif Domain Network Involved in Kidney Development

A dissertation submitted in partial satisfaction of the requirements for

the degree Doctor of Philosophy

in Biochemistry and Molecular Biology

by

Catherine Nicole Leettola

2015

ABSTRACT OF THE DISSERTATION

A Sterile Alpha Motif Domain Network Involved in Kidney Development

by

Catherine Nicole Leettola

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2015

Professor James U. Bowie, Chair

Cystic kidney diseases including polycystic kidney disease (PKD) and nephronophthisis (NPHP) are the most common genetic disorders leading to end-stage renal failure in humans. Animal models and human cases of PKD and NPHP have implicated the sterile alpha motif (SAM) domain containing proteins bicaudal C homolog 1 (BICC1) and ankyrin repeat and SAM-domain containing protein 6 (ANKS6) as being involved in these conditions and important for renal development. SAM domains are known protein-protein interaction domains that are capable of binding each other to form polymers and heterodimers. Using a negGFP native gel assay, we have identified the SAM domain of the previously uncharacterized protein ankyrin repeat and SAM-domain containing protein 3 (ANKS3) as a direct binding partner of the BICC1 and ANKS6 SAM domains. We found the ANKS3 SAM domain to polymerize with moderate affinity and determined the ANKS6 SAM domain can bind to a single end of this polymer. Crystal structures of the ANKS3 SAM domain polymer and the ANKS3 SAM-ANKS6 SAM

heterodimer are presented to reveal typical ML-EH SAM domain interaction interfaces with a pronounced charge complementarity. A crystal structure of the BICC1 SAM domain, which forms a helical polymer with moderate affinity, is also presented. The SAM domains of BICC1 and ANKS3 exhibit high sequence and structural homology, allowing these SAM domains to bind each other using both of their conserved ML and EH interaction surfaces. The measured binding affinities of the two possible interfaces between the BICC1 and ANKS3 SAM domains are nearly identical, suggesting these SAM domains may associate to form the first observed alternating SAM domain co-polymer. The BICC1 and ANKS6 SAM domains are also shown to bind each other as a heterodimer with strong affinity. The R823W and I817N point mutations in the SAM domain of ANKS6 are responsible for cystic kidney disease in the PKD/Mhm(cy/+) rat and a newly identified mouse model, respectively. We show that these point mutations dramatically destabilize the ANKS6 SAM domain and are responsible for a loss of interaction with the SAM domains of ANKS3 and BICC1, suggesting these lost interactions are part of a possible disease mechanism. The network of interactions between the SAM domains of BICC1, ANKS3, and ANKS6 may allow regulation of polymer formation and the creation of diverse cellular scaffolds that are suggested to be important for renal development. This work provides a structural and biochemical foundation for the continued investigation of how the interactions between the BICC1, ANKS3, and ANKS6 SAM domains mediate cellular development.

The dissertation of Catherine Nicole Leettola is approved.

Margot Quinlan

Pascal F. Egea

James U. Bowie, Committee Chair

University of California, Los Angeles

2015

Dedicated to my grandparents

George and Irene Coble

for their constant encouragement and assurance,

and to Anthony Adamucci for supporting me on this journey

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 6

Appendix I

ACKNOWLEDGEMENTS

Firstly I would like to thank my parents and brother for their love and support and unfailing belief in my abilities, even when I doubted myself. I owe a huge debt of gratitude to my longtime partner in love and life, Anthony Adamucci, who supported my decision 6 years ago to attend graduate school, followed me to Los Angeles, and has celebrated each success and supported me through each failure along the way. I would like to thank my advisor, Jim Bowie, for allowing me to work on this exciting project. I entered graduate school interested in protein-protein interactions and specifically wanted to learn crystallography to study protein structure-function relationships; thankfully this project afforded the opportunity to do just that. I would also like to thank Jim Bowie for his guidance and mentorship over the years in helping me to become a better scientist. I would like to thank the crystallography gurus Duilio Cascio and Mike Sawaya for their patience and willingness in teaching me crystallography. I owe a special thanks to Duilio for serving as a mentor and providing much needed inspiration, motivation, and career advice over the years. I would like to thank Mary Jane Knight who first exposed me to this project and graciously let me take over and continue her work when she graduated and left the Bowie lab. She continues to be an excellent SAM domain reference and a source of encouragement. Additionally I would like to thank all members of the Bowie lab for their support, suggestions, and scientific critiques. Scientific research truly is a team effort and discussions with labmates have inspired and made this research possible. Finally, I would like to thank my mentor Dr. Calvin Roff whom I worked with in San Diego before attending graduate school and who continues to provide life and career advice and stimulates my curiosity to continue in research.

Chapter 1 contains figures reprinted from other sources. Figure 1-2 is reprinted from Cell, Vol. 118, Qiao, F., et al., Derepression by Depolymerization: Structural Insights into the Regulation of Yan by Mae, Pages 163-173, Copyright (2004), with permission from Elsevier.

Figure 1-3 is reprinted from Structure, Vol.16, Harada, B., et al., Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1, Pages 380-387, Copyright (2008), with permission from Elsevier. Figure 1-4 is reprinted from TRENDS in Biochemical Sciences, Vol.31, No.7, Gundelfinger, E. D., et al., A role for zinc in postsynaptic density asSAMbly and plasticity?, Pages 366-373, Copyright (2006), with permission from Elsevier.

Chapter 2 describes the results of a negGFP native gel screen to identify novel hetero-SAM-domain interactions. I would like to thank Mary Jane Knight for cloning each of the negGFP human SAM domain fusion constructs and for allowing me to join her in the development of the negGFP native gel screen. I would also like to acknowledge James U. Bowie for guidance and conception of this screen. This work was supported by a National Institutes of Health grant R01GM093393.

Research presented in chapter 3 is a modified version of an article published in BMC Structural Biology, (2014), **14**:17, doi:10.1186/1472-6807-14-17. I would like to acknowledge the following authors who contributed to this work: Mary Jane Knight, Duilio Cascio, Sigrid Hoffman, and James U. Bowie. I would also like to thank Mike Collazo and Mike Sawaya at the UCLA-DOE X-ray Crystallization and Crystallography Core Facilities, which is supported by DOE Grant DE-FC02-02ER63421. I additionally want to thank Dan McNamara for assistance with the SEC-MALS experiments and the staff of the NECAT beamline 24-ID-C at APS which is supported by grants from the National Center for Research Resources (5P41RR015301-10) and the National Institute of General Medical Sciences (8 P41 GM103403-10) from the National Institutes of Health. Use of the APS was supported by DOE under Contract DE-AC02-06CH11357. The research presented in chapter 3 was supported by NIH Grant 5R01DK100482 and a Ruth L. Kirschstein National Research Service Award GM007185.

The research presented in chapter 4 is in preparation for publication. I would like to acknowledge the other authors, Duilio Cascio and James U. Bowie. I would like to thank Brendan Amer for assistance with the SEC-MALS experiments. I would also like to thank Mike Collazo and Mike Sawaya at the UCLA-DOE Macromolecular Crystallization Core Technology Center for their assistance with preparing crystal trays and helping to solve the crystal structures presented in this chapter. I owe additional thanks to thank Dominique Gauguier and Brigitte Lelongt for sharing information about the ANKS6 I747N mutation and to Sigrid Hoffman for putting us in contact with these researchers. This work was supported by NIH Grant 5R01DK100482, a Ruth L. Kirschstein National Research Service Award GM007185, and a UCLA Dissertation Year Fellowship.

Chapter 5 includes a preliminary analysis of the coiled coil region C-terminal to the ANKS3 SAM domain. I would like to acknowledge Mary Jane Knight and James U. Bowie for contributing to this work. This research was supported by NIH Grant 5R01DK100482 and a Ruth L. Kirschstein National Research Service Award GM007185.

The future experiments described in Chapter 6 involve several collaborations. I would like to acknowledge the laboratory of Michael Sheets at the University of Wisconsin who is currently developing cell culture assays to test BICC1 function and cellular localization. I would also like to acknowledge Sigrid Hoffman at the University of Heidelberg in Mannheim, Germany, Dominique Gauguier at the University Pierre & Marie Curie in Paris, France, and Brigitte Lelongt at Sorbonne Universités in Paris, France for their current and on-going work to develop and characterize ANKS3 mutant rats. Finally I would like to thank Tyler Luu, Christopher Koo, and Mark Arbing at the UCLA Protein Expression Technology Center for developing a purification protocol for the ankyrin repeat domains of ANKS3 and ANKS6. I also thank Soeren Lienkamp at the University of Freiburg Medical Center, Freiburg, Germany for generously providing full-length ANKS6 cDNA. This research is currently supported by NIH Grant 5R01DK100482.

The research presented in appendix I is in preparation for publication as part of a collaboration led by Dr. Guntram Borck. I would like to acknowledge the other authors N. Kakar, M. Spielmann, S. Kühl, N. Tayebi, G. Nürnberg, N. Sowada, J. Altmüller, D. Lupianez, R. Flöttmann, M. Radenz, H. van Bokhoven, C. E. Schwartz, H. Thiele, P. Nürnberg, M. Kühl, J. U. Bowie, C. Kubisch, J. Ahmad, S. Mundlos, and G. Borck. I thank Guntram Borck for providing wild-type and mutant human ZAK cDNA. I would also like to thank Brendan Amer for assistance with SEC-MALS. This work was supported by a UCLA Dissertation Year Fellowship.

VITA

2002-2006	Bachelors of Science, Biochemistry and Cell Biology University of California, San Diego
2006	Outstanding Achievement Award in Biochemistry and Cell Biology University of California, San Diego
2006-2007	Research Associate eBioscience, Inc. San Diego, CA
2007-2009	Research Associate II BioLegend, Inc. San Diego, CA
2009-2015	Graduate Student Researcher Department of Chemistry and Biochemistry University of California, Los Angeles Advisor: Professor James U. Bowie
2009-2010	UCLA Chancellor's Prize
2010-2011	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles
2010-2013	Cellular and Molecular Biology Training Grant Fellowship
2013-2014	Teaching Assistant Department of Chemistry and Biochemistry University of California, Los Angeles
2014-2015	UCLA Graduate Division Dissertation Year Fellowship

PUBLICATIONS AND PRESENTATIONS

Leettola, CN., Knight, MJ., Cascio, D., Hoffman, S., Bowie, JU. (2014) Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. *BMC Structural Biology* **14**:17.

Leettola, CN., Knight, MJ., Cascio, D., Bowie JU. The PKD-related proteins ANKS6, BICC1, and ANKS3 form a SAM domain interaction network. Poster presented at the 28th Annual Symposium of the Protein Society; 2014; San Diego, CA.

Knight, MJ., **Leettola, C.**, Gingery, M., Li, H., Bowie, JU. (2011) A human sterile alpha motif domain polymerizome*. Protein Science* **20**:1697-706.

Chapter 1

What is a Sterile Alpha Motif domain?

Introduction

The majority of eukaryotic proteins are composed of multiple domains, which are structurally and functionally distinct protein modules that when strung together provide a single protein with a diverse set of functionalities. Protein domains can be classified as being either catalytic or involved in the binding of other biological molecules including proteins, nucleic acids, lipids, and other small molecules (Pawson et al., 2002). Among these, domains responsible for protein-protein interactions are particularly important since they mediate cellular signaling, gene expression, protein and vesicular trafficking, cytoskeletal rearrangements, and control over the cell cycle.

One such protein-protein interaction domain is the sterile alpha motif (SAM) domain. SAM domains were first identified in 1995 by Christopher P. Ponting on the basis of sequence similarity among 65-70 amino acid residue segments in fourteen eukaryotic proteins (Ponting, 1995). Sterile alpha motif domains were so named due to the highly alpha-helical predicted secondary structure and the presence of this domain in four yeast proteins (Byr2, Ste11, Ste4, and Ste50) essential for sexual differentiation and that when mutated cause sterility (Ponting, 1995). The early detection of this domain by others led to a variety of initial names including HLH (helix-loop-helix), SEP (yeast sterility, Ets-related, PcG proteins), NCR (N-terminal conserved region), SPM (a subgroup of SAM domains with high homology to those of the *Drosophila* Scm and Ph proteins), and PNT (pointed, which refers to a high-homology subgroup of SAM domains found in Ets-type transcription factors) (Golub et al., 1994; Klambt, 1993; Peterson et al., 1997; Qiao and Bowie, 2005). However, SAM domain is now the universally accepted domain name.

Interaction properties of SAM domains

SAM domains are characterized as being approximately 70 amino acids in length and forming a five helix bundle. They are primarily found in eukaryotes, although some bacteria and a small handful of viruses also contain SAM domains. Within eukaryotes SAM domains are quite common. A search of the SMART protein database compiled from completely sequenced genomes reveals 6823 SAM domains among 5586 eukaryotic proteins (Letunic et al., 2015). This number is comparable to that of the well-known SH2 domain, for which SMART identifies 7407 domains in 6736 eukaryotic proteins. The number of SAM domains within the genome also generally correlates with organismal complexity (Qiao and Bowie, 2005; Vogel and Chothia, 2006).

As a protein-protein interaction domain, SAM domains commonly bind each other. In the first structure of a SAM domain, that of the EphA4 receptor solved by NMR, N- and C-terminal residues made contacts to generate dimers (Stapleton et al., 1999). In the first structure of a SAM domain solved by x-ray crystallography, that of the EphB2 receptor, N-terminal arm exchange and packing of C-terminal helices against each other formed a polymer (Thanos et al., 1999). While these types of SAM domain associations remain a possibility for these proteins, all subsequent structurally characterized SAM-SAM interactions involve two distinct surfaces which lie on opposite sides of the globular domain and are termed the mid-loop (ML) and end-helix (EH) surfaces. These interaction surfaces were first identified in the crystal structure of the TEL SAM domain and were so named because residues forming the ML surface are on loops in the central part of the domain and residues forming the EH surface localize to the N-terminal portion of the last helix (Fig. 1-1A) (Kim et al., 2001). Via the head-to-tail association of complementary ML and EH surfaces, SAM domains can self-associate to generate helical polymers (Fig. 1-1B). All observed SAM domain polymers form left-handed helices with the N- and C-termini of each individual domain splaying outward from the polymer

axis, thereby allowing polymer formation to be compatible in the context of the full-length protein. Observed polymers consist of 5-8 SAM domain subunits per helical repeat and appear flexible in terms of helical pitch (Baron et al., 2006; Harada et al., 2008; Kim, 2005; Kim et al., 2001, 2002; Nanyes et al., 2014; Stafford et al., 2011). In fact, single SAM domains have been observed to generate polymers of different dimensions (Nanyes et al., 2014; Stafford et al., 2011). SAM domains also partake in heterotypic interactions, where binding between complementary ML and EH surfaces of two different SAM domains generates heterodimers and co-polymers (Fig. 1-1B) (Kim, 2005; Lee et al., 2012; Leone et al., 2008, 2009; Qiao et al., 2004; Rajakulendran et al., 2008).

Additionally, individual SAM domain polymers are capable of associating side-by-side into large supramolecular sheets (Baron et al., 2006; Harada et al., 2008; Stafford et al., 2011). In the case of the SAM domains of Shank3 and DGKδ, the coordination of zinc ions creates interpolymer contacts that stabilize the formation of a sheet-like structure (Baron et al., 2006; Gundelfinger et al., 2006; Knight et al., 2010). Point mutations that eliminate zinc binding and preclude sheet formation impair the cellular localization and function of these SAM-domain containing proteins, demonstrating that sheet formation by SAM domain polymers is biologically significant.

However, not all SAM domains form polymers (Knight et al., 2011). Of the monomeric SAM domains, some form heterodimers with other SAM domain-containing proteins while others have been observed to bind RNA (Aviv et al., 2003), lipids (Barrera, 2003; Bhunia et al., 2009; Rufini et al., 2011), and non-SAM domain-containing proteins (Seidel and Graves, 2002; Wu et al., 2015; Yan et al., 2010).

There is also an increasing amount of evidence that some SAM domains interact with neighboring regions or domains of the proteins in which they reside. In some instances, these

interactions stabilize the folded state of the SAM domain and the region to which it is bound. For example, human STIM1 resides in the ER membrane and on the lumenal side contains 2 EF-hands immediately N-terminal to a SAM domain. When Ca^{2+} is bound to one of the EF-hands, the EF-hands and SAM domain form a single, cooperative and compact globular fold with extensive hydrophobic contacts. However, when ER-lumenal Ca^{2+} is depleted, the EF-SAM structure is destabilized, causing partial unfolding of the domain and exposure of hydrophobic residues which causes oligomerization of STIM1 and downstream activation of store-operated Ca^{2+} channels (Stathopulos et al., 2008). In another example, the neighboring coiled coil and SAM domains in the protein KSR-1 fold cooperatively into a single globular domain, termed CC-SAM. Neither the coiled coil nor SAM domain is stable independently. The coiled coil mediates binding of the protein to the plasma membrane and the SAM domain is thought to be needed for stability (Koveal et al., 2012).

In other instances, interaction of a SAM domain with a neighboring peptide segment affects protein function. For example, the SAM domain of the yeast protein Sla1p is capable of forming polymers but can also interact with a variant clathrin-box (vCB) binding motif which lies immediately C-terminal to the SAM domain. Since the ML surface of the Sla1p SAM domain binds the vCB region, polymer formation and vCB binding are mutually exclusive. It is thought that vCB binding of the SAM domain may act as a molecular switch, preventing clathrin binding in the cytosol but allowing clathrin binding at endocytic sites where Sla1p concentrations are high enough to drive SAM-domain mediated polymer formation (Di Pietro et al., 2010). In another instance, the SAM domain of the *Drosophila* polyhomeotic protein interacts with a unstructured N-terminal linker region and this reduces polymer formation (Robinson et al., 2012).

SAM domains are involved in diverse biological functions

For many proteins, the presence of a domain suggests a particular function. For example, an SH2 domain suggests the protein binds phospho-tyrosine and presence of a K-homology (KH) domain suggests nucleic acid binding (Grucza et al., 2000; Valverde et al., 2008). However, SAM domains are fairly unique in that the presence of a SAM domain does not indicate a single function but rather a variety of possible functions including regulation of gene expression, control of cellular localization, involvement in signaling cascades, and cellular scaffolding.

The human protein Transcription-Ets-Leukemia (TEL) and its *Drosophila melanogaster* homologue Yan are both members of the Ets family of transcription factors that function in transcriptional repression. Each of these proteins contains a C-terminal DNA binding domain and an N-terminal SAM domain that forms polymers, resulting in a transcriptionally silenced chromosomal region (Kim et al., 2001; Qiao et al., 2004). In the case of Yan, transcriptional repression can be regulated through the *D. melanogaster* protein Mae, which contains a monomeric SAM domain that binds the Yan SAM domain with ~1,000 fold higher affinity than Yan-SAM binds itself (Qiao et al., 2004). This heterodimeric interaction effectively depolymerizes Yan, freeing it from the DNA and allowing transcription to occur (Fig. 1-2). The *Drosophila* polyhomeotic (Ph) proteins also contain SAM domains and are developmental regulators responsible for maintaining a repressed state of homeotic genes (Kim et al., 2002). In these proteins, oligomerization of the SAM domains is required for gene silencing (Robinson et al., 2012).

SAM domains also have a demonstrated role in controlling enzyme function through control of cellular localization. The enzyme diacylglycerol kinase δ (DGKδ) catalyzes the conversion of diacylglycerol (DG) to phosphatidic acid (PA). As DG is produced at the plasma

membrane, DGKδ must localize to the membrane in order to be active (Harada et al., 2008). However, the SAM domain of DGKδ oligomerizes to form polymers and sheets of polymers that sequester the enzyme in an inactive state within the cytosol (Fig 1-3) (Harada et al., 2008; Knight et al., 2010). Since DG and PA are both second messengers, this SAM-domain mediated control of enzyme localization also regulates cellular signaling.

Other SAM domains are also involved in signal transduction and cellular signaling. The yeast proteins Byr2 and Ste11 are MEK kinases that are part of the mating pheromone response pathway and are required for yeast sexual differentiation from haploid to diploid cells (Ponting, 1995). The SAM domains of Byr2 and Ste11 bind the SAM domains of Ste4 and Ste50, respectively, and these SAM-SAM interactions are essential for proper signaling (Grimshaw et al., 2003; Ramachander, 2002). The SAM domains of the metazoan proteins CNK and HYP also bind each other and this SAM-SAM interaction is required for RAF kinase signaling in a MAPK cascade that controls cellular growth and differentiation (Rajakulendran et al., 2008).

Many SAM domain-containing proteins also contain additional protein-protein interaction motifs, allowing SAM domain interactions to organize large cellular scaffolds (Qiao and Bowie, 2005). Perhaps the best example of this is the SAM domain-containing protein SHANK3 which is found at the postsynaptic density of neuronal cells (Fig 1-4). The SAM domain of SHANK3 polymerizes and polymers assemble side-by-side into sheets that act as master scaffolds (Baron et al., 2006). The additional SH3 and ankyrin repeat domains of SHANK3 interact with the actin-based cytoskeleton and with proteins in the membrane proximal layer, which are responsible for clustering synaptic receptors and cell adhesion molecules at the cell surface (Gundelfinger et al., 2006). In this manner, the scaffold created by the SHANK3 SAM domain is responsible for the organization of proteins at the postsynaptic density.

Other SAM domain-containing proteins are also thought to function as scaffolds. These include Caskin1, which has two SAM domains in tandem that form a helical polymer. Additional functional domains of Caskin1 interact with proteins found in synaptic vesicles and at the synaptic membrane, thus creating a model where Caskin1 polymers act as a presynaptic organizational scaffolds (Stafford et al., 2011). The SAM domain of Tankyrase1 also polymerizes and through additional functional domains, Tankyrase1 interacts with a plethora of different proteins to form large lattice-like complexes. This master scaffolding activity of Tankyrase1 is important for its cellular functions in the regulation of telomere length and vesicular trafficking (De Rycker and Price, 2004).

SAM domains are involved in a number of human diseases

Besides having a diverse array of biological functions, SAM domains are also involved in a variety of human diseases. As shown in Table 1-1, chromosomal translocations of SAM domains or mutations within SAM domains are responsible for a range of conditions. Perhaps the best-characterized are chromosomal translocations which fuse the strongly polymeric TEL SAM domain to tyrosine kinases. The polymerization induced by the TEL SAM domain brings the tyrosine kinase domains within close proximity, resulting in constitutive kinase activation which leads to cell transformation and leukemia (Golub et al., 1994, 1996; Qiao and Bowie, 2005).

The SAM domain of the transcription factor P63 is also clinically significant as there are currently 24 different missense mutations within the SAM domain that are associated with Hay-Wells syndrome, also known as ankyloblepharon–ectodermal dysplasia–clefting (AEC) syndrome (Berk et al., 2009; Rinne et al., 2007, 2009). Clinical features of this syndrome include skin erosions and eyelid, nail, and tooth defects. A biochemical and structural analysis of these mutations within the P63 SAM domain suggests that they destabilize the domain

(Sathyamurthy et al., 2011). It was recently shown that the P63 SAM domain binds GM1 monogangliosides (Rufini et al., 2011). The related SAM domain of P73 has also been shown to bind lipids and it was demonstrated that lipid binding induced structural changes within the SAM domain including a measurable increase in alpha-helicity (Barrera, 2003). However, the molecular basis of how P63 SAM domain mutations cause disease still remains largely unknown.

As a final example, point mutations within the SAM domains of the proteins ANKS6 and BICC1 are associated with kidney disorders including nephronophthisis (NPHP), polycystic kidney disease (PKD), and cystic renal dysplasia (Brown et al., 2005; Hoff et al., 2013; Kraus et al., 2012). Since these disorders share similar phenotypes, it suggests that these SAM domain-containing proteins may be involved in related cellular pathways. During the course of my graduate work, we have discovered that the SAM domain of the protein ANKS3 is capable of interacting with the SAM domains of both ANKS6 and BICC1. The scaffolding complexes formed by this network of SAM domain-mediated interactions are suggested to be important for kidney development. The structural and biochemical characterization of these interactions and an understanding of how the disease-associated mutations affect their respective SAM domains is presented within the scope of this work.

In summary, SAM domains are an abundant and common motif within eukaryotic proteins. They engage in a variety of different interactions with each other to form polymers and heterodimers, and with other biological molecules. They are found throughout the cell and perform a variety of diverse functional tasks necessary for cellular maintenance and development. They are clinically significant as mutations within SAM domains and SAM-domain containing proteins result in several human disease states. Continued research into the interactions and function of each SAM domain will continue to reveal new insights into the behavior and significance of this protein domain.

Figures and Tables

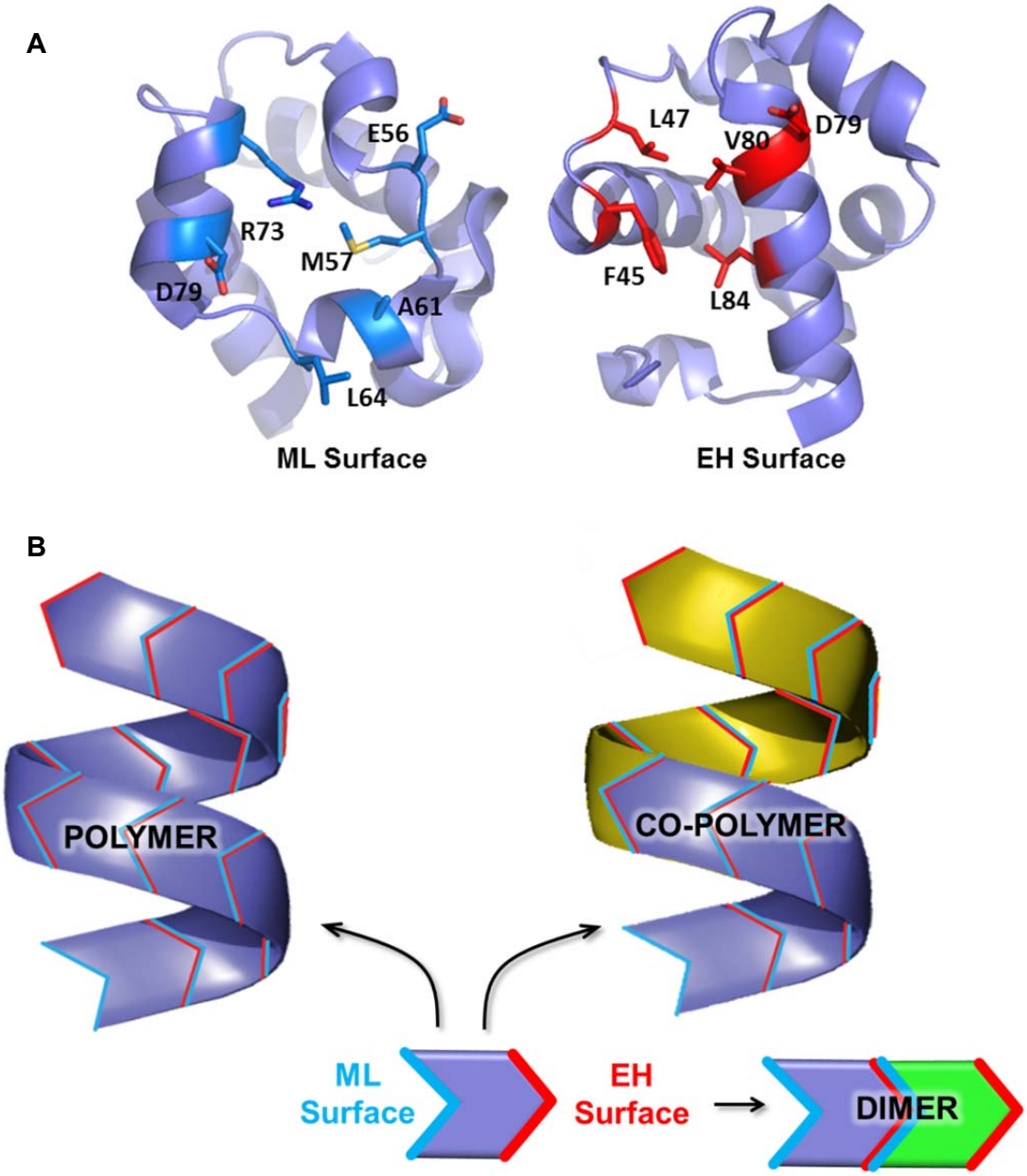


Figure 1-1. SAM domains bind each other using two distinct surfaces.

A) Residues that form the mid-loop (ML) and the end-helix (EH) surfaces are highlighted in blue and red, respectively, on the crystal structure of the TEL SAM domain (PDB:1JI7). B) Interactions between the ML and EH surfaces of SAM domains mediate homotypic interactions to produce polymers or heterotypic interactions to form co-polymers and dimers.

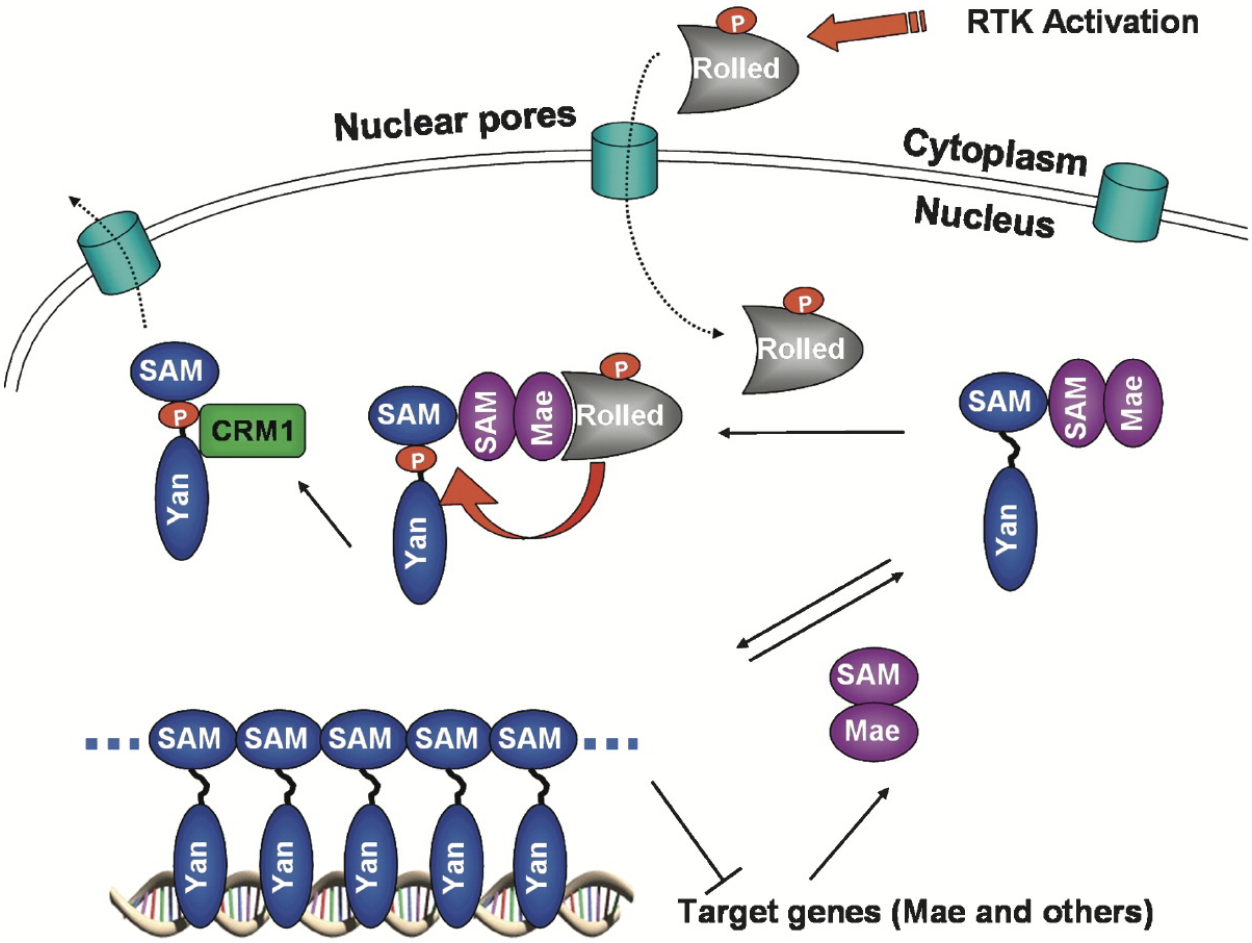


Figure 1-2. SAM domain interactions regulate gene expression.

Interactions between the SAM domains of the *Drosophila* proteins Yan and Mae regulate gene transcription. In a mechanistic model, polymerization by the SAM domain of the DNA-binding protein Yan precludes gene transcription. The SAM domain of Mae binds the SAM domain of Yan, effectively depolymerizing Yan and allowing for gene transcription to occur. This figure is reprinted from Cell, Vol. 118, Qiao, F., et al., Derepression by Depolymerization: Structural Insights into the Regulation of Yan by Mae, Pages 163-173, Copyright (2004), with permission from Elsevier.

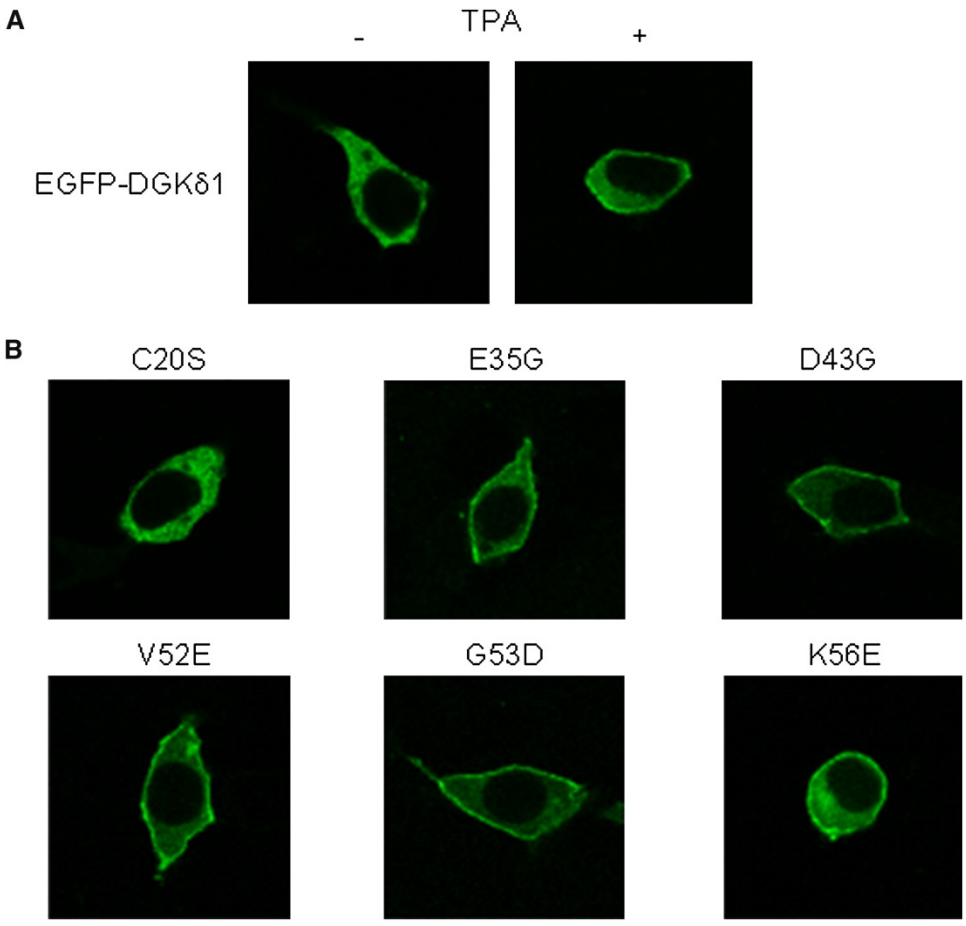


Figure 1-3. The SAM domain of DGKδ controls enzyme localization.

A) The wild-type SAM domain of DGKδ forms polymers which sequester the enzyme (tagged with GFP) within the cytosol in an inactive state. Stimulation of cells by TPA causes the protein to migrate to the plasma membrane where it is active. B) Point mutations within the DGKδ SAM domain which prevent polymerization result in localization to the plasma membrane in the absence of TPA stimulation. Note, C20S is not monomeric and has a cytoplasmic localization. This figure is reprinted from Structure, Vol.16, Harada, B., et al., Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1, Pages 380-387, Copyright (2008), with permission from Elsevier.

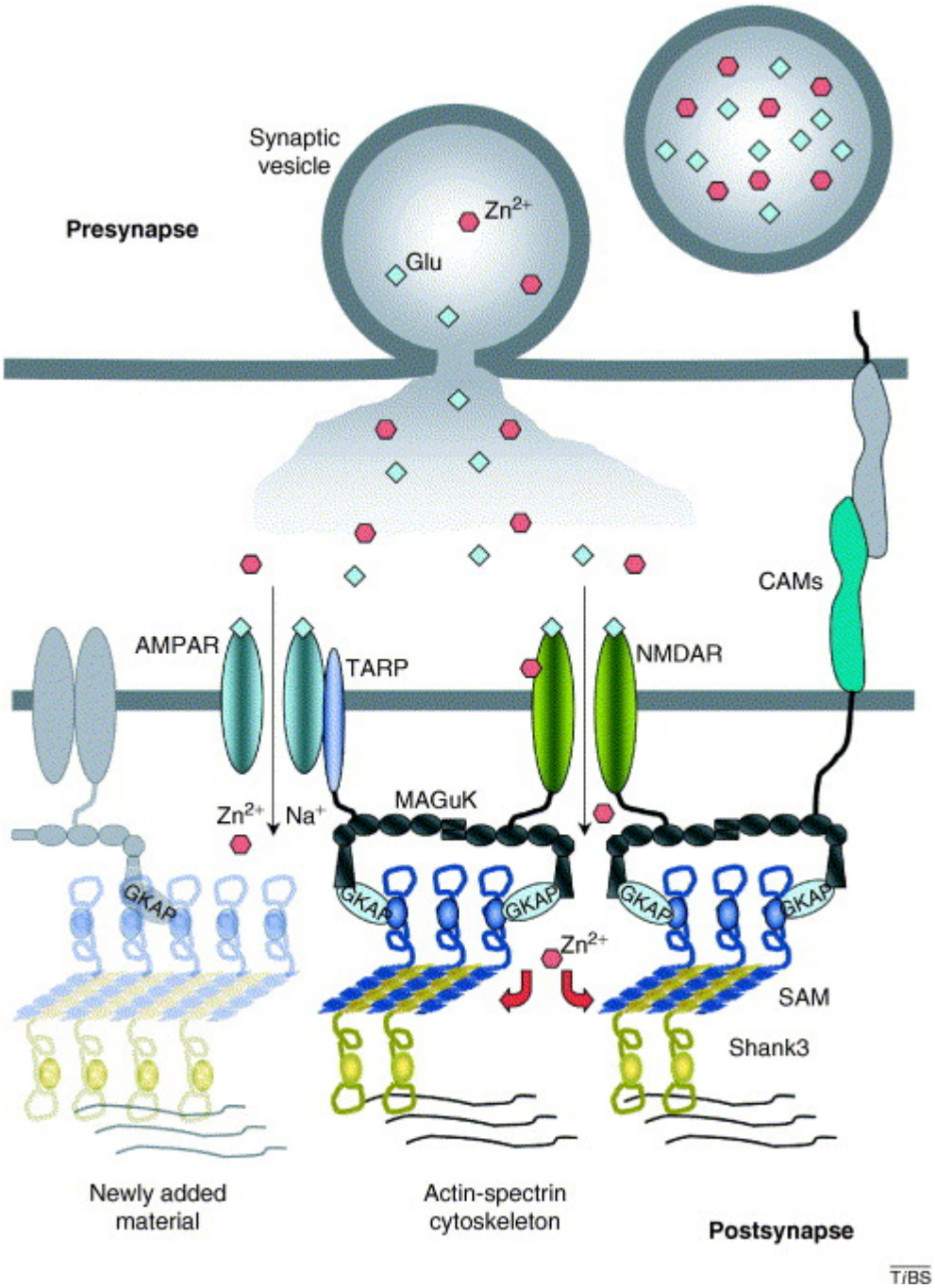


Figure 1-4. The SAM domain of SHANK3 forms a scaffold at the postsynaptic density.

A model depicting how polymers formed by the SAM domain of SHANK3 assemble side-by-side into sheets, creating a scaffold to organize proteins at the postsynaptic density. Additional functional domains of SHANK3 protrude from the SAM domain sheet and link the cytoskeleton with the membrane proximal layer. This figure is reprinted from TRENDS in Biochemical Sciences, Vol.31, No.7, Gundelfinger, E. D., et al., A role for zinc in postsynaptic density asSAMbly and plasticity?, Pages 366-373, Copyright (2006), with permission from Elsevier.

SAM domain-containing protein	Medical Relation
ANKS6	Mutations within the SAM domain are associated with nephronophthisis (NPHP) in humans and causative of polycystic kidney disease (PKD) in mice and rats (Bakey et al., 2015; Hoff et al., 2013; Neudecker et al., 2010).
Atherin	Found in atherosclerotic lesions and binds LDL. May play a role in immobilizing LDL in the arterial wall (Lees et al., 2005).
BICC1	Mutation within the SAM domain is associated with cystic renal dysplasia (Kraus et al., 2012). In-frame fusions of BICC1 with the tyrosine kinase FGFR2 are associated with cholangiocarcinomas (Wu et al., 2013).
EphA2	Mutations within the SAM domain are associated with congenital cataracts (Park et al., 2012).
L3MBTL4	Mutations within the protein and the SAM domain are found in breast cancer (Addou-Klouche et al., 2010).
P63	Mutations within the SAM domain result in Hay-Wells syndrome, also known as ankyloblepharon–ectodermal dysplasia–clefting (AEC) syndrome (Sathyamurthy et al., 2011).
Sans/Usher1G	Mutations within the SAM domain destabilize interaction with the protein harmonin and are associated with the hereditary hearing-vision loss disease, Usher syndrome 1 (Yan et al., 2010).
Shank2	Mutations throughout the protein are associated with autism spectrum disorder (Berkel et al., 2010).
Shank3	Mutations throughout the protein and within the SAM domain are associated with autism spectrum disorder (Cochoy et al., 2015; Moessner et al., 2007; Nemirovsky et al., 2015).
TEL	Chromosomal translocations fuse the TEL SAM domain to tyrosine kinases and result in leukemia (Golub et al., 1994, 1996; Qiao and Bowie, 2005).

Table 1-1. Medical significance of select human SAM domain-containing proteins.

References

Addou-Klouche, L., Adélaïde, J., Finetti, P., Cervera, N., Ferrari, A., Bekhouche, I., Sircoulomb, F., Sotiriou, C., Viens, P., Moulessehoul, S., et al. (2010). Loss, mutation and deregulation of L3MBTL4 in breast cancers. Mol. Cancer *9*, 213–213.

Aviv, T., Lin, Z., Rendl, L., M., Sicheri, F., and Smibert, C.A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. Nat. Struct. Biol. *10*, 614–621.

Bakey, Z., Bihoreau, M.-T., Piedagnel, R., Delestré, L., Arnould, C., de Villiers, A. d'Hotman, Devuyst, O., Hoffmann, S., Ronco, P., Gauguier, D., et al. (2015). The SAM domain of ANKS6 has different interacting partners and mutations can induce different cystic phenotypes. Kidney Int.

Baron, M.K., Boeckers, T.M., Vaida, B., Faham, S., Gingery, M., Sawaya, M.R., Salyer, D., Gundelfinger, E.D., and Bowie, J.U. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. Science *311*, 531–535.

Barrera, F.N. (2003). Binding of the C-terminal Sterile Motif (SAM) Domain of Human p73 to Lipid Membranes. J. Biol. Chem. *278*, 46878–46885.

Berk, D.R., Crone, K., and Bayliss, S.J. (2009). AEC syndrome caused by a novel p63 mutation and demonstrating erythroderma followed by extensive depigmentation. Pediatr. Dermatol. *26*, 617–618.

Berkel, S., Marshall, C.R., Weiss, B., Howe, J., Roeth, R., Moog, U., Endris, V., Roberts, W., Szatmari, P., Pinto, D., et al. (2010). Mutations in the SHANK2 synaptic scaffolding gene in autism spectrum disorder and mental retardation. Nat. Genet. *42*, 489–491.

Bhunia, A., Domadia, P.N., Mohanram, H., and Bhattacharjya, S. (2009). NMR structural studies of the Ste11 SAM domain in the dodecyl phosphocholine micelle. Proteins Struct. Funct. Bioinforma. *74*, 328–343.

Brown, J.H., Bihoreau, M.-T., Hoffmann, S., Kranzlin, B., Tychinskaya, I., Obermuller, N., Podlich, D., Boehn, S.N., Kaisaki, P.J., Megel, N., et al. (2005). Missense Mutation in Sterile Motif of Novel Protein SamCystin is Associated with Polycystic Kidney Disease in (cy/+) Rat. J. Am. Soc. Nephrol. *16*, 3517–3526.

Cochoy, D.M., Kolevzon, A., Kajiwara, Y., Schoen, M., Pascual-Lucas, M., Lurie, S., Buxbaum, J.D., Boeckers, T.M., and Schmeisser, M.J. (2015). Phenotypic and functional analysis of SHANK3 stop mutations identified in individuals with ASD and/or ID. Mol. Autism *6*.

Golub, T.R., Barker, G.F., Lovett, M., and Gilliland, D.G. (1994). Fusion of PDGF Receptor B to a Novel ets-like Gene, tel, in Chronic Myelomonocytic Leukemia with t(5;12) Chromosomal Translocation. Cell *77*, 307–316.

Golub, T.R., Goga, A., Barker, G.F., Afar, D.E., McLaughlin, J., Bohlander, S.K., Rowley, J.D., Witte, O.N., and Gilliland, D.G. (1996). Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. Mol. Cell. Biol. *16*, 4107–4116.

Grimshaw, S.J., Mott, H.R., Stott, K.M., Nielsen, P.R., Evetts, K.A., Hopkins, L.J., Nietlispach, D., and Owen, D. (2003). Structure of the Sterile Motif (SAM) Domain of the Saccharomyces cerevisiae Mitogen-activated Protein Kinase Pathway-modulating Protein STE50 and Analysis of Its Interaction with the STE11 SAM. J. Biol. Chem. *279*, 2192–2201.

Grucza, R.A., Bradshaw, J.M., Mitaxov, V., and Waksman, G. (2000). Role of Electrostatic Interactions in SH2 Domain Recognition: Salt-Dependence of Tyrosyl-Phosphorylated Peptide Binding to the Tandem SH2 Domain of the Syk Kinase and the Single SH2 Domain of the Src Kinase. Biochemistry *39*, 10072–10081.

Gundelfinger, E.D., Boeckers, T.M., Baron, M.K., and Bowie, J.U. (2006). A role for zinc in postsynaptic density asSAMbly and plasticity? Trends Biochem. Sci. *31*, 366–373.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Hoff, S., Halbritter, J., Epting, D., Frank, V., Nguyen, T.-M.T., van Reeuwijk, J., Boehlke, C., Schell, C., Yasunaga, T., Helmstädter, M., et al. (2013). ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3. Nat. Genet. *45*, 951–956.

Kim, C.A. (2005). Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. J. Biol. Chem. *280*, 27769–27775.

Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S., and Bowie, J.U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J. *20*, 4173–4182.

Kim, C.A., Gingery, M., Pilpa, R.M., and Bowie, J.U. (2002). The SAM domain of polyhomeotic forms a helical polymer. Nat. Struct. Biol. *9*, 453–457.

Klambt, C. (1993). The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. Development *117*, 163–176.

Knight, M.J., Joubert, M.K., Plotkowski, M.L., Kropat, J., Gingery, M., Sakane, F., Merchant, S.S., and Bowie, J.U. (2010). Zinc Binding Drives Sheet Formation by the SAM Domain of Diacylglycerol Kinase δ. Biochemistry *49*, 9667–9676.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Koveal, D., Schuh-Nuhfer, N., Ritt, D., Page, R., Morrison, D.K., and Peti, W. (2012). A CC-SAM, for Coiled Coil–Sterile α Motif, Domain Targets the Scaffold KSR-1 to Specific Sites in the Plasma Membrane. Sci. Signal. *5*, ra94.

Kraus, M.R.-C., Clauin, S., Pfister, Y., Di Maïo, M., Ulinski, T., Constam, D., Bellanné-Chantelot, C., and Grapin-Botton, A. (2012). Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. Hum. Mutat. *33*, 86–90.

Lee, H.J., Hota, P.K., Chugha, P., Guo, H., Miao, H., Zhang, L., Kim, S.-J., Stetzik, L., Wang, B.-C., and Buck, M. (2012). NMR Structure of a Heterodimeric SAM:SAM Complex:

Characterization and Manipulation of EphA2 Binding Reveal New Cellular Functions of SHIP2. Structure *20*, 41–55.

Lees, A.M., Deconinck, A.E., Campbell, B.D., and Lees, R.S. (2005). Atherin: a newly identified, lesion-specific, LDL-binding protein in human atherosclerosis. Atherosclerosis *182*, 219–230.

Leone, M., Cellitti, J., and Pellecchia, M. (2008). NMR Studies of a Heterotypic Sam−Sam Domain Association: The Interaction between the Lipid Phosphatase Ship2 and the EphA2 Receptor. Biochemistry *47*, 12721–12728.

Leone, M., Cellitti, J., and Pellecchia, M. (2009). The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. BMC Struct. Biol. *9*, 59.

Letunic, I., Doerks, T., and Bork, P. (2015). SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. *43*, D257–D260.

Moessner, R., Marshall, C.R., Sutcliffe, J.S., Skaug, J., Pinto, D., Vincent, J., Zwaigenbaum, L., Fernandez, B., Roberts, W., Szatmari, P., et al. (2007). Contribution of SHANK3 Mutations to Autism Spectrum Disorder. Am. J. Hum. Genet. *81*, 1289–1297.

Nanyes, D.R., Junco, S.E., Taylor, A.B., Robinson, A.K., Patterson, N.L., Shivarajpur, A., Halloran, J., Hale, S.M., Kaur, Y., Hart, P.J., et al. (2014). Multiple polymer architectures of human polyhomeotic homolog 3 sterile alpha motif: PHC3 SAM Can Form a Fivefold Repeat Polymer. Proteins Struct. Funct. Bioinforma. *82*, 2823–2830.

Nemirovsky, S.I., Córdoba, M., Zaiat, J.J., Completa, S.P., Vega, P.A., González-Morón, D., Medina, N.M., Fabbro, M., Romero, S., Brun, B., et al. (2015). Whole Genome Sequencing Reveals a De Novo SHANK3 Mutation in Familial Autism Spectrum Disorder. PloS One *10*.

Neudecker, S., Walz, R., Menon, K., Maier, E., Bihoreau, M.-T., Obermüller, N., Kränzlin, B., Gretz, N., and Hoffmann, S.C. (2010). Transgenic Overexpression of Anks6(p.R823W) Causes Polycystic Kidney Disease in Rats. Am. J. Pathol. *177*, 3000–3009.

Park, J.E., Son, A.I., Hua, R., Wang, L., Zhang, X., and Zhou, R. (2012). Human Cataract Mutations in EPHA2 SAM Domain Alter Receptor Stability and Function. PLoS ONE *7*, e36564.

Pawson, T., Raina, M., and Nash, P. (2002). Interaction domains: from simple binding events to complex cellular behavior. FEBS Lett. *513*, 2–10.

Peterson, A.J., Kyba, M., Bornemann, D., Morgan, K., Brock, H.W., and Simon, J. (1997). A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. Mol. Cell. Biol. *17*, 6683–6692.

Di Pietro, S.M., Cascio, D., Feliciano, D., Bowie, J.U., and Payne, G.S. (2010). Regulation of clathrin adaptor function in endocytosis: novel role for the SAM domain. EMBO J. *29*, 1033–1044.

Ponting, C.P. (1995). SAM: a novel motif in yeast sterile and Drosophila polyhomeotic proteins. Protein Sci. *4*, 1928–1930.

Qiao, F., and Bowie, J.U. (2005). The many faces of SAM. Sci. Signal. *2005*, re7.

Qiao, F., Song, H., Kim, C.A., Sawaya, M.R., Hunter, J.B., Gingery, M., Rebay, I., Courey, A.J., and Bowie, J.U. (2004). Derepression by depolymerization: structural insights into the regulation of Yan by Mae. Cell *118*, 163–173.

Rajakulendran, T., Sahmi, M., Kurinov, I., Tyers, M., Therrien, M., and Sicheri, F. (2008). CNK and HYP form a discrete dimer by their SAM domains to mediate RAF kinase signaling. Proc. Natl. Acad. Sci. *105*, 2836–2841.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Rinne, T., Brunner, H.G., and van Bokhoven, H. (2007). p63-Associated Disorders. Cell Cycle *6*, 262–268.

Rinne, T., Bolat, E., Meijer, R., Scheffer, H., and van Bokhoven, H. (2009). Spectrum of *p63* mutations in a selected patient cohort affected with ankyloblepharon-ectodermal defects-cleft lip/palate syndrome (AEC). Am. J. Med. Genet. A. *149A*, 1948–1951.

Robinson, A.K., Leal, B.Z., Chadwell, L.V., Wang, R., Ilangovan, U., Kaur, Y., Junco, S.E., Schirf, V., Osmulski, P.A., Gaczynska, M., et al. (2012). The Growth-Suppressive Function of the Polycomb Group Protein Polyhomeotic Is Mediated by Polymerization of Its Sterile Alpha Motif (SAM) Domain. J. Biol. Chem. *287*, 8702–8713.

Rufini, S., Lena, A.M., Cadot, B., Mele, S., Amelio, I., Terrinoni, A., Desideri, A., Melino, G., and Candi, E. (2011). The sterile alpha-motif (SAM) domain of p63 binds in vitro monoasialoganglioside (GM1) micelles. Biochem. Pharmacol. *82*, 1262–1268.

De Rycker, M., and Price, C.M. (2004). Tankyrase Polymerization Is Controlled by Its Sterile Alpha Motif and Poly(ADP-Ribose) Polymerase Domains. Mol. Cell. Biol. *24*, 9802–9812.

Sathyamurthy, A., Freund, S.M.V., Johnson, C.M., Allen, M.D., and Bycroft, M. (2011). Structural basis of p63α SAM domain mutants involved in AEC syndrome: Mutants involved in AEC syndrome. FEBS J. *278*, 2680–2688.

Seidel, J.J., and Graves, B.J. (2002). An ERK2 docking site in the Pointed domain distinguishes a subset of ETS transcription factors. Genes Dev. *16*, 127–137.

Stafford, R.L., Hinde, E., Knight, M.J., Pennella, M.A., Ear, J., Digman, M.A., Gratton, E., and Bowie, J.U. (2011). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK. Structure *19*, 1826–1836.

Stapleton, D., Balan, I., Pawson, T., and Sicheri, F. (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. Nat. Struct. Biol. *6*, 44–49.

Stathopulos, P.B., Zheng, L., Li, G.-Y., Plevin, M.J., and Ikura, M. (2008). Structural and Mechanistic Insights into STIM1-Mediated Initiation of Store-Operated Calcium Entry. Cell *135*, 110–122.

Thanos, C.D., Goodwill, K.E., and Bowie, J.U. (1999). Oligomeric Structure of the Human EphB2 Receptor SAM Domain. Science *283*, 833–836.

Valverde, R., Edwards, L., and Regan, L. (2008). Structure and function of KH domains: Structure and function of KH domains. FEBS J. *275*, 2712–2726.

Vogel, C., and Chothia, C. (2006). Protein Family Expansions and Biological Complexity. PLoS Comput. Biol. *2*, e48.

Wu, Y., Koharudin, L.M.I., Mehrens, J., DeLucia, M., Byeon, C.-H., Byeon, I.-J.L., Calero, G., Ahn, J., and Gronenborn, A.M. (2015). Structural basis of clade-specific engagement of SAMHD1 restriction factors by lentiviral Vpx virulence factors. J. Biol. Chem. doi: 10.1074/jbc.M115.665513.

Wu, Y.-M., Su, F., Kalyana-Sundaram, S., Khazanov, N., Ateeq, B., Cao, X., Lonigro, R.J., Vats, P., Wang, R., Lin, S.-F., et al. (2013). Identification of Targetable FGFR Gene Fusions in Diverse Cancers. Cancer Discov. *3*, 636–647.

Yan, J., Pan, L., Chen, X., Wu, L., and Zhang, M. (2010). The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc. Natl. Acad. Sci. *107*, 4040–4045.

Chapter 2

A negGFP screen to identify novel hetero-SAM domain interactions

Abstract

Sterile alpha motif (SAM) domains are a common protein-protein interaction domain. Many SAM domains are known to self-associate into helical polymers, however a large portion of SAM domains are classified as monomeric. Some monomeric SAM domains have been shown to associate through heterotypic interactions with the SAM domains of other proteins. Thus, the population of monomeric SAM domains may engage in heterotypic interactions which have not yet been identified. Here we developed a negGFP native gel screen which was used to detect hetero-interactions of 41 different monomeric human SAM domains with each other or with 23 different polymeric human SAM domains. From this screen we identified three novel human SAM domain hetero-interactions: ARAP2/CentaurinD1 and INPPL1, ANKS3 and ANKS6, Caskin1 and EPHA6. By identifying new SAM domain mediated interactions, we further define the complexes these proteins are part of and gain additional function insight into the roles of the SAM-domain containing proteins.

Introduction

Eukaryotic proteins are commonly composed of several protein domains that when linked together in a modular fashion imbue a single protein with multiple functionalities (Pawson et al., 2002). Many protein domains are involved in generating protein-protein interactions and in general, the number of protein-protein interaction domains in a proteome correlates with organismal complexity (Xia et al., 2008). Therefore, an understanding of cellular processes at the molecular level undoubtedly involves a detailed understanding of the various associations mediated by these protein-protein interaction domains.

The sterile alpha motif (SAM) domain is a well-studied protein-protein interaction domain. Many SAM domains are known to oligomerize into helical polymers by forming successive associations between the mid-loop (ML) and end-helix (EH) surfaces of the SAM domain (Kim et al., 2001; Qiao et al., 2004; Baron et al., 2006; Harada et al., 2008; Stafford et al., 2011). This polymeric nature of some SAM domains generates a variety of functions, including the creation of cellular scaffolds (Baron et al., 2006; Stafford et al., 2011), regulation of enzyme activity by controlling cellular localization (Harada et al., 2008), and control of gene expression (Kim et al., 2001; Qiao et al., 2004). However, there are a large number of SAM domains which do not form polymers (Meruelo and Bowie, 2009; Knight et al., 2011). These monomeric SAM domains have been shown to bind RNA (Aviv et al., 2003), lipids (Barrera, 2003; Bhunia et al., 2009; Rufini et al., 2011), engage in heterotypic interactions with the SAM domains of other proteins (Kwan et al., 2004; Qiao et al., 2004; Ramachander, 2002), or bind other proteins altogether (Seidel and Graves, 2002; Wu et al., 2015; Yan et al., 2010). Given the propensity of SAM domains to bind each other, we reasoned that the large number of monomeric SAM domains may actually engage in heterotypic SAM domain interactions which have not yet been identified.

We had previously attempted to identify novel hetero-SAM domain interactions using a split-DHFR assay, described in the dissertation of Mary Jane Knight (Knight, 2010). In this approach, human SAM domains were fused to either the N- or C-terminal fragment of dihydrofolate reductase (DHFR) such that a SAM-SAM interaction would reconstitute an enzymatically active DHFR and allow transformed *E. coli* to survive in the presence of trimethoprim. However, when libraries containing SAM domains fused to N- and C-terminal DHFR fragments were mixed, we observed that many human SAM domains appeared promiscuous, meaning that they engaged in hetero-interactions with many other human SAM domains. Such a wide range of interactions suggested these interactions were not highly specific or of high affinity and therefore likely to be artifacts of the selection method.

To address this problem we developed a negGFP native gel screen to detect SAM domain hetero-interactions with higher stringency. We applied this method to screen 64 different human SAM domains and detected three novel hetero-SAM interactions between ARAP2/CentaurinD1 and INPPL1, ANKS3 and ANKS6, Caskin1 and EPHA6. The identification of heterotypic SAM domain interactions and the resulting protein complexes allows a further understanding of the functional roles of SAM-domain containing proteins.

Results

Development of a negGFP native gel screen to identify hetero-interactions

To develop a rapid and high-throughput method to search for hetero-SAM domain interactions with binding affinities strong enough to be biologically relevant we developed a negGFP native gel screen using the negGFP fusion proteins from our previous analysis of the human SAM domain polymerizome (Knight et al., 2011). In this work, human SAM domains were fused to a green fluorescent protein modified to have a net charge of -30 (negGFP) (Lawrence et al., 2007). This negGFP fusion provided several advantages: 1) the charge repulsion from neighboring negGFPs reduced the binding affinity and polymeric character of strongly polymeric SAM domains, allowing them to migrate into the gel; 2) the large negative charge of the GFP caused the fusion proteins to migrate towards the cathode of a native gel, with migration distance inversely proportional to the extent of polymerization; 3) because negGFP can be imaged using a fluorescence imager, this allowed crude cellular lysate to be loaded on native gels as opposed to purified protein.

In the set-up for this experiment (Fig. 2-1) negGFP fusions of human SAM domains were run individually and as mixes on a native gel. From our experience with the human SAM domain polymerizome project (Knight et al., 2011), negGFP fusions of monomeric SAM domains run as discrete bands while negGFP fusions of polymeric SAM domains run as smeared bands with reduced migration. Hetero-interactions are identified as a band shift when the negGFP fusions of the interacting SAM domains are mixed together.

Prior research using monomeric point mutants of the well characterized SAM domains from the human proteins transcription factor ETV6 (TEL) and diacylglycerol kinase delta (DGKδ) showed in a proof of concept that mixing constructs capable of forming dimers resulted in gel shifts (Knight et al., 2011). We repeated this experiment and verified that mixing monomeric

point mutants of TEL and DGKδ indeed created heterodimers which ran with reduced migration (Fig. 2-2A,B). The native TEL heterodimer has a strong measured binding affinity (K_d = 1.7 ± 0.5 nM) (Kim et al., 2001) and the native DGKδ heterodimer has a weaker measured binding affinity (K_d = 6.3 ± 1.2 μM) (Harada et al., 2008). Thus, this negGFP native gel screen allows the detection of both relatively weak and strong interactions which are expected to biologically relevant.

Using the negGFP native gel screen in a high-throughput manner required multiple hetero-SAM interactions to be tested in a single gel lane. To show that this was feasible, we mixed monomeric point mutants of TEL and DGKδ in the presence of additional negGFP SAM domain fusions and assessed that the heterodimeric interactions were still observable (Figs. 2-2A,B). Indeed, TEL and DGKδ heterodimers produced the same gel shift in the presence of three additional negGFP human SAM fusions. Based on this, we set up the screen to test four different negGFP SAM constructs per gel lane.

With this method we also tested the validity of some of the hetero-SAM interactions that were detected in the split-DHFR assay. negGFP fusions of the SAM domains of TP63 and SAMD13 were mixed in 1:1 ratios with negGFP fusions of human SAM domains that were identified as interactants in the split-DHFR screen (Fig. 2-2C). However, in all cases, these 1:1 mixes did not result in any gel shift, indicating no interaction. This reinforced that the interactions detected in the split-DHFR assay were either artifacts or of sufficiently low binding affinity and therefore biologically irrelevant.

A single hetero-interaction between monomeric SAM domains identified by negGFP native gel screen

Given that SAM domains are protein-protein interaction domains and typically associate with each other, we reasoned that monomeric SAM domains may form heterotypic interactions with other monomeric SAM domains. To test this, we selected negGFP SAM domain fusion constructs from 41 different human proteins for which the SAM domain was identified as being monomeric (Knight et al., 2011). Cellular lysate from these constructs was organized into "pots" (Fig 2-3A), such that multiple SAM domains could be screened per lane. To normalize the amount of negGFP fusion protein of each constituent in each lane, equal amounts of protein based on fluorescence were loaded. A single negGFP fusion of a monomeric SAM domain was then added to each "pot" in a 1:1 ratio with all "pot" constituents, such that each lane contained equal amounts of all proteins, based on fluorescence. After allowing the mixes to equilibrate, the samples were run on native gels (Fig. 2-4) and the banding pattern of each gel was compared to that of the "pots" alone (Fig 2-3B) to identify gel shifts indicative of an interaction.

From this screen, we observed gel shifts for the proteins ARAP2/CentaurinD1 and INPPL1. The gel shift for ARAP2/CentaurinD1 occurred when this protein was mixed with "pot" 6 and the gel shift for INPPL1 occurred when this protein was mixed with "pot" 5. Since "pot" 6 contained INPPL1 and "pot" 5 contained ARAP2/CentaurinD1 this suggested a hetero-interaction between these SAM domains. However, although our discovery of this specific hetero-interaction was novel, a highly similar hetero-interaction between the SAM domains of INPPL1 and ARAP3 had been previously identified, biochemically characterized, and structurally mapped using NMR (Raaijmakers et al., 2007; Leone et al., 2009). Many of the residues of ARAP3 that were identified by chemical shift perturbations as being involved in binding INPPL1 are conserved in the ARAP2/CentaurinD1 SAM domain sequence, which is overall 40% identical and 60% similar to the SAM domain of ARAP3 (Fig. 2-5A). Additionally,

the structures of the SAM domains of ARAP3 and ARAP2 solved by NMR align very well (RMSD across all atoms = 1.085Å, RMSD across backbone atoms = 0.776Å) (Fig. 2-5B). Moreover, the related SAM domain of ARAP1/CentaurinD2 has been proposed to interact with INPPL1 as well (Cuthbert et al., 2008). Due to the considerable amount of existing research on the ARAP3-INPPL1 SAM domain hetero-interaction and that the sequence and structural homology between ARAP3 and ARAP2/CentaurinD1 provides a simple explanation for our observed hetero-interaction, we chose not to pursue this further.

Two novel hetero-interactions between polymeric and monomeric SAM domains identified by negGFP native gel screen

We next focused our attention on identifying polymer-capping SAM domains. In this scenario, a monomeric SAM domain binds to a single end of a SAM domain polymer, effectively capping the extent of polymerization and pending the correct stoichiometry and binding affinity, de-polymerizing the SAM domain polymer to form heterodimers. This type of heterotypic SAM domain interaction has been observed previously with the SAM domains of the *Drosophila* proteins Yan and Mae (Qiao et al., 2004). To search for polymer-capping hetero-interactions we mixed negGFP fusions of 23 different polymeric human SAM domains (Fig 2-6) with the "pots" of monomeric SAM domains described above. To shift the equilibrium of binding towards the heterotypic interaction we mixed the polymeric SAM constructs in a 1:2 ratio based on fluorescence with the monomeric SAM domains. This was to influence complete interaction of the polymeric SAM domain to create a complete and obvious gel shift. Gels were then compared against the monomeric SAM domain "pots" alone (Fig. 2-3B) to identify gel shifts.

Compared to the monomeric SAM domain native gels above, gel shifts of the polymeric SAM domains were much more difficult to distinguish (Fig. 2-7). In some cases, addition of the

monomeric SAM domains caused the protein to uniformly migrate faster (SCMH1, PH2). In other instances, some of the negGFP polymeric SAM domain fusions ran unevenly across all lanes (Atherin/SAMD1, PH3, PH-like1/PHC1). For strongly polymeric SAM domains that do not enter the gel otherwise (TEL/ETV6, SCML1, LBP1, Kazrin-3 SAMs) we looked for downward gel shifts with the appearance of a new band within the gel accompanied by disappearance of the band at the top of the gel. However, for these strongly polymeric constructs we detected no hetero-interactions. All gels were closely examined, searching for small gel shifts and/or changes in the intensity and pattern of the monomeric SAM constructs in each lane. From this we detected possible hetero-interactions involving the SAM domains of ANKS3, SAMD9, Caskin1, and PH-like1/PHC1 which were followed up on by mixing each SAM domain with its potential monomeric SAM domain interacts and looking for band shifts when these were run on native gels (Fig. 2-8).

An interaction between the SAM domains of ANKS3 and ANKS6 was confirmed with the presence of a strong gel shift when the negGFP fusions of these SAM domains were mixed in a 1:2 ratio, respectively (Fig. 2-8A). This novel hetero-SAM interaction formed the springboard for the basis of the work in this dissertation and is discussed in detail in Chapter 3.

The potential hetero-interactions involving the SAM domains of SAMD9 and PH-like1/PHC1 could not be reproduced (Fig 2-8B). When these constructs were mixed in 1:1 ratios with potential hetero-interactants we observed no gel shifts aside from a faster apparent migration of SAMD9 mixed with SAMD12. However, since there were no gel upshifts as had previously been observed and because there were no decreases in the fluorescence intensity of the monomeric bands in the 1:1 mixes, we were unable to identify any interactions.

An interaction between the tandem SAM domains of Caskin1 and the SAM domain of EPHA6 was confirmed with the presence of a slight gel upshift accompanied by a decrease in

the intensity of the monomeric EPHA6 band (Fig. 2-8B). Caskin1 contains two SAM domains in tandem and a crystal structure (PDB:3SEI) shows that the ML surface of the first SAM domain binds the EH surface of the second SAM domain (Stafford et al., 2011). This leaves the EH surface of the first SAM domain and the ML surface of second SAM domain free, allowing Caskin1 to polymerize through the sequential association of these surfaces. To further understand the interaction between the SAM domains of Caskin1 and EPHA6 we tested whether the SAM domain of EPHA6 bound the free ML or free EH surface of Caskin1. To probe this interaction, we mixed the SAM domain of EPHA6 with the Caskin1 point mutants G520E and G566K, which inhibit binding to either the free EH or free ML surface of the Caskin1 tandem SAM domains, respectively (Stafford et al., 2011) (Fig. 2-8C). The SAM domain of EPHA6 is able to bind both the wild-type Caskin1 and the G520E point mutant, as indicated by the slight gel upshift and disappearance of the band corresponding to EPHA6. However, the point mutation G566K prevents this interaction, meaning that the SAM domain of EPHA6 binds the free ML surface of the second SAM domain in Caskin1.

Discussion

Of the 79 different human SAM-domain containing proteins for which the SAM domains were successfully cloned (Knight et al., 2011), we screened SAM domains from 64 different proteins for hetero-interactions. From this, novel hetero-interactions were detected between the SAM domains of ARAP2/CentaurinD1 and INPPL1, ANKS3 and ANKS6, and Caskin1 and EPHA6. Although the interaction of the SAM domains of ARAP2/CentuarinD1 and INPPL1 had not previously been identified, the SAM domain of ARAP3 which shares close sequence and structural homology with the SAM domain of ARAP2/CentaurinD1 did have a documented and characterized interaction with INPPL1. The interaction between the SAM domains of ANKS3

and ANKS6 is discussed in detail in Chapter 3. Regarding the interaction of the SAM domains of Caskin1 and EPHA6, we identified that the SAM domain of EPHA6 binds the free ML surface of the second tandem SAM domain in Caskin1. Since SAM domains typically associate via ML-EH interfaces, this suggests that the EH surface of EPHA6 is responsible for binding Caskin1. A sequence alignment of the SAM domains of the human Eph receptors shows that all have highly similar sequences in the region consistent with the EH surface (Fig. 2-9). Thus, the SAM domains of other Eph receptors may also interact with the SAM domains of Caskin1. Further research into these possible interactions and their biological significance will need to be undertaken.

Given the large number of SAM domains screened, we expected to identify significantly more interactions. There are however several possible explanations for why more interactions were not detected. First, it is possible that some heterotypic SAM domain interactions require the presence of nearby functional domains within their respective proteins. For example, a hetero-interaction between the SAM domains of the yeast proteins Byr2 and Ste4 has weak affinity (K_d = 56 μM) but when the Ste4 construct contains an additional leucine-zipper domain which is found in the wild-type protein immediately C-terminal to the SAM domain, this SAM domain hetero-interaction becomes significantly stronger (K_d = 19 nM).

Secondly, our lack of detection of many hetero-interactions could indicate that a number of heterotypic SAM domain interactions are of low affinity and therefore undetectable in our screen. In support of this, the SAM domain of EPHA4 is known to self-associate very weakly (K_d = 500 μM – 5 mM) (Stapleton et al., 1999) and the SAM domain of EPHB2 also binds itself weakly (K_d > 100 μM) (Smalla et al., 1999; Thanos et al., 1999). It is thought that the membrane localization of the EPHA4 and EPHB2 proteins increases their local concentration to compensate for the effects of otherwise low binding affinities (Stapleton et al., 1999). This could be a trend among some of the SAM-domain containing proteins.

It is also possible that the specific peculiarities of this screen caused interactions to be missed. In support of this, there were a number of hetero-interactions which should have been detected. For example, the SAM domain of ARAP3, which shares sequence and structural homology with ARAP2, binds the SAM domain of INPPL1 with a strong measured binding affinity of approximately 100 nM (Raaijmakers et al., 2007). However, the SAM domain of INPPL1 also binds the SAM domain of EPHA2 with a measured binding affinity of 0.75 ± 0.12 μM (Leone et al., 2009). However, no band shifts were detected in the lanes where EPHA2 and INPPL1 were mixed. Additionally, this negGFP native gel screen failed to detect an interaction between the SAM domains of BICC1 and ANSK6 which are described in Chapter 4 and have a strong measured binding affinity (K_d = 460.5 ± 10 nM). Given that the relatively weak dimerization of monomeric point mutants of DGKδ (K_d = 6.3 ± 1.2 μM) could be detected, the above-mentioned interactions should have been observed. Many SAM domain binding interfaces are composed of a hydrophobic region surrounded by charged residues which form salt bridges and hydrogen bonds across the interface. In particular, the EH surface of the SAM domain of BICC1, which binds the SAM domain of ANKS6, is positively charged. It may be possible that the negGFP interacted with this positively charged surface, thereby shielding it from interaction with the SAM domain of ANKS6. However, if this were the case, this same shielding would have been expected to interfere with BICC1 SAM polymerization. This same potential charge shielding would have also been expected to interfere with the interaction between the SAM domains of ANKS6 and ANKS3, which has an EH surface nearly identical to that of BICC1 and binds ANKS6 in the same manner.

Nevertheless, the negGFP fusion appears to affect SAM domain interactions unusually. For instance, the SAM domain of BICC1 is able to bind the SAM domain of ANKS3 using both its ML and EH surfaces (Chapter 4). This creates a BICC1-EH/ANKS3-ML interface and a BICC1-ML/ANKS3-EH interface which have roughly the same measured binding affinities.

However, the gel shift observed by formation of the BICC1-ML/ANKS3-EH interface is of greater magnitude than the gel shift observed by formation of the BICC1-EH/ANKS3-ML interface. Other migration idiosyncrasies include the dimeric TEL SAM domain, which migrates slower than the dimeric DGKδ SAM domain and slower also than the wild-type DGKδ SAM domain that exists as a short polymer. Hence, although extent of migration of the negGFP fusion proteins tends to correlate with molecular weight, this is not always the case.

This screen was also not well suited to detect hetero-interactions involving polymeric SAM domains. If a monomeric SAM domain acted as a polymer capper, the small change in size due to the addition of a single SAM domain to an otherwise long polymer is not sufficient to cause an obvious gel shift. Polymeric and monomeric SAM domain constructs were mixed in a 1:2 ratio to shift the equilibrium toward depolymerization of the polymeric SAM domain and formation of heterodimers. However, complete heterodimerization would require that the binding affinity between the monomeric and polymeric SAM domains is stronger than the affinity of the polymeric SAM domain for itself, which may not hold true. An alternate way to detect these interactions was to look for the disappearance or weakened intensity of a monomeric SAM domain band. With the screen set-up used here this was not always feasible since the monomeric SAM domain “pots” contained multiple SAM domains which ran with similar migration and the sensitive detection of negGFP fluorescence made distinguishing altered band intensity difficult. If this screen were to be repeated, fewer SAM domains should be mixed in each lane such that the SAM domains in each lane run as distinct bands where the fluorescence intensity can be more easily monitored.

Overall, this negGFP native gel screen has proved useful for identifying several novel SAM domain hetero-interactions. However the lack of a detectable interaction does not mean the interaction does not exist. Rather, this method finds more utility in quickly identifying and interpreting how mutations or altered constructs affect a known and detectable interaction.

Methods

Cloning

All negGFP human SAM domain fusion constructs were as described previously (Knight et al., 2011). negGFP fusion constructs of the Caskin1 tandem SAM domain point mutants were taken from prior work (Stafford et al., 2011).

Preparation of negGFP SAM domain fusion proteins

Lysate of polymeric negGFP SAM domain fusions is the same as was used previously (Knight et al., 2011). Lysate of monomeric negGFP SAM domain fusions was prepared again following the method described in Leettola, et al., 2014. Briefly, negGFP human SAM fusion constructs transformed into ARI814 cells were induced with 0.2% arabinose and expressed at 16^{o}C for 16 hours. Harvested cells were resuspended with 0.5mL of 20mM Tris pH 7.5, 0.3M NaCl, 2mM TCEP supplemented with 40μg/mL DNase1, 10mg/mL lysozyme, 1mM PMSF, 10mM $MgCl_2$ and ½ of a cOmplete mini, EDTA-free protease inhibitor tablet (Roche) per sample. Cells were lysed by three freeze-thaw cycles and pulsed by sonication for 10 seconds. Samples were centrifuged at 16,060*g* for 20 minutes and supernatant was transferred to a clean tube. Cellular lysate was used for all tested SAM domains except Usher1G, LBP1, and Kazrin-3 SAMs. Because of low expression, these negGFP fusion proteins were purified as described in Knight et al., 2011 and pure protein was screened against. Prior to use, the fluorescence of each cellular lysate or purified protein was measured as described previously, using an excitation wavelength of 488nm, an emission wavelength of 510nm and a cutoff filter of 515nm (Knight et al., 2011).

negGFP native gel screen

Initial proof-of-concept gels

Lysates containing wild-type and monomeric point mutants of the negGFP fusions of the SAM domains of TEL and DGKδ were either run alone or mixed in equal volumes with each other and the lysates of the negGFP fusions of TP63, MOB, and Neurabin1. After equilibration at 4°C for 3 hours, RunBlue Native Sample Buffer (Expedeon) was added to a final concentration of 1X. Samples were loaded onto 20% RunBlue 12-well Native gels and run on ice at 4°C for 24 hours at 90V. Gels were visualized using a Bio-Rad Molecular Imager FX Pro-Plus as described previously (Knight et al., 2011) The native gel to test interactions observed in a split-DHFR assay was run similarly, with negGFP fusion protein lysates either prepared individually or mixed in equal volumes prior to equilibration and gel running.

Monomeric SAM domain interaction screen

Lysates containing negGFP fusions of 41 different monomeric SAM domains were organized into 11 "pots" for interaction screening. Pots were prepared to maintain equal amounts of each constituent negGFP fusion protein, based on fluorescence. For screening, lysates were mixed individually with aliquots of each monomeric SAM domain "pot" in 1:1 ratios with each of the "pot" constituents. For each gel, a sample was prepared which contained only the monomeric SAM domain being tested. To control for concentration effects, all samples were brought to 22.5μL using 20mM Tris pH 7.5, 0.3M NaCl, 1mM TCEP for dilution. After equilibrating at 4°C for 3 hours, 7.5μL of 4X RunBlue Native Sample Buffer (Expedeon) was added and samples were loaded onto 20% RunBlue 12-well Native gels (Expedeon), run at 90V for 24 hours at 4°C on ice, and visualized as described above.

Polymeric SAM domain interaction screen

negGFP fusions of the polymeric SAM domains from 23 different proteins were mixed individually with aliquots of the monomeric SAM domain “pots” in 1:2 ratios with each of the “pot” constituents, based on fluorescence. For each gel, a sample was also prepared which contained only the polymeric SAM domain being tested. All samples were diluted to 22.5μL as above to control for concentration effects. Samples were allowed to equilibrate at 4^{o}C for 4 hours prior to the addition of loading dye and running on a native gel, as above.

Figures

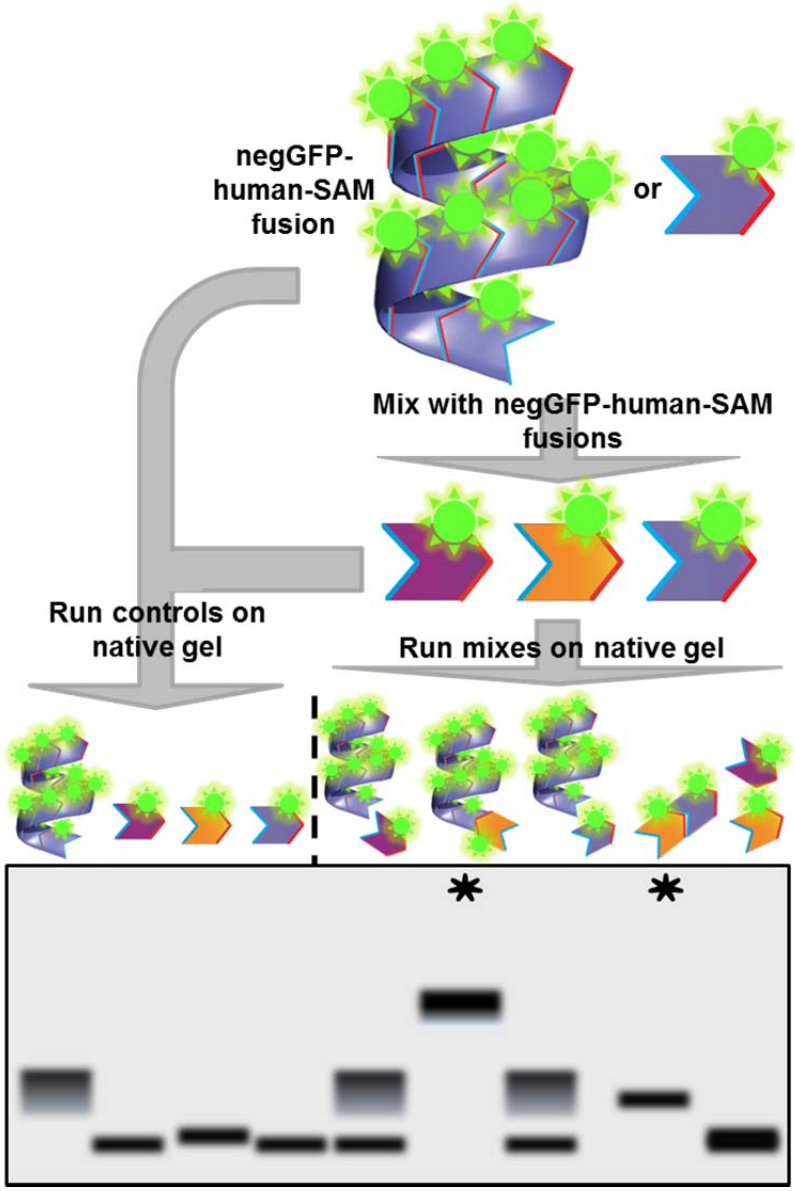


Figure 2-1. Schematic of the negGFP native gel screen to detect SAM domain hetero-interactions.

Polymeric and monomeric human SAM domains fused to negGFP are run either alone or as mixes on a native gel. negGFP fusions of monomeric SAM domains run as discrete bands with higher mobility than negGFP fusions of polymeric SAM domains, which run with reduced mobility and a smeared band character. Hetero-interactions between SAM domains are detected as a band shift accompanied by the disappearance of the bands corresponding to the proteins run alone. Shown in our schematic are expected results for hetero-interactions (marked by *) involving polymeric and monomeric SAM domains. This figure is an adapted version from Leettola, et al., 2014.

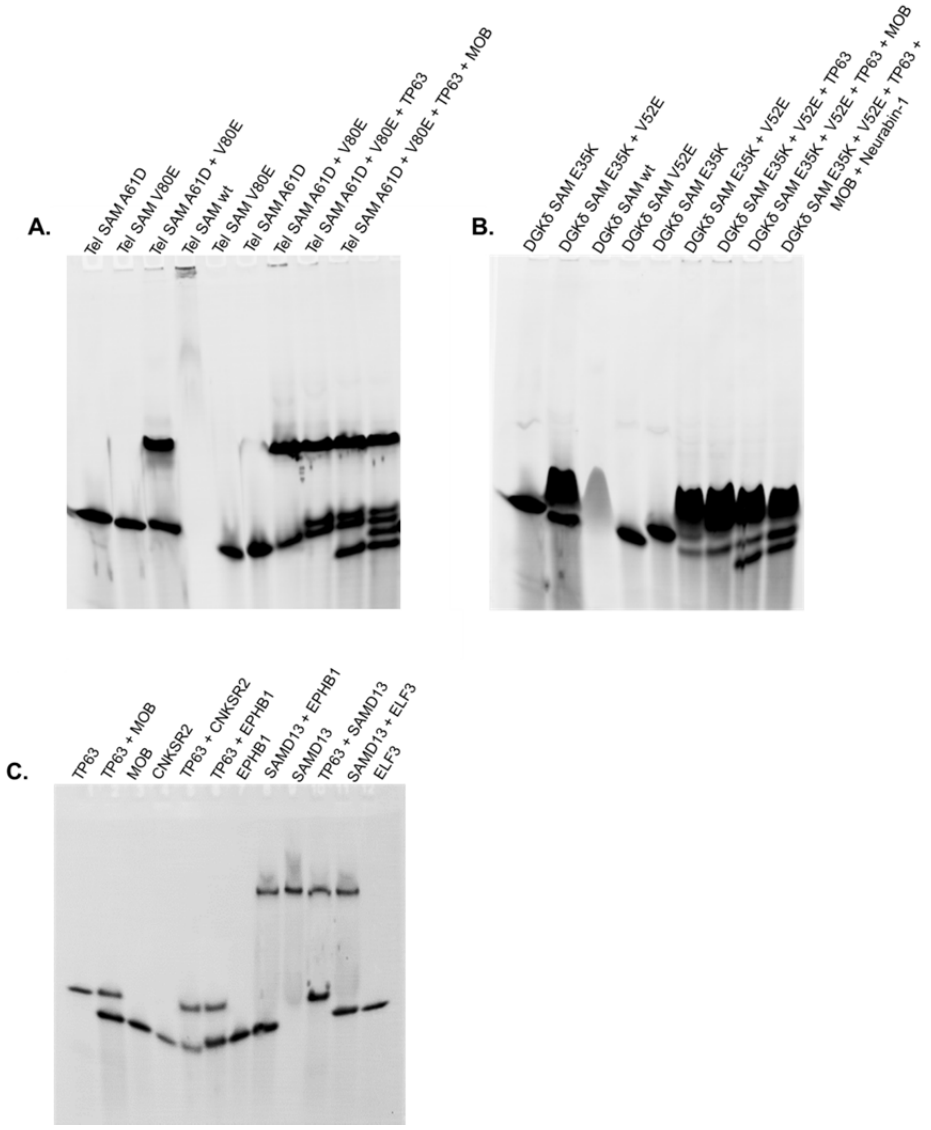


Figure 2-2. negGFP native gel screen proof of concept.

A) negGFP fusions of monomeric and wild-type constructs of the TEL SAM domain run on a native gel. Point mutations at the ML (A61D) and EH (V80E) surfaces cause the strongly polymeric (K_d = 1.7 ± 0.5 nM) TEL SAM domain to become monomeric. A 1:1 ratio mix of these monomeric point mutants forms a heterodimer with obvious gel shift. The presence of additional negGFP SAM domain fusions does not affect the formation or migration of the TEL SAM heterodimer. B) negGFP fusions of monomeric and wild-type constructs of the DGKδ SAM domain run on a native gel. Point mutations at the ML (E35K) and EH (V52E) surfaces cause the weakly polymeric (K_d = 6.3 ± 1.2 μM) DGKδ SAM domain to become monomeric. A 1:1 ratio mix of these monomeric point mutants forms a heterodimer with obvious gel shift. The presence of additional negGFP SAM domain fusions does not affect the formation or migration of the DGKδ SAM heterodimer. C) Heterotypic SAM domain interactions detected in a split-DHFR screen do not form interactions when expressed as negGFP fusions, mixed in 1:1 ratios, and run on a native gel.

A.

"POT"	Human SAM	UNIPROT ID	Residues
1	ARAP1/CentaurinD2	Q96P48	1-70
	ETS1	P14921	53-136
	ELF3	P78545	48-135
	CNKSR2	Q8WXI2	1-77
2	ELF5	Q9UKW6	44-129
	EPHA5	P54756	960-1030
	EPHA7	Q15375	919-998
	GA-binding protein/GABPA	Q06546	170-251
3	EPHA2	P29317	899-970
	ESE3	Q9NZC4	41-120
	BAR	Q9NZS9	177-252
	Neurabin-1	Q9ULJ8	983-1054
4	MOB	Q86VZ5	7-78
	LRSAM1	Q6UWE0	564-634
	SAMSN1	Q9NSI8	237-306
	Usher1G-optimized	Q495M9	381-461
5	EPHA6	Q9UF33	956-1023
	SAMD4B	Q5PRF9	296-361
	ARAP2/CentaurinD1	Q8WZ64	1-71
6	INPPL1	O15357	194-1258
	EPHB1	P54762	906-984
	EPHB2	P29323	906-985
	EPHB6	O15197	945-1021

"POT"	Human SAM	UNIPROT ID	Residues
7	SASH1 SAM1	O94885	625-699
	SASH1 SAM2	O94885	1173-1247
	SASH3	O75995	247-318
	TP63	Q9H3D4	541-609
8	WDSUB1	Q8N9V3	329-396
	Smaug/SAMD4A	Q9UPU9	320-383
	SAMHD1	Q9Y3Z3	1-111
	STIM2	Q9P246	129-205
9	CNKSR3	Q6P9H4	1-73
	SEC23-interacting protein/SEC23IP	Q9Y6Y8	642-703
	SLP76/LCP2	Q13094	10-80
	StarD13-deletion	Q9Y3MB	154-241
10	ANKS6	Q68DC2	771-840
	C14orf174	Q9P1V8	541-610
	DDHD-containing 2	O94830	382-449
11	SAMD12	Q8N8I0	71-145
	ANKS1B/AIDA1C/ cajalin both SAMs	Q7Z6G8	808-950
	ANKS1A/odin both SAMs	Q92625	692-836

B.

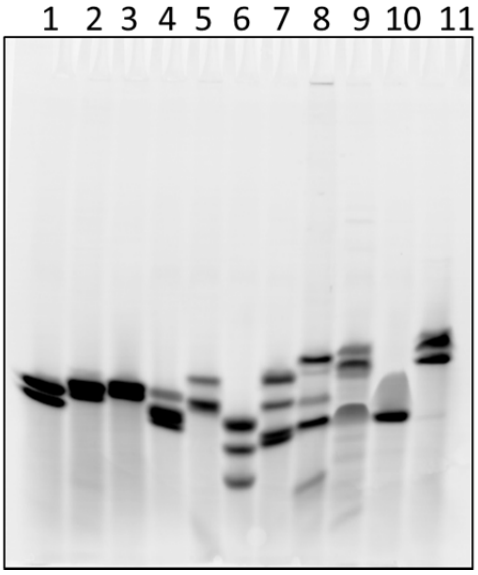


Figure 2-3. Organization of monomeric negGFP human SAM fusions for interaction testing.

A) 41 different monomeric human SAM domains fused to negGFP were organized into "pots" so that multiple SAM hetero-interactions could be tested in each native gel lane. The human SAM domains in each "pot" are listed, along with the Uniprot ID and the cloned residues encompassing each protein's SAM domain. B) Each "pot" of monomeric negGFP human SAM domains run on a native gel. Numbering corresponds to the "pot" in each lane.

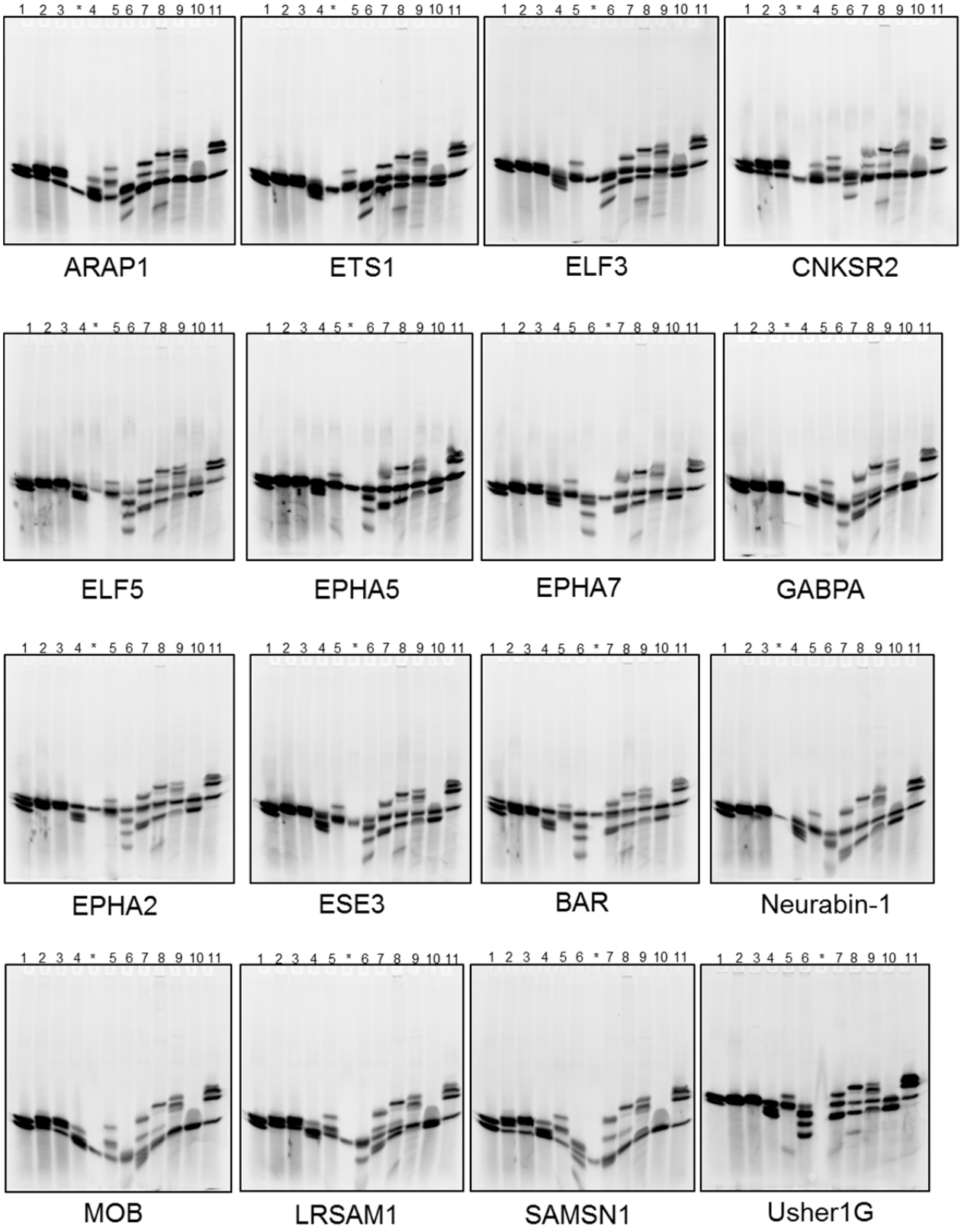


Figure 2-4. Human monomeric SAM domain hetero-interaction screen. *Continued on next page.*

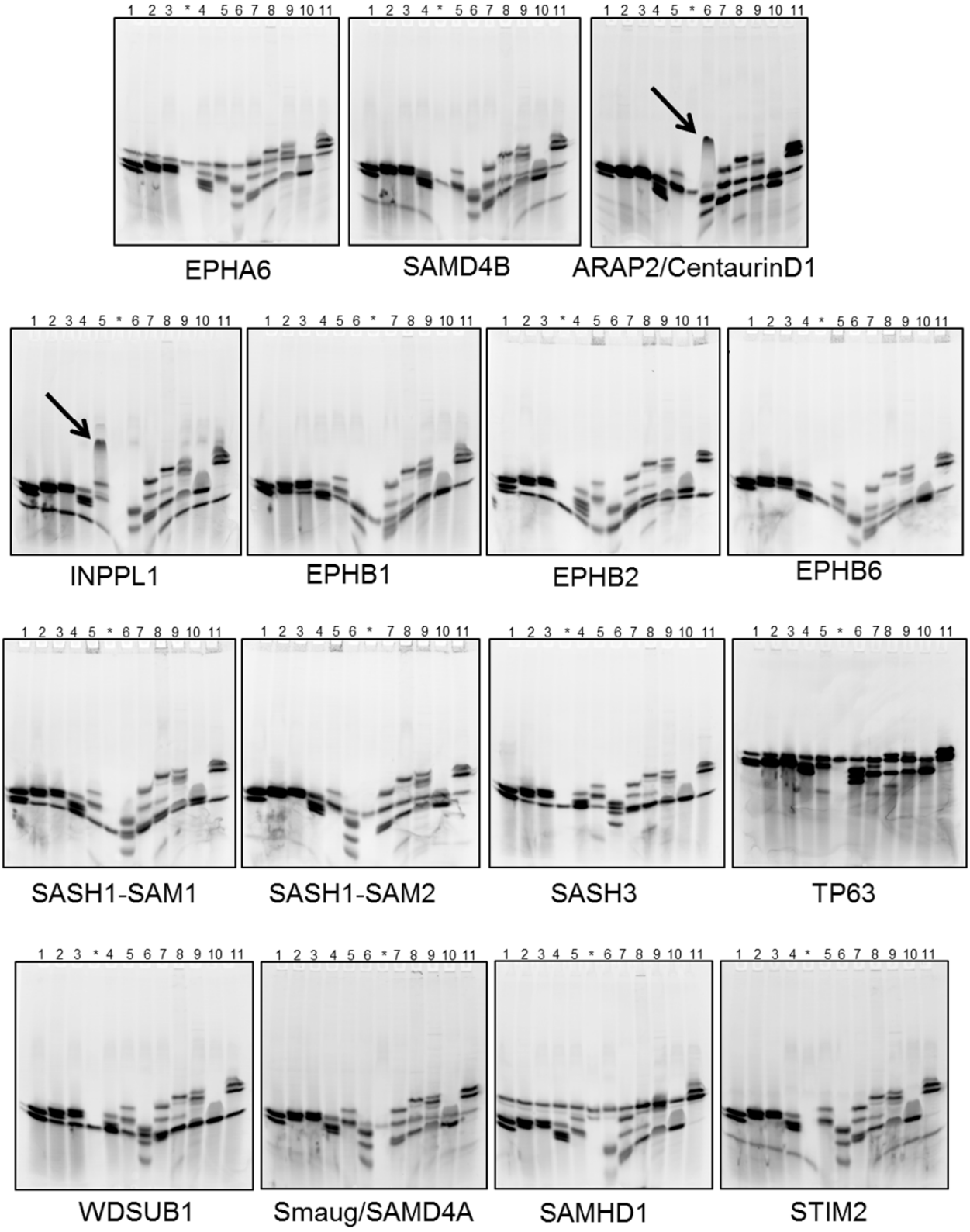


Figure 2-4. Human monomeric SAM domain hetero-interaction screen. *Continued on next page*

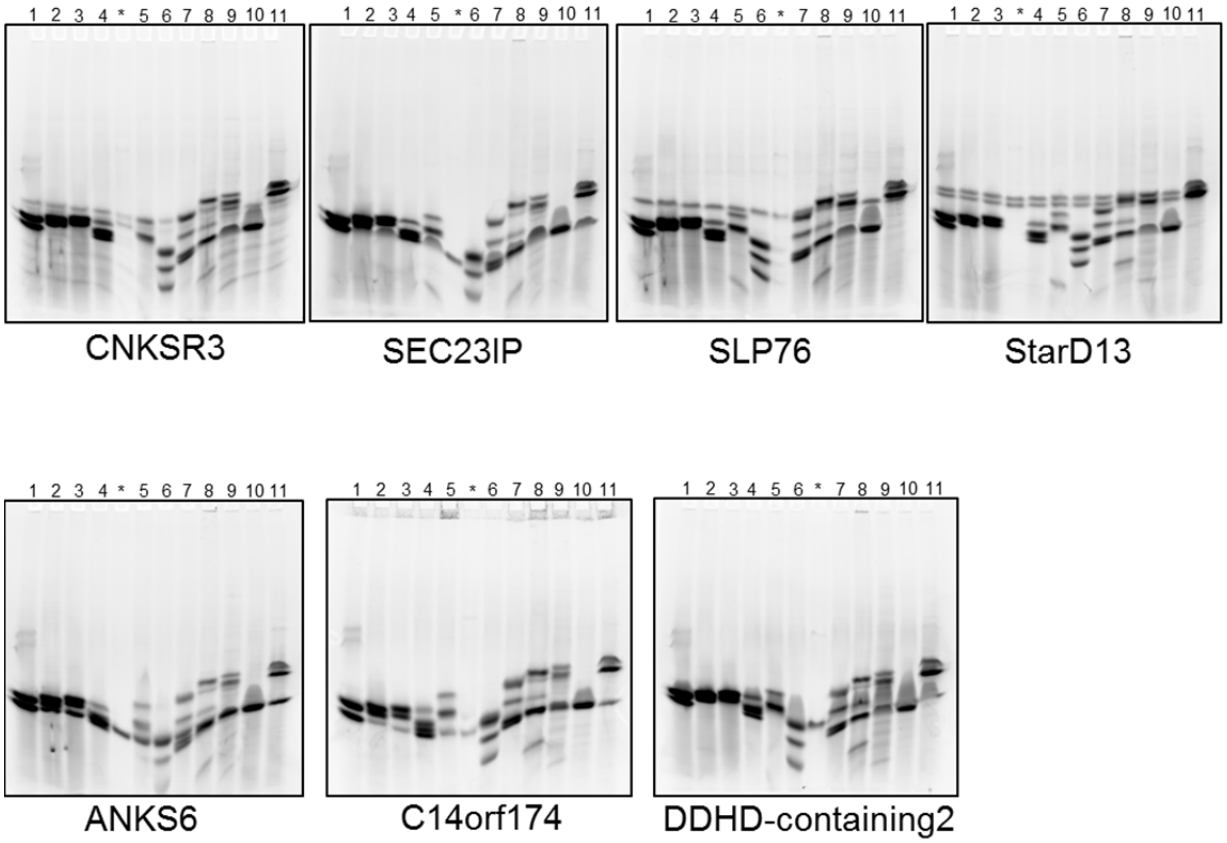


Figure 2-4. Human monomeric SAM domain hetero-interaction screen.

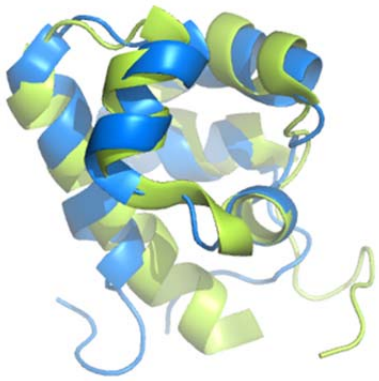
On each gel, a single monomeric human SAM domain negGFP fusion is mixed with the "pots" (Fig. 2-3) containing 41 different human monomeric SAM domains. The banding pattern was compared against a gel run with the monomeric "pots" alone (Fig. 2-3B) to identify band shifts or changes, indicative of an interaction. From this screen only band shifts for INPPL1 and ARAP2/CentaurinD1 were identified (arrows). Numbering above each lane indicates the monomeric SAM "pot" in that lane. * = control lanes with the SAM domain being screened for run alone.

A.

```
ARAP3                MAAPQD--LDIAVWLATVHLEQYADTFRRHGLATAGAARGLGHEELKQLGISATGHRKRI 58
ARAP2/CentaurinD1    MSSVSEVNVDIKDFLMSINLEQYLLHFHESGFTTVKDCAAINDSLLQKIGISPTGHRRRI 60
ARAP1/CentaurinD2    MAEAGDAALSVAEWLRALHLEQYTGLFEQHGLVWATECQGLSDTRLMDMGMLLPGHRRRI 60

ARAP3                LR-LLQTGTEEGS 70
ARAP2/CentaurinD1    LKQLQIILSKM- 71
ARAP1/CentaurinD2    LAGLLRAHTS--- 70
```

B.



ARAP3
ARAP2/CentaurinD1

Figure 2-5. ARAP2/CentaurinD1 is homologous with ARAP3

A) Multiple sequence alignment of the SAM domains of ARAP3 [Uniprot: Q8WWN8, residues 1-70], ARAP2/CentaurinD1 [Uniprot:Q8WZ64, residues 1-71], and ARAP1/CentaurinD2 [Uniprot:Q96P48, residues 1-70]. The SAM domain of ARAP2/CentaurinD1 is 40% identical and 60% similar to the SAM domain of ARAP3. The SAM domain of ARAP1/CentaurinD2 is 45% identical and 58% similar to the SAM domain of ARAP3. Residues of ARAP3 identified by Leone et al., 2009 as being involved in binding the SAM domain of INPPL1 are shaded blue. Of these residues, those that are conserved in ARAP2/CentaurinD1 and ARAP1/CentaurinD2 have been shaded grey. B) Alignment of the structures of the SAM domains of ARAP3 (PDB:2KG5) shown in blue and ARAP2/CentaurinD1 (PDB:1X40) shown in green. The structures align closely (RMSD across all atoms = 1.085Å, RMSD across backbone atoms = 0.776Å).

	Human SAM	UNIPROT ID	Residues
Weak Polymer	ANKS3	Q6ZW76	421-490
	DGKδ	Q16760	1141-1214
	SAMD9	Q5K651	1-80
	SHANK2	Q9UPX8	1402-1470
Medium Polymer	BICC1	Q9H694	867-939
	Caskin1-both SAMS	Q8WXD9	466-607
	Caskin2-both SAMS	Q8WXE0	486-624
	LMBTL3	Q96JM7	703-780
	SCMH1	Q96GD3	587-660
Strong Polymer	Atherin/SAMD1	Q6SPF0	457-538
	PH1	P78364	939-1003
	PH2	Q8IXK0	784-858
	PH3	Q8NDX5	919-983
	PH-like1/PHC1	Q6GMQ3	888-957
	SamD13-long	Q5VXD3	21-122
	SAMD13-short	Q5VXD3	47-116
	SAMD3	Q8N6K7	1-67
	SAMD8	Q96LT4	1-79
	SARM1-both SAMS	Q6SZW1	408-548
	TEL2/ETV7	Q9Y603	41-115
Very Strong Polymer	TEL/ETV6	P41212	38-125
	SCML1	Q9UN30	252-329
Low Expression	LBP1	Q9NZI7	367-437
	Kazrin - 3 SAMs	Q674X7	443-679

Figure 2-6. Polymeric negGFP human SAM domain fusions tested for hetero-interactions.

negGFP fusions of polymeric human SAM domains tested for hetero-interactions with monomeric human SAM domains are organized by strength of polymer formation. For constructs labeled as low expression, purified protein instead of cellular lysate was used in the negGFP native gel screen. The Uniprot ID and cloned residues encompassing each protein's SAM domain are listed. Constructs labeled as "both SAMs" or "3 SAMs" have multiple SAM domains in tandem.

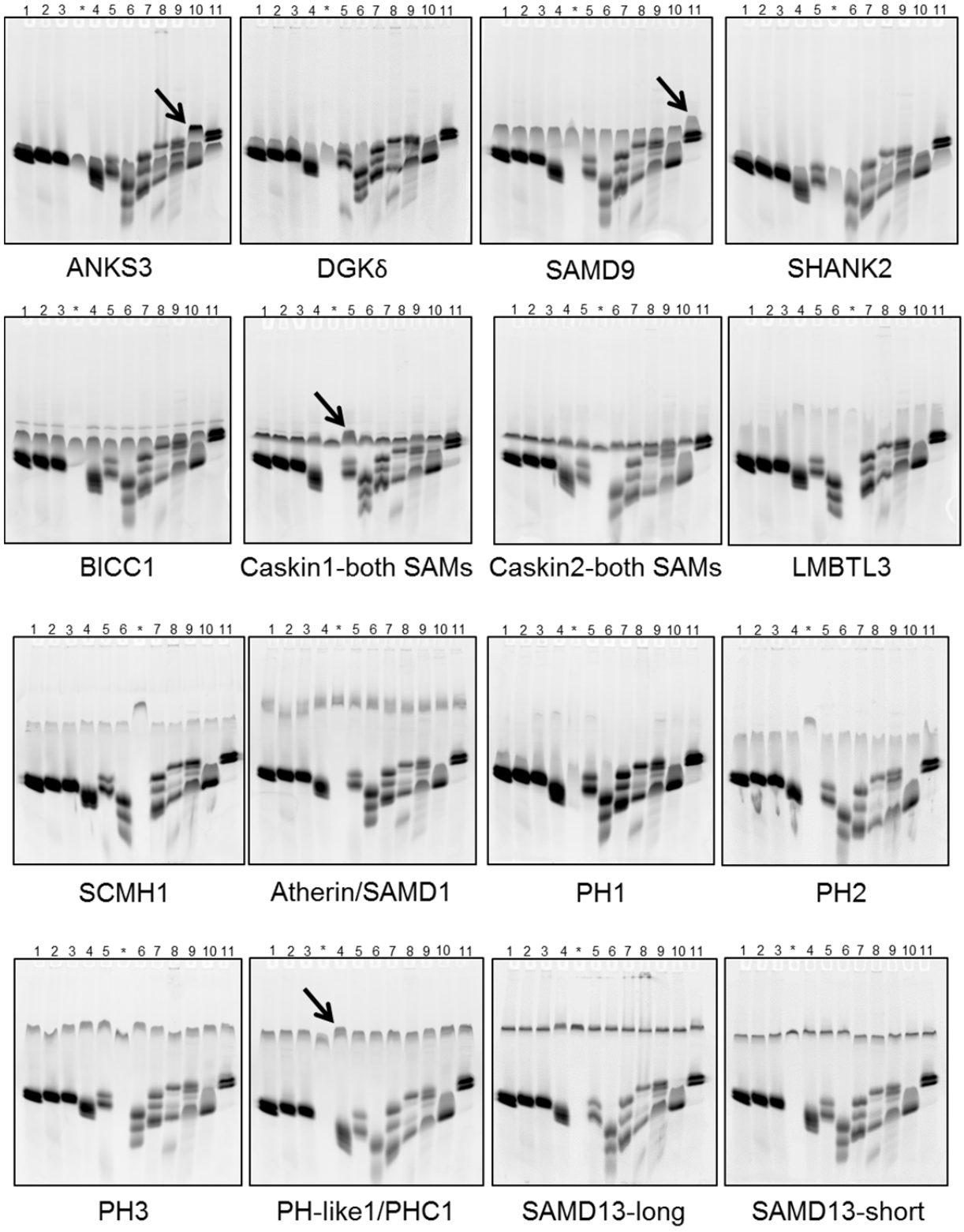


Figure 2-7. Human polymeric SAM domain hetero-interaction screen.

Continued on next page.

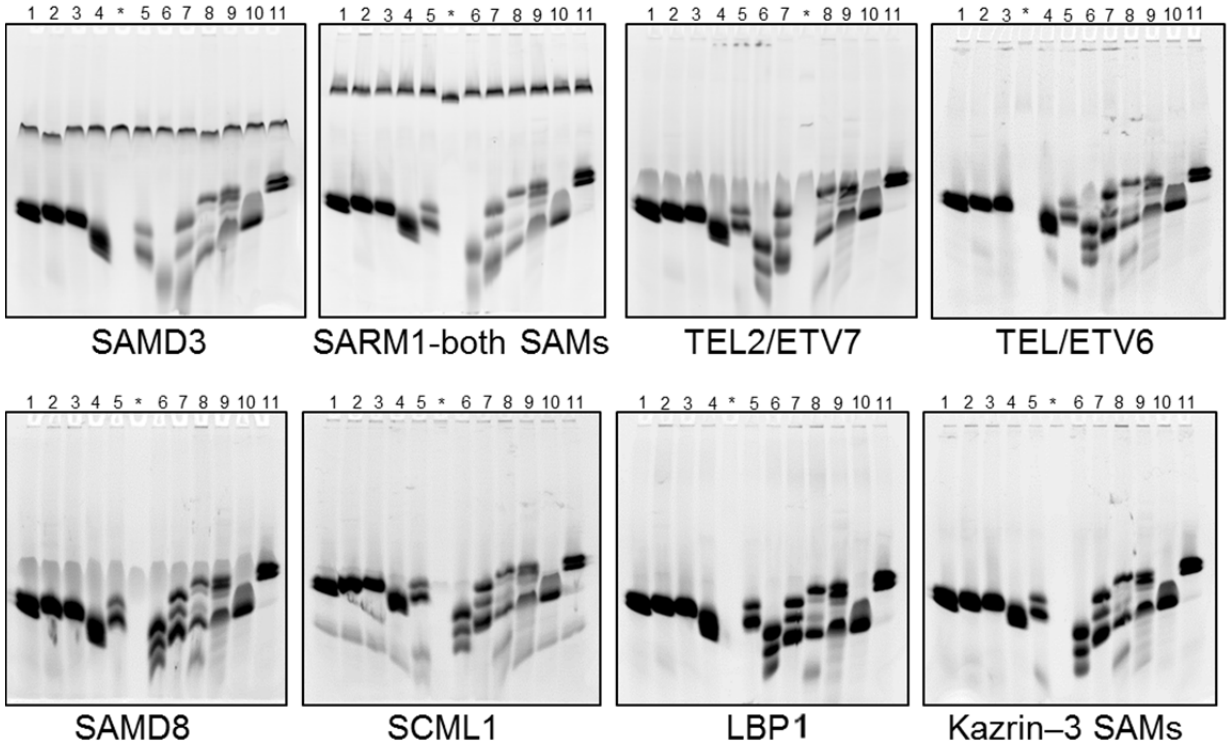


Figure 2-7. Human polymeric SAM domain hetero-interaction screen.

On each gel, a single polymeric human SAM domain negGFP fusion is mixed with the "pots" (Fig. 2-3) containing 41 different human monomeric SAM domains. The banding pattern was compared against a gel run with the monomeric "pots" alone (Fig. 2-3B) to identify band shifts or changes, indicative of an interaction. Numbering above each lane indicates the monomeric SAM "pot" in that lane. * = control lanes with the SAM domain being screened for run alone. Arrows indicate possible hetero-interactions.

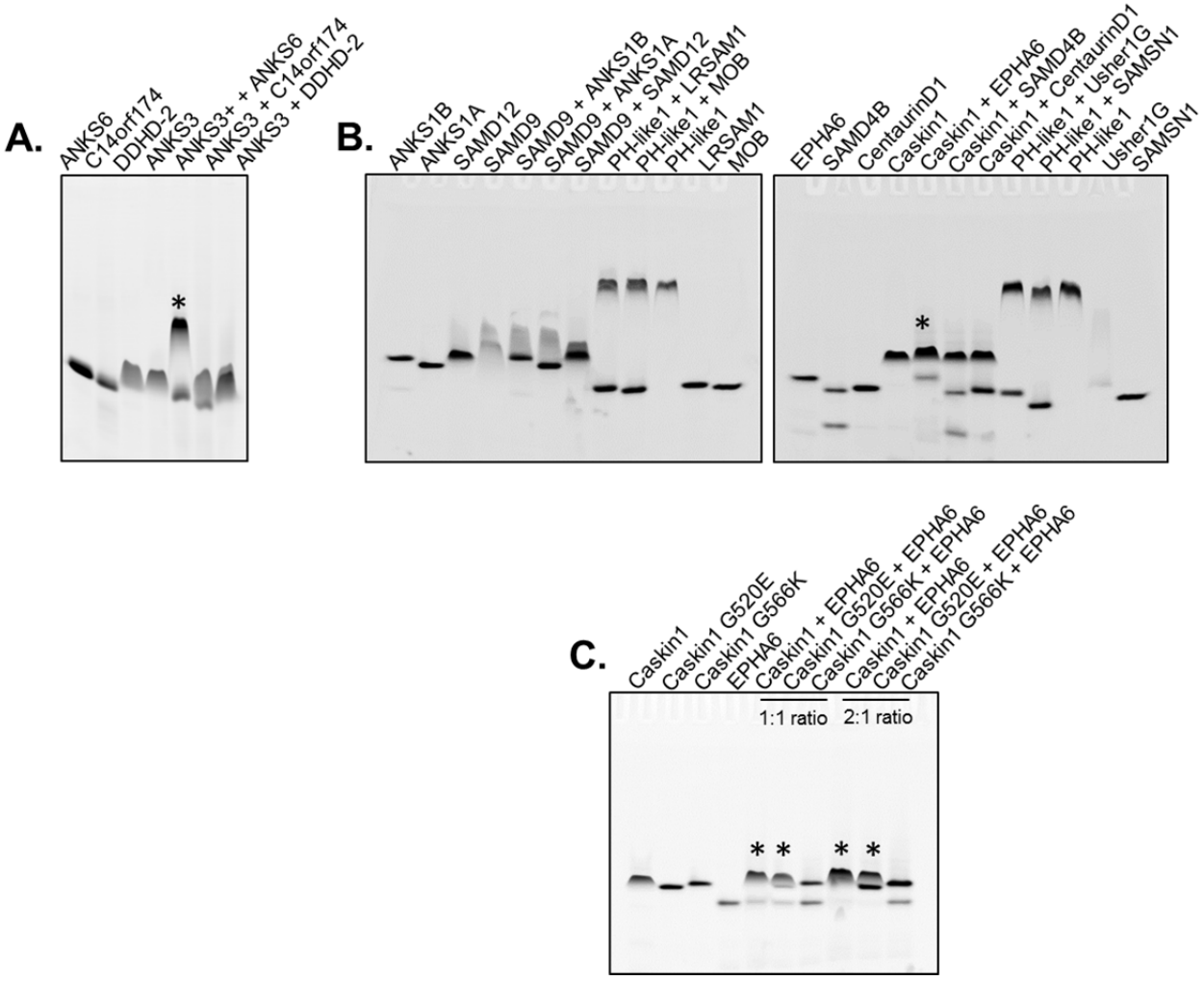


Figure 2-8. Novel SAM domain hetero-interactions.

A) The SAM domains of ANKS3 and ANKS6 interact. When the weakly polymeric SAM domain of ANKS3 is mixed in a 1:2 ratio based on fluorescence with the monomeric SAM domain of ANKS6, the appearance of a new band (*) accompanied by the disappearance of the ANKS3 band indicates an interaction. B) The SAM domains of SAMD9, Caskin1, and PH-like1 were mixed in 1:1 ratios based on fluorescence with monomeric SAM domains suspected of forming interactions. A gel shift accompanied by weakened intensity of the monomeric SAM domain band was only observed for Caskin1 mixed with EPHA6 (indicated by *). The Caskin1 construct contains two SAM domains in tandem. C) Wild-type and mutant Caskin1 tandem SAM domains were mixed in 1:1 and 2:1 ratios, based on fluorescence, with the SAM domain of EPHA6. Interactions (indicated by *) of the wild-type and G520E mutant of Caskin1 with EPHA6 are visible as a slight gel shift accompanied by disappearance of the EPHA6 band. The point mutation G520E at the free EH surface of the first SAM domain of Caskin1 does not affect binding to EPHA6. The point mutation G566K at the free ML surface of the second SAM domain of Caskin1 prevents interaction with EPHA6.

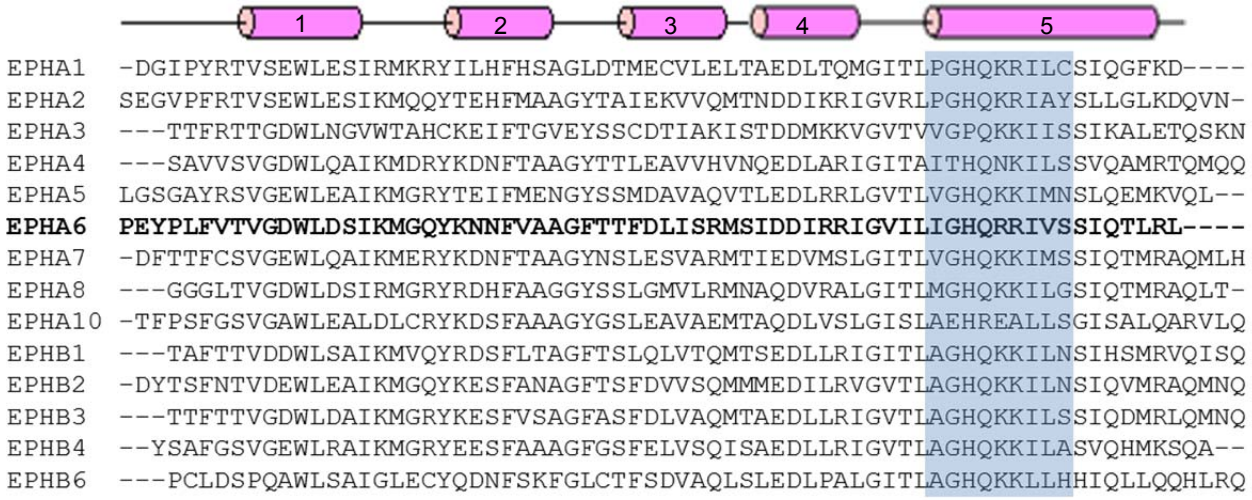


Figure 2-9. Alignment of human Ephrin Receptor family member SAM domains.

Multiple sequence alignment of the SAM domains from human EPH family members. A secondary structure prediction performed using PSIPRED is shown for the EPHA6 SAM domain and indicates the position of 5 helices, which are numbered. Residues which correspond to the EH surface (blue shading) are highly conserved among family members.

References

Aviv, T., Lin, Z., Rendl, L., M., Sicheri, F., and Smibert, C.A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. Nat. Struct. Biol. *10*, 614–621.

Baron, M.K., Boeckers, T.M., Vaida, B., Faham, S., Gingery, M., Sawaya, M.R., Salyer, D., Gundelfinger, E.D., and Bowie, J.U. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. Science *311*, 531–535.

Barrera, F.N. (2003). Binding of the C-terminal Sterile Motif (SAM) Domain of Human p73 to Lipid Membranes. J. Biol. Chem. *278*, 46878–46885.

Bhunia, A., Domadia, P.N., Mohanram, H., and Bhattacharjya, S. (2009). NMR structural studies of the Ste11 SAM domain in the dodecyl phosphocholine micelle. Proteins Struct. Funct. Bioinforma. *74*, 328–343.

Cuthbert, E.J., Davis, K.K., and Casanova, J.E. (2008). Substrate specificities and activities of AZAP family Arf GAPs in vivo. AJP Cell Physiol. *294*, C263–C270.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S., and Bowie, J.U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J. *20*, 4173–4182.

Knight, M.J. (2010). Sterile Alpha Motif Domain Oligomerization. University of California, Los Angeles.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Kwan, J.J., Warner, N., Pawson, T., and Donaldson, L.W. (2004). The Solution Structure of the S.cerevisiae Ste11 MAPKKK SAM Domain and its Partnership with Ste50. J. Mol. Biol. *342*, 681–693.

Lawrence, M.S., Phillips, K.J., and Liu, D.R. (2007). Supercharging Proteins Can Impart Unusual Resilience. J. Am. Chem. Soc. *129*, 10110–10112.

Leettola,, C.N., Knight, M.J., Cascio, D., Hoffmann, S., and Bowie, J.U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. BMC Struct. Biol. *14*.

Leone, M., Cellitti, J., and Pellecchia, M. (2009). The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. BMC Struct. Biol. *9*, 59.

Meruelo, A.D., and Bowie, J.U. (2009). Identifying polymer-forming SAM domains. Proteins Struct. Funct. Bioinforma. *74*, 1–5.

Pawson, T., Raina, M., and Nash, P. (2002). Interaction domains: from simple binding events to complex cellular behavior. FEBS Lett. *513*, 2–10.

Qiao, F., Song, H., Kim, C.A., Sawaya, M.R., Hunter, J.B., Gingery, M., Rebay, I., Courey, A.J., and Bowie, J.U. (2004). Derepression by depolymerization: structural insights into the regulation of Yan by Mae. Cell *118*, 163–173.

Raaijmakers, J.H., Deneubourg, L., Rehmann, H., de Koning, J., Zhang, Z., Krugmann, S., Erneux, C., and Bos, J.L. (2007). The PI3K effector Arap3 interacts with the PI(3,4,5)P3 phosphatase SHIP2 in a SAM domain-dependent manner. Cell. Signal. *19*, 1249–1257.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Rufini, S., Lena, A.M., Cadot, B., Mele, S., Amelio, I., Terrinoni, A., Desideri, A., Melino, G., and Candi, E. (2011). The sterile alpha-motif (SAM) domain of p63 binds in vitro monoasialoganglioside (GM1) micelles. Biochem. Pharmacol. *82*, 1262–1268.

Seidel, J.J., and Graves, B.J. (2002). An ERK2 docking site in the Pointed domain distinguishes a subset of ETS transcription factors. Genes Dev. *16*, 127–137.

Smalla, M., Schmieder, P., Kelly, M., Ter Laak, A., Krause, G., Ball, L., Wahl, M., Bork, P., and Oschkinat, H. (1999). Solution structure of the receptor tyrosine kinase EphB2 SAM domain and identification of two distinct homotypic interaction sites. Protein Sci. *8*, 1954–1961.

Stafford, R.L., Hinde, E., Knight, M.J., Pennella, M.A., Ear, J., Digman, M.A., Gratton, E., and Bowie, J.U. (2011). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK. Structure *19*, 1826–1836.

Stapleton, D., Balan, I., Pawson, T., and Sicheri, F. (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. Nat. Struct. Biol. *6*, 44–49.

Thanos, C.D., Goodwill, K.E., and Bowie, J.U. (1999). Oligomeric Structure of the Human EphB2 Receptor SAM Domain. Science *283*, 833–836.

Wu, Y., Koharudin, L.M.I., Mehrens, J., DeLucia, M., Byeon, C.-H., Byeon, I.-J.L., Calero, G., Ahn, J., and Gronenborn, A.M. (2015). Structural basis of clade-specific engagement of SAMHD1 restriction factors by lentiviral Vpx virulence factors. J. Biol. Chem. doi: 10.1074/jbc.M115.665513.

Xia, K., Fu, Z., Hou, L., and Han, J.-D.J. (2008). Impacts of protein-protein interaction domains on organism and network complexity. Genome Res. *18*, 1500–1508.

Yan, J., Pan, L., Chen, X., Wu, L., and Zhang, M. (2010). The structure of the harmonin/sans complex reveals an unexpected interaction mode of the two Usher syndrome proteins. Proc. Natl. Acad. Sci. *107*, 4040–4045.

Chapter 3

Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3

Abstract

Autosomal dominant polycystic kidney disease (ADPKD) is the most common genetic disorder leading to end-stage renal failure in humans. In the PKD/Mhm(cy/+) rat model of ADPKD, the point mutation R823W in the sterile alpha motif (SAM) domain of the protein ANKS6 is responsible for disease. SAM domains are known protein-protein interaction domains, capable of binding each other to form polymers and heterodimers. Despite its physiological importance, little is known about the function of ANKS6 and how the R823W point mutation leads to PKD. Recent work has revealed that ANKS6 interacts with a related protein called ANKS3. Both ANKS6 and ANKS3 have a similar domain structure, with ankyrin repeats at the N-terminus and a SAM domain at the C-terminus. Herein we identify the SAM domain of ANKS3 as a direct binding partner of the ANKS6 SAM domain. We find that ANKS3-SAM polymerizes and ANKS6-SAM can bind to one end of the polymer. We present crystal structures of both the ANKS3-SAM polymer and the ANKS3-SAM/ANKS6-SAM complex, revealing the molecular details of their association. We also learn how the R823W mutation disrupts ANKS6 function by dramatically destabilizing the SAM domain such that the interaction with ANKS3-SAM is lost. By structurally and biochemically characterizing the interaction between the ANKS3 and ANKS6 SAM domains, our work provides a basis for future investigation of how the interaction between these proteins mediates kidney function.

Introduction

Autosomal dominant polycystic kidney disease (ADPKD) is the most common inherited renal cystic disease, with a prevalence of approximately 1 in 1,000 individuals (Chapin and Caplan, 2010; Torres et al., 2007; Wilson, 2004). ADPKD is characterized by the progressive formation of fluid-filled cysts within the kidney which ultimately disrupt renal function, as well as various extrarenal manifestations (Gabow, 1993; Torres et al., 2007). Mutations in two genes, *PKD1* (polycystic kidney disease-1) and *PKD2* (polycystic kidney disease-2) account for approximately 85% and 15% of disease cases, respectively (Chapin and Caplan, 2010). The occurrence of disease in patients lacking mutations in either of these genes suggests the involvement of other genetic loci, but this is still uncertain (Ariza et al., 1997). *PKD1* and *PKD2* encode the proteins polycystin-1 (PC1) and polycystin-2 (PC2), respectively. PC1 is a multidomain membrane receptor capable of binding and interacting with proteins, lipids, and carbohydrates and stimulating intracellular signaling pathways. PC2 is a membrane protein which acts as a Ca^{2+}-permeable non-selective cation channel. PC1 and PC2 can interact, and in the process, modulate each other's activity (Chapin and Caplan, 2010; Hanaoka et al., 2000).

Among the animal models of polycystic kidney disease, the PKD/Mhm(cy/+) rat recapitulates many of the hallmarks of human ADPKD (Gretz et al., 1996; Guay-Woodford, 2003; Nagao et al., 2012; Schafer et al., 1994). The Cy mutation was found to be a missense mutation in the *ANKS6* gene, encoding ankyrin repeat and SAM-domain containing protein 6 (ANKS6) (Brown et al., 2005). Human ANKS6 is an 871 amino acid protein containing 11 ankyrin repeats at its N-terminus and a SAM domain near its C-terminus (Fig. 3-1A). The Cy mutation occurs in the SAM domain and generates an arginine to tryptophan mutation at amino acid 823 (Brown et al., 2005). The R823W point mutation acts in a dominant-negative fashion, as evidenced by the PKD phenotype of transgenic rats over-expressing mutated *ANKS6*$^{(p.R823W)}$ (Neudecker et al., 2010).

Recently ANKS6 has also been implicated in human renal disease. In particular, the work of Hoff and Halbritter *et al*, (2013) has placed ANKS6 as a central node in a network of nephronophthisis (NPHP)-associated proteins, including direct binding between the ankyrin-repeat domain of ANKS6 and NEK8, a NimA (never in mitosis A)-related serine-threonine kinase which is mutated in nephronophthisis. Moreover, 8 individuals from 6 different families bearing homozygous mutations in *ANKS6* all presented with nephronophthisis, which is essentially an infantile or juvenile onset of PKD (Hoff et al., 2013). While several of the detected mutations localize to the ankyrin repeats, one generates a truncation within the SAM domain at Tyr790. Thus, the SAM domain has been shown in both rats and humans to be essential for normal ANKS6 function.

SAM domains consist of approximately 70 amino acids and adopt a globular structure generally containing a core of five α-helices (Kim and Bowie, 2003; Qiao and Bowie, 2005). Most SAM domains that have been characterized are protein-protein interaction modules that either self-associate (Baron, 2006; Harada et al., 2008; Kim et al., 2001, 2002; Knight et al., 2011; Di Pietro et al., 2010; Stafford et al., 2011), bind to other SAM-domain containing proteins (Grimshaw et al., 2003; Kwan et al., 2004; Leone et al., 2009; Qiao and Bowie, 2005; Qiao et al., 2004; Ramachander, 2002), or bind other proteins altogether (Qiao and Bowie, 2005; Qiao et al., 2005; Zhang et al., 2000). Some SAM domains have been found to bind RNA and lipids, however (Aviv et al., 2003; Barrera, 2003). The protein-protein interactions of SAM domains are typically mediated by two distinct surfaces on the domain, termed the mid-loop (ML) and end-helix (EH) surfaces. SAMs can bind each other via their ML and EH surfaces, to generate open-ended polymers (Baron, 2006; Harada et al., 2008; Kim et al., 2001, 2005; Qiao et al., 2004), closed oligomers (Qiao et al., 2004; Ramachander, 2002), and heterodimers (Leone et al., 2008, 2009; Qiao et al., 2004; Rajakulendran et al., 2008). Through polymer formation and heterotypic interactions, SAM domains confer a diverse array of biological functions including

gene regulation (Kim et al., 2001; Qiao et al., 2004), enzyme localization (Harada et al., 2008; Knight et al., 2010), and scaffolding (Baron, 2006; Gundelfinger et al., 2006; Stafford et al., 2011).

In spite of the important role of the ANKS6-SAM domain in cystic kidney disease, the function of ANKS6 remains unknown. Prior work using a proteomics screen of tandem affinity purified (TAP)-tagged NPHP-associated proteins found that ANKS6 and ankyrin repeat and SAM-domain containing protein 3 (ANKS3) are potential binding partners (Hoff et al., 2013). ANKS3 has a similar domain structure to ANKS6 including a C-terminal SAM domain (Fig. 3-1 A). Here we define the ANKS6/ANKS3 interaction by discovering that the ANKS6 SAM domain binds to the SAM domain of ANKS3. We show that ANKS3-SAM forms polymers and that ANKS6 binds to one end of these polymers. The R823W mutation was found to disrupt the structure of the ANKS6 SAM domain and negatively affects binding to ANKS3-SAM. Our results provide a structural explanation for the defect in PKD/Mhm(cy/+) rats and a potential new pathway to cystic disease via the SAM domain of ANKS3.

Results

negGFP Native Gel Screen Identifies ANKS3-SAM + ANKS6-SAM Interaction

To identify new human SAM domain hetero-interactions involving ANKS3, we employed a rapid screen for binding activity, outlined in Fig. 3-2A. SAM domains were fused to an engineered green fluorescent protein modified to have a net charge of -30 (negGFP) (Knight et al., 2011; Lawrence et al., 2007). The high charge on the negGFP effectively solubilizes even insoluble proteins, and leads to consistent migration of the fusion proteins towards the cathode on a native gel (Knight et al., 2011). Binding between two negGFP fusion proteins is detected as the appearance of a new band with retarded migration on native gels. Because the fusions are fluorescent, assays can be performed using crude extracts. Thus, many different combinations of proteins can be rapidly screened using this technique.

Extract containing negGFP-ANKS3-SAM fusion protein was mixed with extracts containing negGFP-SAM fusions from 40 different human SAM domain-containing proteins (Table 3-2) and binding was tested by native gel electrophoresis. Among the human SAMs assayed, we detected a novel interaction between ANKS3-SAM and ANKS6-SAM (Fig. 3-2B). Consistent with prior results, negGFP-ANKS6-SAM appears monomeric as it runs as a discrete band with migration similar to other monomeric SAM domains (Knight et al., 2011). In contrast, the negGFP-ANKS3-SAM fusion runs as a more diffuse band with slightly slower migration, behavior that was observed previously and is typical of weakly polymeric SAM domains (Knight et al., 2011). When negGFP-ANKS3-SAM and negGFP-ANKS6-SAM were mixed in a 1:1 ratio, the appearance of a new slower migrating species indicates that ANKS3-SAM and ANKS6-SAM bind to each other. We therefore decided to investigate the polymeric character of ANKS3-SAM and the novel ANKS3-SAM + ANKS6-SAM hetero-interaction further.

ANKS3-SAM is Polymeric

We first chose to characterize the ANKS3-SAM domain and investigate whether ANKS3-SAM forms a polymer. In a prior screen for SAM polymers, negGFP-ANKS3-SAM was observed by negative stain transmission electron microscopy (TEM) to form short polymeric structures (Knight et al., 2011). We re-examined the negGFP-ANKS3-SAM fusion and obtained results consistent with earlier work, showing short polymers 11.4 ± 2.2 nm wide on average and varying in length from 20-90 nm, with an average length of approximately 36 nm (Fig. 3-3A).

Since the addition of negGFP in the SAM fusion proteins introduces charge repulsion that weakens SAM domain interaction and polymerization, we next examined the ANKS3-SAM domain by itself. Without the negGFP fusion, ANKS3-SAM was much less soluble and precipitated after purification. We examined the precipitate by TEM and saw enormous sheets of polymers, some extending more than 1 μm long and 0.4 μm wide (Figs. 3-3B-D). Individual polymers within the fiber-like sheets were approximately 4-6 nm wide. The ability of SAM domains to organize as sheets of polymers has been seen previously with the SAM domains of DGKδ and Shank3 (Baron, 2006; Knight et al., 2010), however it is unclear whether sheet formation is physiologically relevant for ANKS3-SAM.

Mapping the ANKS3-SAM Polymer

To determine the interfaces of ANKS3-SAM responsible for polymer formation, we employed our negGFP binding assay to rapidly screen for point mutants that blocked polymerization. A similar approach was used in our previous study of the Caskin1 tandem SAM domains (Stafford et al., 2011). We targeted putative ML and EH surface residues to find those that yielded faster migration on a native gel. As shown in Fig. 3-4A, the mutations D31K, I36E

and E47K on the ML-surface and L52E, F53E and K58E on the EH-surface inhibited polymerization. Thus, ANKS3-SAM appears to form a polymer using an interface common to other SAM polymers (Qiao and Bowie, 2005).

The identification of monomeric ANKS3-SAM mutants allowed us to measure binding affinity between subunits using surface plasmon resonance (SPR). To measure affinity of the native interface, we immobilized an ANKS3-SAM EH-surface mutant, F53E, on an SPR chip and detected equilibrium binding to an ML-surface mutant, I36E. At 0.15M NaCl, we observe hyperbolic binding with a K_d of 5.8 ± 0.4 μM (Fig. 3-4D). As many of the mutations that reduce polymerization involve charged residues, we also examined the salt dependence of binding. As shown in Fig. 3-4F, where the salt sensitivity of a protein-protein interaction is indicated by the slope of a log K_d versus log [salt] plot (Grucza et al., 2000; Hileman et al., 1998), the binding affinity is indeed strongly dependent on salt concentration.

Structure of the ANKS3-SAM Polymer

To better understand how the ANKS3-SAM domain forms polymers we sought a crystal structure of the polymer. Because the wild-type SAM domain is relatively insoluble, forming heterogeneous polymers, it cannot be crystallized directly. We therefore used a strategy that has proven successful for a number of other SAM domain polymers, where we attempt to crystallize SAM domains with mutations in the polymer interface (Harada et al., 2008; Kim et al., 2001; Qiao et al., 2004). The mutations weaken subunit association so that the protein remains soluble during purification but under the high concentrations required for crystallization, the polymer interface remains a favorable site for crystal contacts, thereby generating the polymer in the crystal.

We were able to obtain a crystal structure of the L52A mutant (Fig. 3-5). ANKS3-SAM L52A crystallized in space group $P4_1$ with two molecules in the asymmetric unit and the structure was solved to 1.6 Å resolution. Each SAM domain has the characteristic five α-helical fold and chains in the asymmetric unit are nearly identical (RMSD of 0.29 Å on backbone atoms and 0.3 Å on all atoms). Examination of the crystal packing reveals a triple helix of intertwined SAM polymers (Fig. 3-5A). Individual SAM polymers contain 8 SAMs per helical repeat, with each helical repeat measuring 72Å in diameter and 100Å in length. As expected from the mutational studies described above, the ANKS3 SAM domains associate via sequential interactions of ML and EH surfaces (Fig. 3-5B). The ML-surface, formed by residues spanning loop 2 through helix 4, is composed of a shallow hydrophobic patch (residues V32, I36, L40, and I48) flanked by negatively charged residues (D31, D33, D44, E47). The EH-surface encompasses the N-terminal portion of helix 5 and contains a critical Phe (F53) which packs against the hydrophobic patch of a neighboring ML-surface. Several positively charged residues (K22, K56, R57, K58) surround F53, creating a ring of positive charge which binds the ring of negative charge on a neighboring ML-surface through the formation of salt bridges (K22 + D31, K56 + E47, R57 + D44). The striking asymmetric charge distribution is consistent with the strong salt dependence of subunit association (see above).

The crystal structure is consistent with the identified polymer-blocking mutants (Figs. 3-4A, 3-5B). F53E removes a key hydrophobic contact of the EH-surface and I36E introduces steric overlap and disruption of the ML-surface. The L52A mutation used in the crystal structure is found at the end of helix 5 and appears to remove van der Waals packing and complementary hydrophobic surface at the interface. The mutations D31K and E47K remove ionic interactions. Finally, K58 was also found crucial for polymerization and although it does not form a direct salt bridge in the crystal structure, this residue is near the interface and helps to maintain charge complementarity. It is also possible that the interface is altered somewhat in the mutant

structure, weakening and/or breaking some salt bridges. The width of the single polymer (~7nm) is thicker than the polymers seen by TEM (4-6 nm) but SAM domain polymers are known to stretch and compact readily in different crystal forms (Nauli et al., 2007).

Individual ANKS3-SAM polymers pack in the crystal structure as a striking triple helix, but we do not know if this is a biologically relevant structure. Caskin1, which also contained 8 SAMs per helical polymer repeat, adopted a triple helix in the crystal, but it was shown to likely be an artifact of crystallization (Stafford et al., 2011). The N- and C-termini do splay out from the triple helix, so the remaining domains of the full-length protein could be accommodated. Moreover, there is charge complementarity between polymer surfaces. In particular, the polymer surface viewed from the free ML-surface end (Fig. 3-5A right) has negative patches which complement the positive patches seen from the free EH-surface end (Fig. 3-5A left). Unlike the inter-polymer association seen in the Caskin1 structure which was largely mediated by an added His_6-tag, the triple helix of ANKS3-SAM is held together by residues Ala3-Gly7 of chain A (for which there is no equivalent density in chain B) intercalating between SAM polymers of an adjacent triple helix (Fig. 3-8). Nevertheless, the polymers are not packed tightly, as there are ample gaps between individual polymers.

The phenomenon of ANKS3-SAM polymer sheet formation observed by TEM also remains to be tested in the context of the full-length protein. Sheets of triple helices can be constructed from the crystal structure, but the sheets formed in the crystal packing (Additional file 2) would be incompatible with triple helix formation. Although the N- and C- termini extend away from the triple helix polymer axis, additional domains of the full-length protein cannot be obviously accommodated in the sheets. If sheets do form as seen in the EM images, it would require different packing than we see in the crystal structure.

The EH-surface of ANKS3-SAM binds the ML-surface of ANKS6-SAM

To map the binding interface between ANKS3-SAM and ANKS6-SAM, we employed the negGFP binding assay discussed above. As shown in Fig. 3-4B, mutations in residues L52, F53, and K58 on the EH-surface of ANKS3-SAM and mutations in residues E29 and D42 on the ML-surface of ANKS6-SAM abolish the ANKS3-SAM + ANKS6-SAM hetero-interaction in this assay.

Since the EH-surface of ANKS3-SAM is required for both polymerization and binding to ANKS6-SAM, these two events are mutually exclusive. Therefore, ANKS6-SAM binding to ANKS3-SAM should block ANKS3-SAM polymerization. To test this possibility, we mixed ANKS6-SAM with the insoluble ANKS3-SAM polymer described above. As shown in Fig. 3-4C, the addition of ANKS6-SAM does indeed lead to solubilization of the ANKS3-SAM precipitate.

The ability of ANKS6-SAM to solubilize ANKS3-SAM polymers and the slower migration of the hetero-interaction on the native gel compared to ANKS3-SAM suggests that ANKS6-SAM has a higher affinity for ANKS3-SAM than ANKS3-SAM does for itself. To determine the affinity of the ANKS3-SAM/ANKS6-SAM interface we again used SPR by immobilizing ANKS6-SAM and measuring equilibrium binding to an ANKS3-SAM ML-surface mutant, I36E. At 0.15M NaCl, we observed a high binding affinity (K_d = 249 ± 8 nM) (Fig. 3-4E). This affinity is more than an order of magnitude tighter than the binding affinity we measured for the native EH-ML interface formed between ANKS3-SAMs (K_d = 5.8 ± 0.4 μM), indicating that ANKS6-SAM binding could effectively compete with polymerization. Binding is also strongly salt dependent as shown in Fig. 3-4F, suggesting that ionic interactions are important features of the interface.

Structure of the ANKS3-SAM/ANKS6-SAM complex

To learn how ANKS3-SAM and ANKS6-SAM bind each other, we determined a crystal structure of the complex (Fig. 3-6A). To prevent ANKS3-SAM polymerization without destroying ANKS6-SAM binding, we mixed an ML-surface mutant of ANKS3-SAM, I36E, with wild-type ANKS6-SAM. These proteins formed a heterodimer when analyzed by SEC-MALS (Fig. 3-6B) and yielded crystals suitable for structure determination.

The ANKS3-SAM/ANKS6-SAM heterodimer crystallized in space group $C222_1$ with four molecules (2 ANKS3-SAM I36E mutants and 2 ANKS6-SAMs) per asymmetric unit. Both of the ANKS3-SAM chains and both ANKS6-SAM chains in the asymmetric unit had closely similar structures, with RMSDs between backbone atoms of 0.279 Å and 0.169 Å, respectively. Additionally, the overall structure of ANKS3-SAM I36E aligns well with ANKS6-SAM (RMSD across backbone atoms is 0.626 Å) with the only obvious difference being a 1.6 Å outward shift of ANKS6-SAM's helix 3.

Similar to the ML-surface of ANKS3-SAM, the ML-surface of ANKS6-SAM contains a shallow hydrophobic patch (residues V30, A34, L38, L46) flanked by negatively charged residues (E29, D31, E33, D42, E45) (Fig. 6A). Many of these charged residues form salt bridges with ANKS3-SAM (ANKS6 E29 + ANKS3 K22, ANKS6 D31 + ANKS3 K58, ANKS6 D42 + ANKS3 R57). Other residues of ANKS6-SAM (E33, E45) form hydrogen bonds with ANKS3-SAM bridged by water. These results are consistent with our native gel analysis of critical interface residues (Fig 3-4B). Similar to the ANKS3-SAM/ANKS3-SAM interface, the F53E mutation in ANKS3-SAM removes a Phe required for packing against the hydrophobic patch of the ANKS6-SAM ML-surface and the L52E mutation removes van der Waals interactions and hydrophobic contacts that otherwise stabilize the interaction interface. Removal of salt bridges by the mutations K58E in ANKS3-SAM, E29K in ANKS6-SAM, and D42K in ANKS6-SAM also

break the interface. The ANKS6-SAM mutation R54W (R823W in the full length protein) breaks the interaction with ANKS3-SAM as well and will be discussed below.

Comparing an ANKS3-SAM/ANKS6-SAM heterodimer with an ANKS3-SAM homodimer by aligning the ANKS3-SAMs, we see that ANKS6-SAM is tilted closer towards ANKS3-SAM (Fig. 3-6C). This ANKS3-SAM/ANKS6-SAM interface buries on average 454 $Å^2$ of surface area, which is approximately 100 $Å^2$ more than was buried at the same interface in the ANKS3-SAM polymer. An isoleucine (I36) in ANKS3-SAM is switched for an alanine (A34) in ANKS6-SAM at the ML-surface. The decreased steric bulk of Ala compared to Ile allows helix 3 of ANKS6-SAM's ML-surface to approach the EH-surface of ANKS3-SAM more closely, resulting in the burial of additional surface area and the formation of more salt bridges and hydrogen bonds, which may explain the higher affinity of the ANKS3-SAM/ANKS6-SAM interface. The ANKS6-SAM structure also shows that the EH-surface is lacking a large hydrophobic residue flanked by positively charged residues, thereby showing that ANKS6-SAM cannot polymerize because its EH-surface is incompatible with binding its ML-surface.

The R823W mutation of Pkd/Mhm(cy/+) rats perturbs the structure of ANKS6-SAM

Using our native gel binding assay, we discovered that the Cy mutation (R54W in our numbering of the construct, R823W in the full-length protein) in ANKS6-SAM destroys the interaction with ANKS3-SAM (Fig. 3-4B). This interaction could not be restored by titrating increasing amounts of ANKS6-SAM R823W with ANKS3-SAM (Fig. 3-7A). To understand how the Cy mutation might be responsible for disease in Pkd/Mhm(cy/+) rats, we examined its position in our structure of ANKS6-SAM. R823 (R54 in the crystal structure numbering, see Fig. 3-1) forms salt bridges with D51 of helix 5 and D40 of helix 4, and forms hydrogen bonds with the backbone carbonyls of I48 and K49 which both lie on a loop between helices 4 and 5 (Fig.

3-7B). Through these interactions, R54 is involved in stapling helices 4 and 5 and in maintaining the overall fold in this segment of the domain.

Since the R823W mutation does not lie on either the ML- or EH-surface of ANKS6-SAM, we reasoned that its ability to abolish binding to ANKS3-SAM must be due to a long range structural alteration. Indeed the far-UV CD spectrum of ANKS6-SAM R823W shows an approximately 10% loss of helical structure with a parallel increase in random coil compared to the wild-type protein (Fig. 3-7C). Moreover, the stability of the protein is dramatically reduced. As shown in Fig. 3-7D, wild-type ANKS6-SAM displayed a broad, reversible unfolding curve with a T_m of approximately 48^{o}C. In contrast, the ANKS6-SAM R823W mutant strikingly exhibited a complete loss of cooperative unfolding. Therefore, the R823W mutation appears to dramatically destabilize the structure. The mutant protein also migrates slightly faster and with a higher molecular weight as assessed by SEC-MALS (Fig. 3-9), consistent with generalized unfolding. Indeed, modeling a tryptophan at position R54 in our structure is impossible as this would introduce severe steric clashes. Thus, coupling the removal of Arg with the insertion of a Trp disturbs the overall tertiary structure of the protein and prevents the ML-surface of the ANKS6-SAM domain from adopting a fold complementary to binding the ANKS3-SAM domain.

Discussion

The PKD/Mhm(cy/+) rat, in which an R823W mutation in the SAM domain of ANKS6 is causal of disease, has been used extensively as an animal model for the study of human ADPKD. However, the underlying mechanism whereby this single mutation leads to improper kidney development and loss of renal function has been unclear. For the first time, we have identified the SAM domain of ANKS3 as a direct binding partner of the ANKS6-SAM domain and shown that this interaction is lost in the ANKS6-SAM R823W mutant. The physiological

relevance of this interaction is supported by the independent identification of the potential ANKS3/ANKS6 interaction in a proteomics screen (Hoff et al., 2013).

By solving the crystal structure of ANKS3-SAM, we have observed that ANKS3-SAM forms polymers via sequential interactions of ML- and EH-surfaces, much like other SAM domains. The polymer formation by ANKS3-SAM suggests that ANKS3 may be capable of scaffolding a larger protein complex. We also observed higher order structure of the polymer, including a triple helix in the crystal and sheet structures in TEM images, but the biological relevance of these polymer associations is unclear. SAM domains have been previously found to organize as sheets of polymers that require divalent metal cation binding to form (Gundelfinger et al., 2006; Knight et al., 2010), but we did not observe any obvious metal binding sites in the ANKS3-SAM crystal structure.

By solving a crystal structure of the ANKS3-SAM/ANKS6-SAM heterodimer, we observed that the EH-surface of ANKS3-SAM binds the ML-surface of ANKS6-SAM. We have also provided a molecular explanation for the defect in the R823W mutation: it disrupts the tertiary structure of the ANKS6-SAM domain and prevents the ML-surface of ANKS6-SAM from adopting a conformation capable of binding ANKS3-SAM. Because the EH-surface of ANKS3-SAM is mutually exclusive for binding either the ML-surface of ANKS3-SAM or the ML-surface of ANKS6-SAM, we expect that ANKS6 may act as a polymer capper (Qiao et al., 2004, 2005). It may shorten the ANKS3-SAM polymers or it may simply recruit ANKS3 polymers to a larger complex of proteins.

ANKS6 has been linked to Bicaudal-C1 (BICC1), a protein that when mutated is responsible for disease in the *jcpk* mouse model of ADPKD and the *bpk* mouse model of ARPKD (autosomal recessive polycystic kidney disease) (Bryda et al., 2003; Stagner et al., 2009). ANKS6 and BICC1 have been shown to co-immunoprecipitate and co-localize in inner

medullary collecting duct (IMCD) cells. An unidentified protein complex has been suggested to link ANKS6 to BICC1, by binding both the ANKS6-SAM domain and a strand of RNA to which the KH domains of BICC1 are bound (Stagner et al., 2009). It is reasonable to suspect that ANKS3 may also be part of a complex with ANKS6 and BICC1.

Transgenic overexpression of ANKS6$^{(p.R823W)}$ can generate a disease phenotype, indicating that ANKS6-SAM R823W acts as a dominant negative (Neudecker et al., 2010). A dominant negative effect would be expected to arise from a gain of function, yet the mutation clearly destabilizes the structure of the SAM domain. A reasonable possibility is that other domains on ANKS6 recruit it to a complex (Bryda et al., 2003; Stagner et al., 2009), but the defective SAM domain fails to appropriately recruit other proteins, such as ANKS3, leading to defective function. Altered RXR-mediated signaling pathways are observed in PKD/Mhm(cy/+) rats (Kugita et al., 2010) and ANKS6 has recently been placed as a central node in a distinct NPHP-associated signaling network (Hoff et al., 2013).

The recent finding of mutations in ANKS6 in individuals presenting with polycystic kidney disease (Hoff et al., 2013) has catapulted this protein from relevance in an animal model to having direct implications for human renal development and function. Furthermore, the fact that one patient with a truncation in ANKS6 at the N-terminal end of the SAM domain also presented with aortic stenosis, causing obstructive cardiomyopathy, implicates ANKS6 in cardiac development and function (Hoff et al., 2013). Whether ANKS3 may also have a role in cardiac development, via its link with ANKS6, remains to be investigated. So far, ANKS3 is the only direct binding partner of ANKS6 identified that is affected by the R823W mutation. Our work provides a structural and biochemical basis for future work on ANKS3 and its interaction with ANKS6.

Conclusions

We have identified and characterized the novel interaction between the SAM domain of ANKS3 and the SAM domain of ANKS6. ANKS3-SAM was found to be capable of polymerization, although polymer formation and binding to ANKS6 appear mutually exclusive, suggesting ANKS6 may act as a polymer capper. The R823W mutation associated with cystic kidney disease causes a destabilization of the ANKS6 SAM domain which disrupts binding to ANKS3. By structurally and biochemically characterizing this new interaction we provide a foundation to support continued research into how the ANKS3/ANKS6 interaction affects kidney function.

Methods

Cloning and Mutagenesis

negGFP-human-SAM fusions and DNA used in cloning were as described previously (Knight et al., 2011). All human ANKS3-SAM constructs contained residues 421-490 [UniProt:Q6ZW76] and all human ANKS6-SAM constructs contained residues 771-840 [UniProt:Q68DC2]. A His_6-tagged construct of ANKS3-SAM was generated by cloning the human ANKS3-SAM sequence into a pET28 vector (Novagen). A His_6-tagged construct of ANKS6-SAM was generated by cloning the ANKS6-SAM sequence into a pBAD-HisA vector (Invitrogen). In both constructs, the residues MARHHHHHHSSG were added to the N-terminus of each SAM to incorporate a His_6-tag. Hexahistidine small ubiquitin-like modifier (SUMO) tagged constructs were generated by cloning the ANKS3-SAM and ANKS6-SAM sequences into a pHis-SUMO vector (Senturia et al., 2010). Site-directed mutagenesis was performed using the Quickchange method (Agilent). All plasmid sequences were verified by DNA sequencing (Genewiz).

negGFP Native Gel Binding Assay

negGFP-human-SAM fusions transformed into ARI814 cells (Schatz et al., 1996) were expressed as described previously (Knight et al., 2011). Harvested cells were resuspended in 0.5mL of 20mM Tris pH 7.5, 0.3M NaCl, 2mM TCEP, supplemented with DNaseI (40µg/mL), lysozyme (10mg/mL), PMSF (1mM), $MgCl_2$ (10mM), and half a tablet of cOmplete mini, EDTA-free protease inhibitor (Roche) and lysed by three freeze-thaw cycles and one 10-sec. round of sonication. Lysate was centrifuged at 16,060*g* for 20 minutes at 4^{o}C and the pellet discarded. The expression levels of the negGFP-human-SAM fusions was determined by fluorescence intensity as described previously (Knight et al., 2011). To identify new ANKS3-SAM interactions, lysates were mixed to maintain equal amounts of negGFP-SAM fusion (dilutions were made

using 20mM Tris pH 7.5, 0.3M NaCl, 1mM TCEP). Mixes were allowed to equilibrate at 4°C for 3 hours, at which point 7.5µL of 4X RunBlue Native Sample Buffer (Expedeon) was added and samples were applied to a 20% RunBlue 12-well Native gel in RunBlue Native Run Buffer (Expedeon) and developed at 90V for 24 hours at 4°C. Gels were visualized on a Bio-Rad Molecular Imager FX Pro-Plus using an excitation wavelength of 488nm and an emission wavelength of 510nm. For characterization of ANKS3-SAM and ANKS6-SAM mutants, lysates were prepared as above and adjusted by fluorescence for equal gel loading. Lysates were mixed in varying ratios based on fluorescence where appropriate, and allowed to equilibrate at 4°C for 4 hours before gel loading. Gels were run at 90V for 15 hours at 4°C and visualized as above.

Protein Expression and Purification

negGFP-ANKS3-SAM Fusion

negGFP-ANKS3-SAM fusion transformed into ARI814 cells (Schatz et al., 1996) was expressed as described previously (Knight et al., 2011). Harvested cells were lysed as described previously, except that 5 cOmplete mini, EDTA-free protease inhibitor tablets (Roche) were included in the lysis buffer and five 1-min. rounds of sonication were used (Knight et al., 2011). Following centrifugation at 13,200*g* for 20 minutes, lysate supernatant was supplemented with 10mM imidazole and rocked with 2mL of Ni-NTA Superflow agarose (Qiagen) for 1 hour at 4°C. The resin was washed with lysis buffer (20mM Tris pH 7.5, 1M NaCl, 1mM TCEP) containing 20mM imidazole, and eluted with lysis buffer containing 75mM imidazole. Eluted protein was dialyzed into 20mM Tris pH 7.5, 0.15M NaCl, 1mM TCEP and loaded onto a 5mL HiTrap Q HP column (GE) 4°C. Protein was eluted using a shallow gradient of NaCl (0.25-0.5M) in 20mM Tris pH 7.5, 1mM TCEP. Purified protein was dialyzed into 20mM Tris pH 7.5, 0.3M NaCl, 1mM TCEP and concentrated using an Amicon Ultra (10kDa MWC0) centrifugal filter unit (Millipore) to a final concentration of 4.3mg/mL.

His_6-tagged Constructs

His_6-tagged ANKS6-SAM was transformed into ARI814 cells and expressed as above in 6L of LB media supplemented with 100µg/mL ampicillin. Harvested cells were resuspended to 80mL in lysis buffer (20mM Tris pH 8, 0.5M NaCl) containing PMSF (1mM), DNaseI (20µg/mL), and $MgCl_2$ (10mM) and lysed on an EmulsiFlex-C3 (Avestin) at 18,000 psi. Lysate was centrifuged at 13,200*g* for 20 minutes at 4°C. Supernatant, with imidazole added to 10mM, was bound at 4°C to a 5mL HiTrap IMAC HP column (GE) charged with $NiCl_2$. The column was washed with lysis buffer containing 20mM imidazole and eluted using a shallow gradient of imidazole (20-260mM). Eluted protein was diluted to approximately 1mg/mL and dialyzed into 20mM Tris pH 8.0, 50mM NaCl. Protein was further purified by loading onto a 1mL HiTrap Q column (GE) equilibrated in 20mM Tris pH 8 at 4°C. Protein was eluted using a shallow gradient of NaCl (0-1.0M) in 20mM Tris pH 8.0. Fractions containing pure protein were dialyzed against 20mM Tris pH 8.0, 0.3M NaCl.

His_6-tagged ANKS3-SAM was transformed into Rosetta(DE3) cells (Novagen) and 4.5L of culture was grown in LB media supplemented with 30µg/mL kanamycin and 34µg/mL chloramphenicol. Cells were grown at 37°C with shaking until cell density reached an OD_{600} of 0.7, at which point cells were chilled to 18°C, induced with 1mM isopropyl β-D-galactopyranoside (IPTG), and grown an additional 16 hours at 18°C. Harvested cells were resuspended to 80mL in lysis buffer (20mM Tris pH 7.0, 1M NaCl, 2mM BME) containing PMSF (1mM), DNaseI (20µg/mL), $MgCl_2$ (5mM) and processed as above. Supernatant, with imidazole added to 10mM, was rocked with 3mL of Ni-NTA Superflow agarose (Qiagen) for 1 hour at 4°C. Resin was washed with 50mL of lysis buffer containing 10mM imidazole, 200mL of lysis buffer containing 20mM imidazole, and eluted using lysis buffer with 200mM imidazole. Elution

fractions were stored at 4°C and developed fluffy precipitate overnight and for several days thereafter.

SUMO-Fusions

pHis-SUMO constructs (ANKS3-SAM mutants L52A, F53E, I36E, and ANKS6-SAM wt and R823W) were transformed into Rosetta(DE3) pLysS cells (Novagen) and 4L of culture was grown and expressed as described above. Harvested cells were resuspended in lysis buffer (50mM $NaHPO_4$ pH 8.0, 0.5M NaCl, 2mM BME) containing PMSF (1mM), DNaseI (20µg/mL), and $MgCl_2$ (5mM) and lysed on an EmulsiFlex-C3 (Avestin) at 18,000 psi followed by centrifugation at 13,200*g* for 20 minutes at 4°C. Supernatant, with imidazole added to 10mM, was rocked with Ni-NTA Superflow agarose (Qiagen) for 1 hour at 4°C. Resin was washed with lysis buffer containing 20mM imidazole and eluted with lysis buffer containing 250mM imidazole. The different proteins were dialyzed into buffers of various ionic strengths as required to maintain solubility: (1) ANKS3-SAM F53E was dialyzed into 20mM Tris pH 8.0, 0.3M NaCl, 2mM BME; (2) ANKS6-SAM wt and R823W were diluted to 5-7mg/mL and dialyzed into 20mM Tris pH 8.0, 0.3M NaCl; (3) ANKS3-SAM L52A and I36E were diluted to 4-5mg/mL and 50mM EDTA was added to prevent precipitation seemingly induced by leached Ni^{2+}, followed by extensive dialysis into 20mM Tris pH 8.0, 0.5M NaCl, 2mM BME. To remove the His_6-SUMO tag, all constructs were digested with the His_6-tagged catalytic domain of SUMO protease 1 (ULP1) at a 50:1 protein:protease molar ratio for 16 hours at 4°C (Malakhov et al., 2004). The cleaved His_6-SUMO tag and ULP1 protease were removed by two rounds of subtractive Ni-NTA. Here the protein was either in the flow-through (ANKS3-SAM F53E, ANKS6-SAM wt and R823W) or remained weakly bound to the resin and was eluted with 20mM $NaHPO_4$ pH 8.0, 0.5M NaCl, 2mM BME, 10mM imidazole (ANKS3-SAM L52A and I36E). Final purification steps are specific to each construct as described below.

ANKS3-SAM F53E

ANKS3-SAM F53E from the subtractive Ni-NTA flow-through was dialyzed into 20mM Tris pH 7.5, 30mM NaCl, 1mM DTT and bound to a 5mL HiTrap Q HP (GE) column equilibrated in 20mM Tris pH 7.5, 2mM DTT. Protein was eluted using a shallow gradient of NaCl (0.1-0.5M). Fractions containing pure ANKS3-SAM F53E were dialyzed into 20mM Tris pH 8.0, 50mM NaCl, 2mM BME and concentrated in an Amicon Ultra (3kDa MWCO) centrifugal filter unit (Millipore) to 27mg/mL.

ANKS6-SAM wt and ANKS6-SAM R823W

ANKS6-SAM wt and R823W from the subtractive Ni-NTA flow-through were each dialyzed into 20mM Tris pH 8.0, 50mM NaCl and bound to a 5mL HiTrap Q HP column (GE) equilibrated in 20mM Tris pH 8.0. Protein was eluted with a shallow gradient of NaCl (0.1-1M). Pure fractions were pooled and dialyzed against 20mM Tris pH 8, 0.15M NaCl. Proteins were concentrated in Amicon Ultra (3kDa MWCO) centrifugal filter units to 12-14 mg/mL.

ANKS3-SAM L52A

ANKS3-SAM L52A exhibited a weak affinity for IMAC resin and was dialyzed into 20mM Tris pH 8.0, 0.5M NaCl, 2mM BME and bound to a 5mL HiTrap IMAC HP column (GE) charged with $NiSO_4$. Protein was eluted using a shallow gradient of imidazole (0-15mM). Fractions containing pure ANKS3-SAM L52A were dialyzed into 20mM Tris pH 8.5, 0.15M NaCl, 1mM DTT and concentrated using an Amicon Ultra (3kDa MWCO) centrifugal filter unit (Millipore) to approx. 5 mg/mL. Surprisingly, at this concentration the protein spontaneously formed crystalline needles with low resolution diffraction. Needles were resolubilized by the addition of NaCl to a final concentration of 0.72M.

ANKS3-SAM I36E

ANKS3-SAM I36E eluted from the subtractive Ni-NTA column in both the flow-through and the 10mM imidazole wash. The protein was dialyzed into 20mM Tris pH 8.0, 0.5M NaCl, 2mM BME and bound to a 5mL HiTrap IMAC HP column (GE) charged with $NiSO_4$. Protein was eluted using a shallow gradient of imidazole (0-15mM). Pure fractions were pooled and dialyzed into 20mM Tris pH 8.0, 0.75M NaCl, 1mM DTT. Protein was concentrated in an Amicon Ultra (3kDa MWCO) centrifugal filter unit (Millipore) to approx. 2.5mg/mL.

Resolubilization of ANKS3-SAM precipitate by ANKS6-SAM

His_6-tagged ANKS3-SAM precipitate was formed by dialyzing protein at 3.2mg/mL into low salt buffer (20mM Tris pH 7, 0.25M NaCl, 1mM DTT). His_6-tagged ANKS6-SAM at 3.3mg/mL in 20mM Tris pH 8.0, 0.3M NaCl buffer or buffer alone (20mM Tris pH 8.0, 0.3M NaCl) was added in varying ratios of 2:1, 1:1, and 1:2 (ANKS3-SAM vs. ANKS6-SAM ratio) to the ANKS3-SAM precipitate, keeping the total volume and total concentration of protein constant. Mixes were allowed to incubate on ice for 3 hours.

Crystallization and Structure Determination

ANKS3-SAM L52A

Initial screening for crystallization conditions was performed at the High Throughput Macromolecular Crystallization Facility at UCLA. The crystallization trials were carried out in hanging drops by vapor diffusion, using commercially available screens and a Mosquito TTP crystallization Robot.

Native crystals were grown by hanging drop vapor diffusion by mixing 1µL of ANKS3-SAM L52A at 4.3mg/mL with 2µL of well solution (0.1M Na/KPO_4 pH 6.6, 0.3M NaCl, 15% PEG-8000). Rod-shaped crystals grew at 4^{o}C over a two-week period and were cryoprotected using well solution supplemented with 25% glycerol. For phasing, crystals were briefly soaked and

cryoprotected (5 sec) in a solution of 0.5M KI prepared in well solution supplemented with 27% glycerol. A data set was collected for a single ANKS3-SAM L52A iodide derivative crystal cryo-cooled to 100K at UCLA using the in-house RIGAKU FRE+ generator and HTC image plate detector at a wavelength of 1.5418nm. Data reduction and scaling were performed using XDS/XSCALE (Kabsch, 2010). Phasing was accomplished by single anomalous dispersion (SAD) using the HKL2MAP interface and SHELX program (Pape and Schneider, 2004; Sheldrick, 2010). Ten iodide atoms were detected and used for phasing. Density modification and model building were accomplished using DM and BUCCANEER, respectively, in the CCP4 suite (Winn et al., 2011). The structure was briefly refined in PHENIX (Adams et al., 2010) with inspection and model rebuilding in COOT (Emsley et al., 2010).

A high resolution data set was collected on a single native crystal cryo-cooled to 100K at the Advanced Photon Source (Argonne National Laboratory), APS-NECAT beamline 24-ID-C on a DECTRIS-PILATUS 6M detector. Single crystals were mounted with CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo). A data set containing 418 1.0° oscillation frames was collected from a single large crystal at a wavelength of 0.9795 Å. This dataset was indexed and merged for scaling using XDS (Kabsch, 2010). The model was refined using PHENIX with TLS parameterization including individual sites and individual atomic displacement parameters (Adams et al., 2010). Data collection and refinement statistics are reported in Table 1. The coordinates have been deposited in the PDB with accession code 4NJ8.

ANKS3-SAM/ANKS6-SAM Heterodimer

Crystals were grown in 2µL hanging drops by mixing equal parts ANKS3-SAM/ANKS6-SAM heterodimer at 30mg/mL (in 20mM Tris pH 8.0, 0.15M NaCl, 2mM BME) with 0.1M Tris pH 8.0, 0.25M $MgCl_2$, 31.5% PEG-4000 reservoir solution. Drops were allowed to equilibrate overnight and were streak seeded the following day using a cat whisker to transfer nuclei from crystals that spontaneously formed overnight in a condition with a higher percentage of PEG-

4000. Seeded crystals grew within one day and reached full size within 4 days. An x-ray diffraction data set containing 300 1^{o} oscillation frames was collected on a single crystal cryo-cooled to 100K at the Advanced Photon Source (Argonne National Laboratory), APS-NECAT beamline 24-ID-C, at a wavelength of 0.9793 Å and processed using XDS (Kabsch, 2010). Molecular replacement was performed using PHASER with ANKS3-L52A as a search model (McCoy et al., 2007) and refined as described above. Data collection and refinement statistics are reported in Table 1. The coordinates have been deposited in the PDB with accession code 4NL9.

Structure analysis

Final structure models were validated with the following structure validation tools: PROCHECK (Laskowski et al., 1993), ERRAT (Colovos and Yeates, 1993), and VERIFY3D (Bowie, et al., 1991). Graphics were prepared using PyMOL (Schrödinger, LLC). Surface electrostatics were prepared using the Adaptive Poisson-Boltzmann Solver (APBS)(Baker et al., 2001) plugin in Pymol and all surfaces were contoured at ±1 kT/e. Analysis of protein-protein binding interfaces was done using the PISA server (Krissinel and Henrick, 2007).

Circular Dichroism

Spectra were collected for protein samples at 0.2 mg/ml in 10mM Tris pH 8.0, 75mM NaCl at 25^{o}C on a JASCO J-715 circular dichroism spectrophotometer equipped with a Peltier temperature control. Spectra were analysed for secondary structure content using the Selcon, Neural Network, and Contin algorithms available in SoftSec (Softwood Inc.). Thermal melts were performed by monitoring the change in CD signal at 222nm across a temperature range of 25-80^{o}C, with ramping of 1^{o}C per minute.

Surface Plasmon Resonance

Experiments were performed at 21°C in 20mM HEPES pH 7.5, 1mM DTT, 0.04% IGEPAL CA-630, and NaCl ranging from 0.1-0.3M using a Biacore T100 (GE). To determine the binding affinity of the ANKS3-SAM/ANKS6-SAM interface, ANKS6-SAM wt was immobilized on a Biacore CM5 chip (GE) via EDC/NHS crosslinking. To determine the binding affinity of the ANKS3-SAM/ANKS3-SAM interface, ANKS3-SAM F53E was immobilized on a Biacore CM5 chip (GE) via EDC/NHS crosslinking. In both cases, ANKS3-SAM I36E at varying concentrations was passed over the chip and equilibrium binding levels were measured. All data points were taken in triplicate and binding data was fit to a 1:1 steady-state model using Biacore T100 Evaluation software. In total, 630.9 response units (RUs) were immobilized on the ANKS6-SAM wt chip surface and 2188.7 RUs were immobilized on the ANKS3-SAM F53E chip surface. At 0.15M NaCl, calculated R_{max} values were 262.7 for the ANKS6-SAM wt chip and 682.3 for the ANKS3-SAM F53E chip, indicating that 30-40% of the surface molecules were active. To determine how ionic strength impacted binding affinity, the ANKS3-SAM/ANKS6-SAM K_d was determined at 0.15, 0.2, 0.25, and 0.3M NaCl and the ANKS3-SAM/ANKS3-SAM K_d was determined at 0.1, 0.15, 0.2, and 0.25M NaCl. Prior to the experiments, ANKS3-SAM I36E at 2.4mg/mL was dialyzed into 20mM HEPES pH 7.5, 0.75M NaCl, 1mM DTT, with the high salt required to maintain stability. Immediately prior to each Biacore run, dilutions of ANKS3-SAM I36E were prepared to match the running buffer and protein concentration was determined by Bradford assay, using known concentrations of ANKS3-SAM I36E for a standard curve.

Electron Microscopy

Samples were applied to carbon/formvar coated copper grids (Ted Pella, catalog number 01754-F) made hydrophilic by glow discharge immediately before use and allowed to bind for several minutes. Grids were rinsed with distilled water and negatively stained with 1%

uranyl acetate. Samples were analysed on a CM120 transmission electron microscope (TEM) operating at 120kV. Images were recorded using a TEITZ F224HD CCD camera and processed using ImageJ (NIH).

SEC-MALS

100µL of protein (either ANKS3-SAM/ANKS6-SAM heterodimer at 15mg/mL, ANKS6-SAM wt at 10mg/mL, or ANKS6-SAM R823W at 10mg/mL) was analysed by SEC-MALS. Protein was loaded onto a WTC-030S5 analytical size-exclusion column (Wyatt Technology Co.) equilibrated in 20mM Tris pH 8.0, 0.15M NaCl, (+ 2mM BME for the ANKS3-SAM/ANKS6-SAM heterodimer) using an AKTA purifier (GE) and run at 0.5mL/min on a miniDAWN TREOS (Wyatt Technology Co.). Eluted protein peaks were analysed for calculated molecular weight and monodispersity using ASTRA software (Wyatt Technology Co.)

Figures and Tables

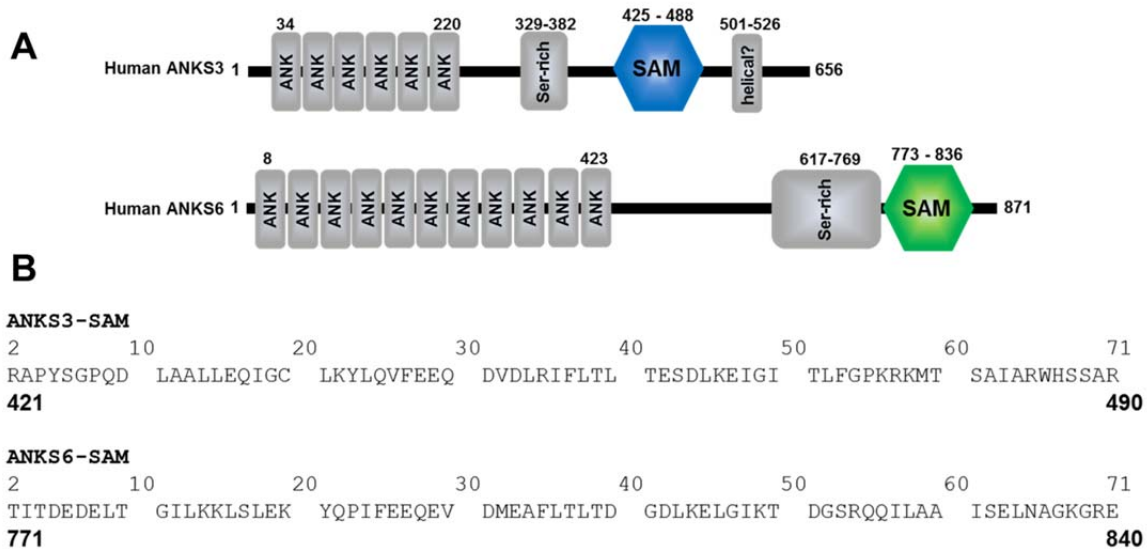


Figure 3-1. The ANKS3 and ANKS6 proteins

A) The domain structure of human ANKS3 and human ANKS6. ANK = ankyrin repeat, SAM = sterile alpha motif, Ser-rich = serine rich region, helical = potential coiled coil domain. B) The sequences of the human ANKS3-SAM [UniProt:Q6ZW76] and ANKS6-SAM domains [UniProt:Q68DC2] used in this work. The top numbering is the numbering used in the crystal structures. Each sequence contains an additional Ser at the N-terminus that is not shown. The numbers in bold below each sequence correspond to the numbering in the full length proteins.

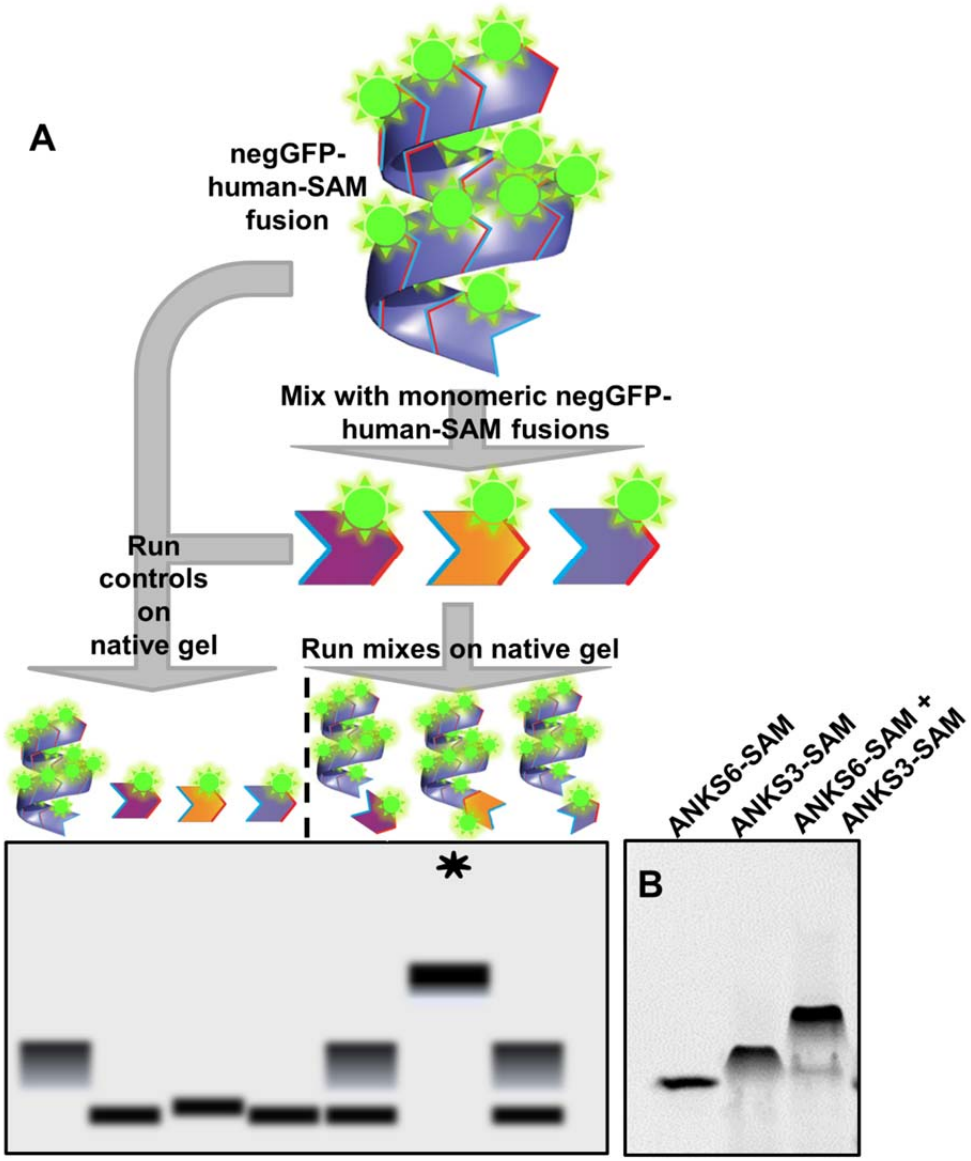


Figure 3-2. ANKS3-SAM binds to ANKS6-SAM

A) Schematic of the negGFP native gel binding assay. Lysate containing a negGFP-human-SAM fusion protein is mixed with lysate containing a negGFP fusion of a different human SAM domain. Each SAM domain contains an ML-surface (red) and an EH-surface (blue). Mixes and individual SAM controls are run on native gels and visualized by fluorescence. Novel hetero-SAM interactions appear as a new upshifted band (asterisk) in the gel schematic. B) negGFP fusions of ANKS6-SAM, ANKS3-SAM, and a 1:1 mix of ANKS6-SAM + ANKS3-SAM assayed by native gel electrophoresis. ANKS6-SAM appears monomeric, ANKS3-SAM appears weakly polymeric, and the appearance of a new distinct band upon mixing the proteins indicates an interaction.

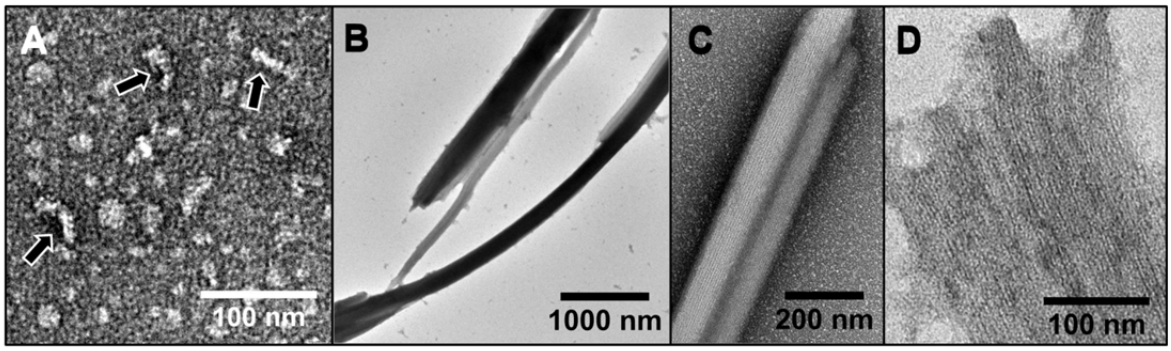


Figure 3-3. ANKS3-SAM assembles into polymers

A) Negative stain TEM of the negGFP-ANKS3-SAM fusion protein reveals short helical polymers (arrows). B-D) His_6-tagged ANKS3-SAM (negGFP removed) precipitates as large fibrous sheets of polymers.

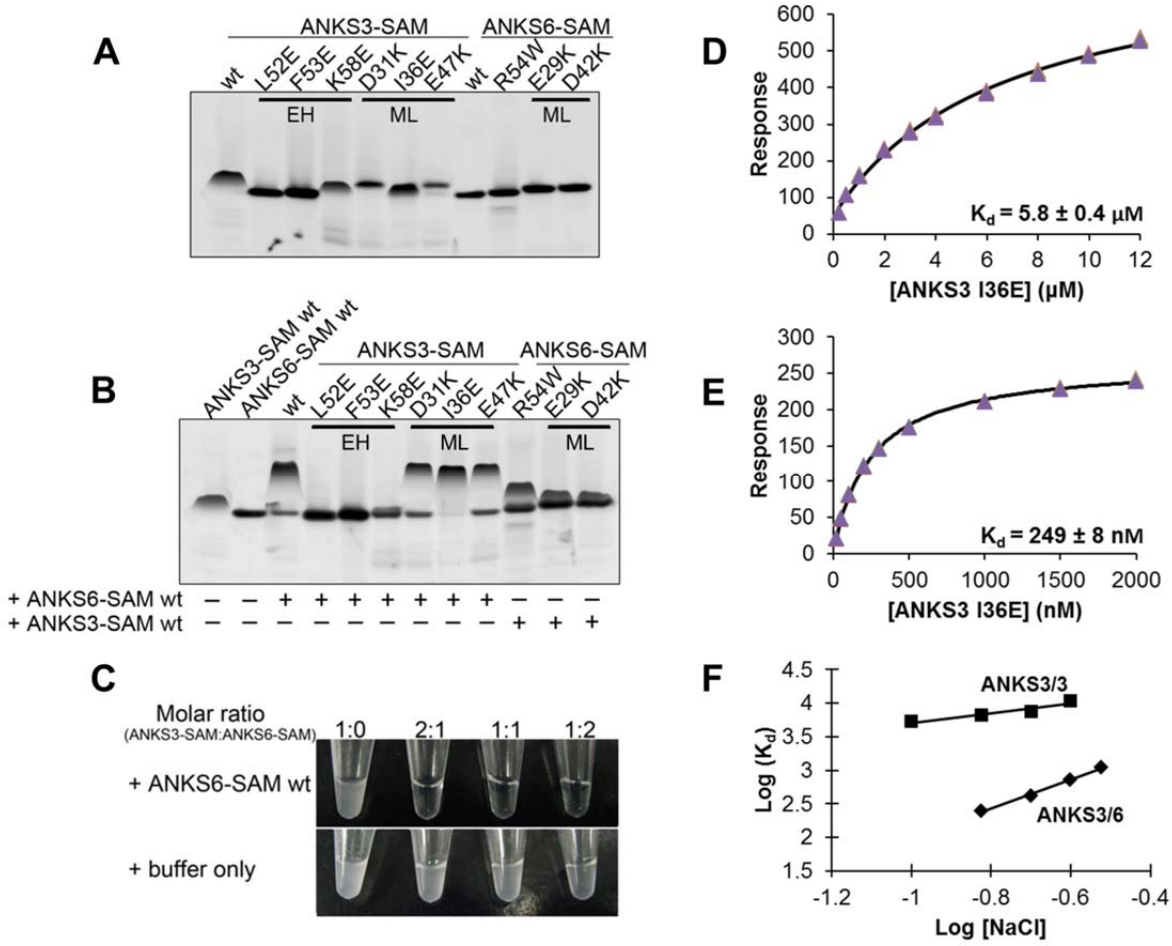


Figure 3-4. Characterizing the ANKS3-SAM polymer interface and the ANKS3-SAM/ANKS6-SAM interaction interface

A) The negGFP native gel screen identifies point mutations in both the EH- and ML-surfaces of ANKS3-SAM that result in a loss of polymeric character. Point mutations in ANKS6-SAM do not impact the native gel migration or monomeric character of this SAM domain. B) Mixing negGFP-ANKS3-SAM and negGFP-ANKS6-SAM mutants in 1:1 molar ratios reveals that the EH-surface of ANKS3-SAM binds the ML-surface of ANKS6-SAM. Point mutations on the EH-surface of ANKS3-SAM inhibit interaction with ANKS6-SAM whereas point mutations on the ML-surface of ANKS3-SAM do not affect the hetero-interaction. Point mutations on the ML-surface of ANKS6-SAM and the Cy mutation (R54W according to our numbering) inhibit interaction with ANKS3-SAM. C) ANKS3-SAM precipitate is resolubilized by the addition of ANKS6-SAM but not by the addition of buffer alone. D) The binding affinity of the native ANKS3-SAM polymer interface is measured by SPR (K_d = 5.8 ± 0.4 μM). Equilibrium binding measurements were performed in triplicate and fit to a 1:1 steady-state model. This calculated K_d is an approximation since tested analyte concentrations were insufficient to reach saturation. The error bars are smaller than the data points. E) The binding affinity of the native ANKS3-SAM/ANKS6-SAM interface is measured by SPR (K_d = 249 ± 8 nM). Equilibrium binding measurements were performed in triplicate and fit to a 1:1 steady-state model. The error bars are smaller than the data points. F)

The binding between both ANKS3-SAM/ANKS3-SAM and ANKS3-SAM/ANKS6-SAM exhibits a salt dependency. The K_d of each interaction was determined by SPR at four different ionic strengths. The slope of the linear fit (~2 for ANKS3-SAM/ANKS6-SAM, ~1 for ANKS3-SAM/ANKS3-SAM) indicates that each interface is salt-dependent and employs ionic interactions.

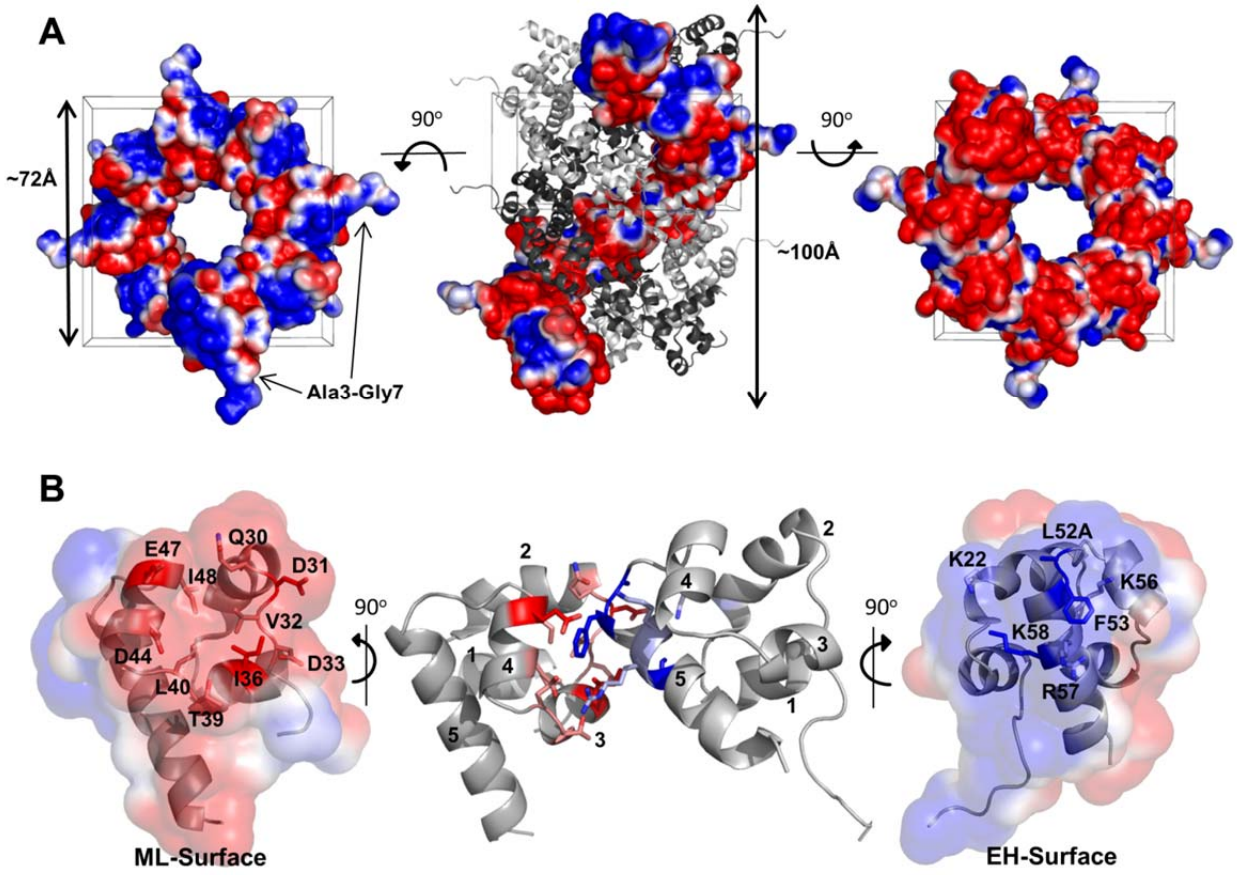


Figure 3-5. Structure of the ANKS3-SAM L52A mutant

A) ANKS3-SAMs pack as a triple helix approximately 72Å in diameter and 100Å long per helical repeat. In the side view of the triple helix, two separate polymer chains are shown as cartoons (grey and black) and the third is shown as a space-filled model colored by surface electrostatics calculated using APBS in Pymol and contoured at ±1 kT/e; red is negatively charged and blue is positively charged. Looking down the polymer axis (left and right) and viewing only the surface electrostatics of a single polymer for clarity reveals the charge complementarity of the polymer surface. B) A single ANKS3-SAM homodimer is shown. Helices have been numbered 1-5. Side chains of residues found critical for the EH-and ML-surface in the negGFP binding assay are shown colored in blue and red, respectively. Both the EH- and ML-surfaces are shown colored by surface electrostatics calculated using APBS and contoured at ±1 kT/e, revealing the charge complementarity of the binding interface.

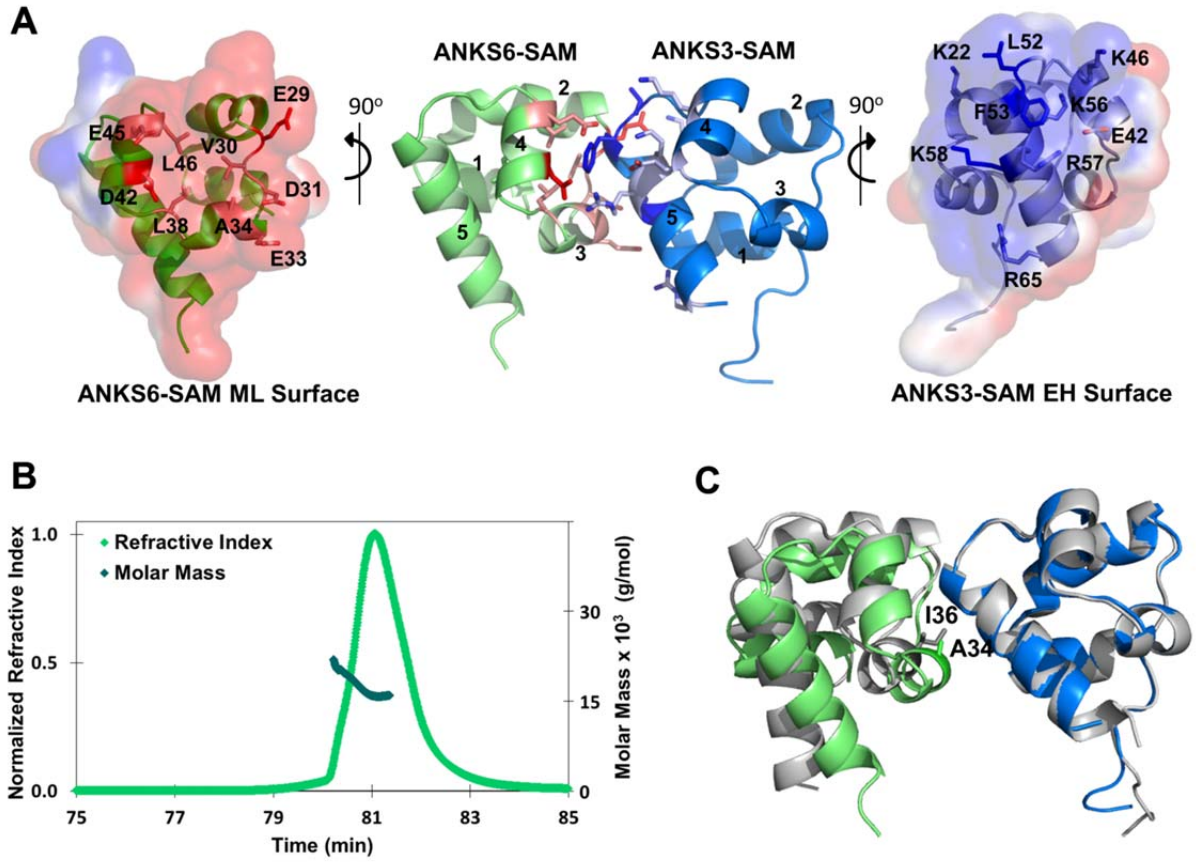


Figure 3-6. Structure of the ANKS3-SAM/ANKS6-SAM heterodimer

A) The ANKS3-SAM EH-surface binds the ANKS6-SAM ML-surface. Helices of the individual SAM domains are numbered 1-5. Side chains of residues which form the ML- and EH-surface are shown and colored red and blue, respectively. Each surface is also colored by surface electrostatics calculated using APBS and contoured at ±1 kT/e, revealing the charge complementarity of the binding interface. B) SEC-MALS analysis of a 1:1 molar ratio mix of ANKS3-SAM I36E + ANKS6-SAM wt produces a single monodisperse peak with a calculated molecular weight of 16.8 kDa, which corresponds to a homogenous population of heterodimer. C) Alignment of an ANKS3-SAM homodimer (grey) with the ANKS3-SAM/ANKS6-SAM heterodimer (blue-green), formed by aligning the backbone atoms of the common ANKS3-SAMs. I36 in ANKS3-SAM is changed to A34 in ANKS6-SAM, which allows ANKS6-SAM to tilt closer to ANKS3-SAM and form more interactions.

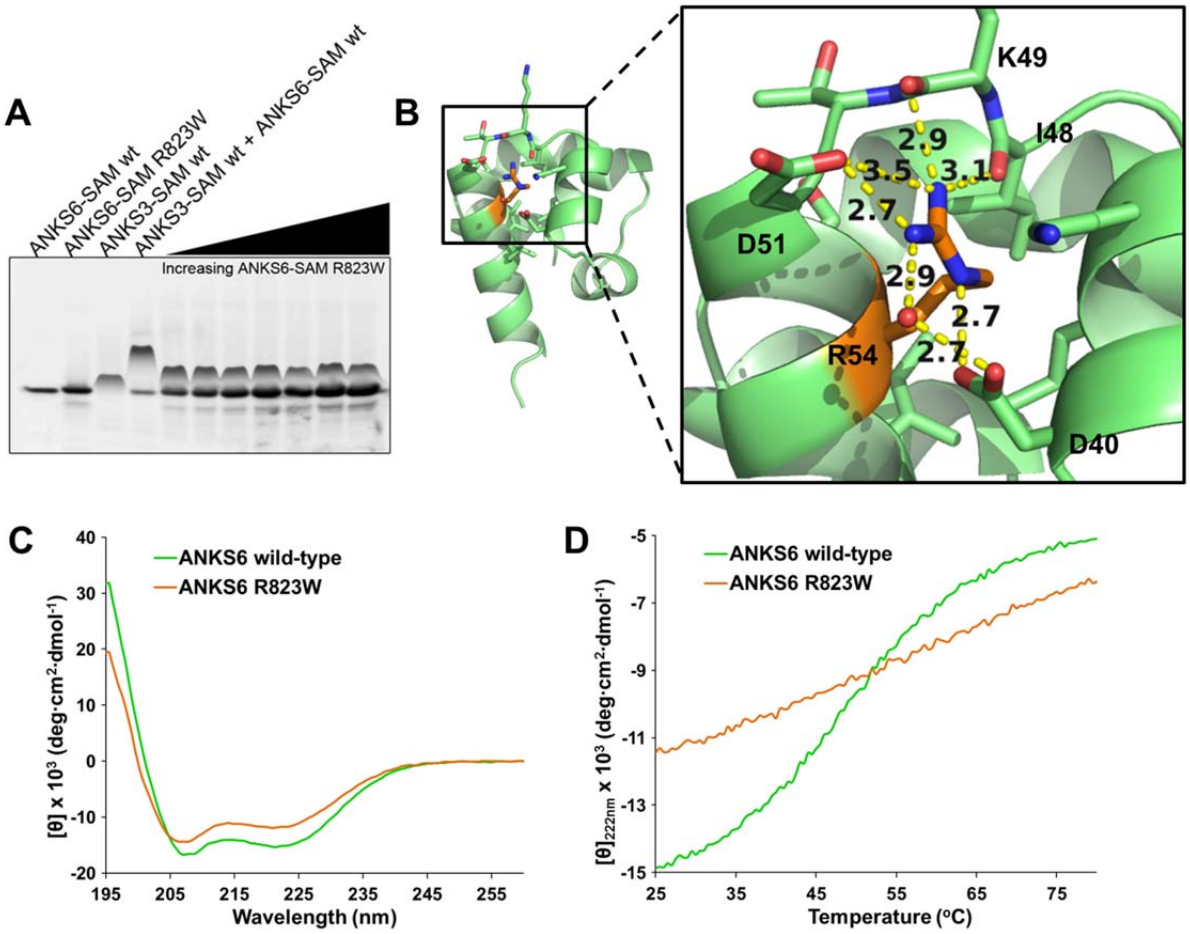


Figure 3-7. Characterizing the Cy mutation

A) negGFP native gel assay of ANKS3-SAM, ANKS6-SAM wt and ANKS6-SAM R823W. A titration series with a constant amount of ANKS3-SAM and increasing amount of ANKS6-SAM R823W is unable to restore the ANKS3-SAM/ANKS6-SAM interaction. B) The Cy mutation (R54) according to our numbering is highlighted on the ANKS6-SAM crystal structure in orange. R54 forms salt bridges and hydrogen bonds with nearby atoms to stabilize the fold in this part of the domain. C) CD spectra of ANKS6-SAM wt and ANKS6-SAM R823W. The reduced CD signal at 222nm and the shifted minima around 208nm correlate with a loss of alpha-helical character and gain of random coil. D) Thermal denaturation curves of ANKS6-SAM wt and ANKS6-SAM R823W monitored by CD signal at 222nm. ANKS6-SAM wt exhibits a broad, weakly cooperative unfolding curve with T_m approx. 48°C, while ANKS6-SAM R823W unfolding is completely uncooperative.

	Iodide derivative of ANKS3-SAM L52A	ANKS3-SAM L52A	ANKS3-SAM/ ANKS6-SAM Heterodimer
PDB Accession #		4NJ8	4NL9
Data collection			
Location	UCLA	APS 24-ID-C	APS 24-ID-C
Space group	$P4_1$	$P4_1$	$C222_1$
Cell dimensions			
a, *b*, *c* (Å)	71.62, 71.62, 33.40	71.89, 71.89, 33.54	47.70, 108.52, 101.74
α, *β*, *γ* (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	2.34	1.60	1.50
R_{sym}	0.118 (.850)	0.087 (.804)	0.069 (.417)
$I/\sigma I$	13.70 (2.62)	20.11 (5.40)	14.83 (4.16)
$CC_{1/2}$	99.7 (74.0)	99.7 (91.7)	99.8 (89.9)
Completeness (%)	99.6 (95.2)	99.3 (99.3)	98.7 (96.9)
Redundancy	6.61 (5.84)	15.15 (15.01)	5.32 (4.94)
Phasing Statistics			
Number of sites	10		
Mean figure of merit			
SAD/after density modification	0.642/0.813		
MapCC (SHELXE)	0.844		
CC (%)	66.73		
Refinement			
Resolution (Å)	2.34	1.60	1.50
No. reflections	7333	22,797	42,203
R_{work}/ R_{free}	0.2709/0.3478	0.1975/0.2177	0.1798/0.2048
No. atoms			
Protein	1030	1022	1952
Water	-	50	246
Magnesium	-	-	1
B-factors (Å^2)			
Protein	36.70	37.42	21.92
Water	-	38.73	28.94
R.m.s deviations			
Bond lengths (Å)	0.011	0.006	0.005
Bond angles (°)	1.303	0.920	0.944

Highest resolution shell is shown in parenthesis.
$R_{sym} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the observed intensity and <I> is the average intensity from observations of symmetry-related reflections. $CC_{1/2}$ = correlation coefficient between two halves of the data [60]. $R_{work} = \sum|F_{obs} - F_{calc}|/\sum F_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. R_{free} is calculated for a set of reflections (10%) that were not included in atomic refinement.

Table 3-1. Crystallographic Data Collection and Refinement Statistics

Human SAM-domain containing protein	UNIPROT ID	Residues
AIDA1C/ANKS1B/ cajalin- 2SAMs	Q7Z6G8	808-950
ANKS1A/ odin- 2SAMs	Q92625	692-836
ANKS6	Q68DC2	771-840
ARAP1/CentaurinD2	Q96P48	1-70
ARAP2/Centaurin D1	Q8WZ64	1-71
BAR	Q9NZS9	177-252
C14orf174	Q9P1V8	541-610
CNKSR2	Q8WXI2	1-77
CNKSR3	Q6P9H4	1-73
DDHD-containing 2	O94830	382-449
ELF3	P78545	48-135
ELF5	Q9UKW6	44-129
EPHA2	P29317	899-970
EPHA5	P54756	960-1030
EPHA6	Q9UF33	956-1023
EPHA7	Q15375	919-998
EPHB1	P54762	906-984
EPHB2	P29323	906-985
EPHB6	O15197	945-1021
ESE3	Q9NZC4	41-120
ETS1	P14921	53-136
GA-binding protein/GABPA	Q06546	170-251
INPPL1	O15357	194-1258
LRSAM1	Q6UWE0	564-634
MOB	Q86VZ5	7-78
Neurabin-1	Q9ULJ8	983-1054
SAMD12	Q8N8I0	71-145
SAMD4B	Q5PRF9	296-361
SAMHD1	Q9Y3Z3	1-111
SAMSN1	Q9NSI8	237-306
SASH1-SAM1	O94885	625-699
SASH1-SAM2	O94885	1173-1247
SASH3	O75995	247-318
SEC23-interacting protein/SEC23IP	Q9Y6Y8	642-703
SLP76/LCP2	Q13094	10-80
Smaug/SAMD4A	Q9UPU9	320-383
StarD13-deletion	Q9Y3MB	154-241
STIM2	Q9P246	129-205
TP63	Q9H3D4	541-609
Usher1G-optimized	Q495M9	381-461
WDSUB1	Q8N9V3	329-396

Table 3-2. Human SAM domains screened for interaction with ANKS3-SAM

negGFP-SAM-domain fusions of the above human SAM-domain containing proteins were screened for binding to negGFP-ANKS3-SAM using the negGFP native gel assay. Sequences of SAM domains used and cloning are as described previously (Knight et al., 2011). Where “2SAMs” is listed, the construct contains two SAM domains in tandem.

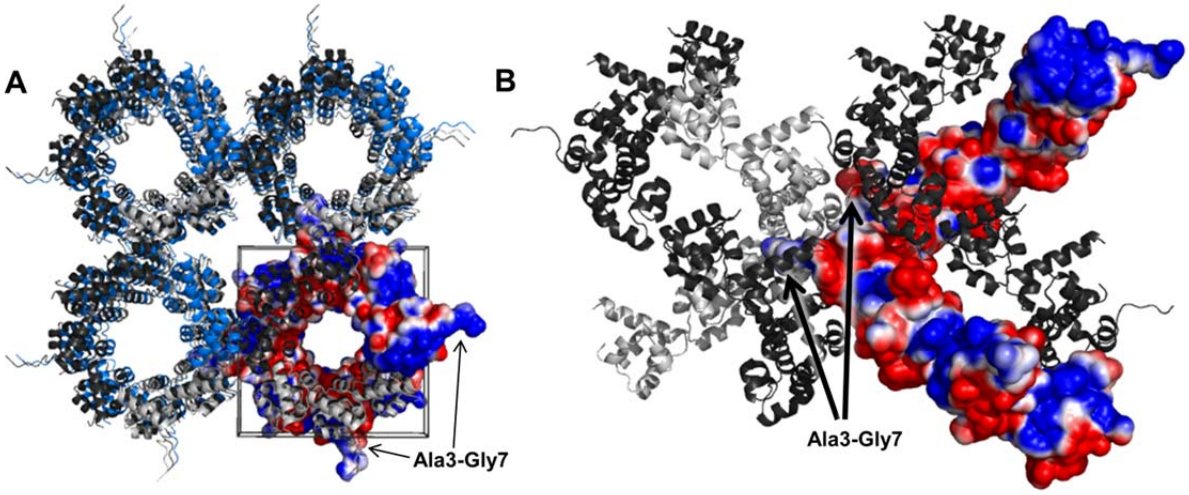


Figure 3-8. Crystal packing of ANKS3-SAM triple helices

A) Individual ANKS3-SAM polymers intertwine to create a triple helix. Triple helices pack side-by-side in the crystal structure. A single ANKS3-SAM triple helix fills the unit cell, shown as a boxed outline. Individual polymers in each triple helix are colored blue, grey, and black. A single polymer is shown as a space-filled model colored by surface electrostatics generated using APBS in Pymol and contouring at ± 1kT/e. Within a single ANKS3-SAM polymer, the N-terminal arm (residues Ala3-Gly7) extends outward from every other SAM domain in the helical segment and forms contacts with a neighboring polymer of triple helices. B) Closer view of the N-terminal arm swapping that occurs between polymers. Two polymers of triple helices are shown, colored as above. Within each, a single polymer has been removed for clarity. Residues Ala3-Gly7 intercalate between two polymers of an adjacent triple helix, thereby forming crystal contacts which stabilize the observed triple helix.

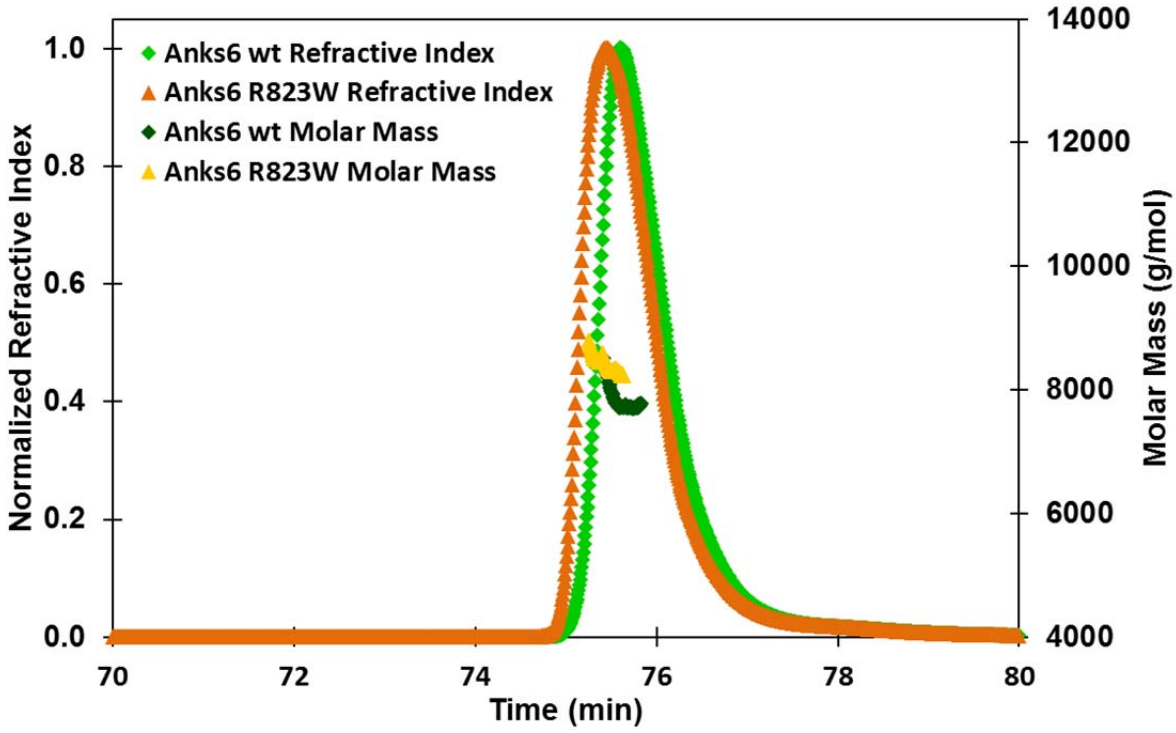


Figure 3-9. Slight unfolding of ANKS6-SAM R823W observed by SEC-MALS

ANKS6-SAM containing the R823W mutation is slightly unfolded compared to ANKS6-SAM wt, as evidenced by the faster migration on SEC-MALS. Faster migration is consistent with a protein that is partially unfolded and therefore exhibits a larger radius of gyration. Both proteins exist as monomers with observed molecular masses of 8.4 kDa for ANKS6-SAM R823W and 7.8 kDa for ANKS6-SAM wt.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Ariza, M., Alvarez, V., Marin, R., Aguado, S., Lopez-Larrea, C., Alvarez, J., Menendez, M.J., and Coto, E. (1997). A family with a milder form of adult dominant polycystic kidney disease not linked to the PKD1 (16p) or PKD2 (4q) genes. J. Med. Genet. *34*, 587–589.

Aviv, T., Lin, Z., Rendl, L., M., Sicheri, F., and Smibert, C.A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. Nat. Struct. Biol. *10*, 614–621.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. *98*, 10037–10041.

Baron, M.K. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. Science *311*, 531–535.

Barrera, F.N. (2003). Binding of the C-terminal Sterile Motif (SAM) Domain of Human p73 to Lipid Membranes. J. Biol. Chem. *278*, 46878–46885.

Bowie,, J.U., Lüthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. Science *253*, 164–170.

Brown, J.H., Bihoreau, M.-T., Hoffmann, S., Kranzlin, B., Tychinskaya, I., Obermuller, N., Podlich, D., Boehn, S.N., Kaisaki, P.J., Megel, N., et al. (2005). Missense Mutation in Sterile Motif of Novel Protein SamCystin is Associated with Polycystic Kidney Disease in (cy/+) Rat. J. Am. Soc. Nephrol. *16*, 3517–3526.

Bryda, E.C., Flaherty, L., Cogswell, C., Price, S.J., Hou, X., and Guay-Woodford, L.M. (2003). Positional cloning of jcpk/bpk locus of the mouse. Mamm. Genome *14*, 242–249.

Chapin, H.C., and Caplan, M.J. (2010). The cell biology of polycystic kidney disease. J. Cell Biol. *191*, 701–710.

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. *2*, 1511–1519.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of *Coot*. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Gabow, P.A. (1993). Autosomal Dominant Polycystic Kidney Disease. N. Engl. J. Med. *329*, 332–342.

Gretz, N., Kränzlin, B., Pey, R., Schieren, G., Bach, J., Obermüller, N., Ceccherini, I., Klöting, I., Rohemeiss, P., and Bachmann, S. (1996). Rat models of autosomal dominant polycystic kidney disease. Nephrol. Dial. Transplant. *11*, 46–51.

Grimshaw, S.J., Mott, H.R., Stott, K.M., Nielsen, P.R., Evetts, K.A., Hopkins, L.J., Nietlispach, D., and Owen, D. (2003). Structure of the Sterile Motif (SAM) Domain of the Saccharomyces cerevisiae Mitogen-activated Protein Kinase Pathway-modulating Protein STE50 and Analysis of Its Interaction with the STE11 SAM. J. Biol. Chem. *279*, 2192–2201.

Grucza, R.A., Bradshaw, J.M., Mitaxov, V., and Waksman, G. (2000). Role of Electrostatic Interactions in SH2 Domain Recognition: Salt-Dependence of Tyrosyl-Phosphorylated Peptide Binding to the Tandem SH2 Domain of the Syk Kinase and the Single SH2 Domain of the Src Kinase. Biochemistry *39*, 10072–10081.

Guay-Woodford, L.M. (2003). Murine models of polycystic kidney diease; molecular and therapeutic insights.pdf. Am. J. Physiol. - Ren. Physiol. *285*, F1034–F1049.

Gundelfinger, E.D., Boeckers, T.M., Baron, M.K., and Bowie, J.U. (2006). A role for zinc in postsynaptic density asSAMbly and plasticity? Trends Biochem. Sci. *31*, 366–373.

Hanaoka, K., Qian, F., Boletta, A., Bhunia, A.K., Piontek, K., Tsiokas, L., Sukhatme, V.P., Guggino, W.B., and Germino, G.G. (2000). Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. Nature *408*, 990–994.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Hileman, R.E., Jennings, R.N., and Linhardt, R.J. (1998). Thermodynamic Analysis of the Heparin Interaction with a Basic Cyclic Peptide Using Isothermal Titration Calorimetry. Biochemistry *37*, 15231–15237.

Hoff, S., Halbritter, J., Epting, D., Frank, V., Nguyen, T.-M.T., van Reeuwijk, J., Boehlke, C., Schell, C., Yasunaga, T., Helmstädter, M., et al. (2013). ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3. Nat. Genet. *45*, 951–956.

Kabsch, W. (2010). *XDS*. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kim, C.A., and Bowie, J.U. (2003). SAM domains: uniform structure, diversity of function. Trends Biochem. Sci. *28*, 625–628.

Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S., and Bowie, J.U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J. *20*, 4173–4182.

Kim, C.A., Gingery, M., Pilpa, R.M., and Bowie, J.U. (2002). The SAM domain of polyhomeotic forms a helical polymer. Nat. Struct. Biol. *9*, 453–457.

Kim, C.A., Sawaya, M.R., Cascio, D., Kim, W., and Bowie, J.U. (2005). Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. J. Biol. Chem. *280*, 27769–27775.

Knight, M.J., Joubert, M.K., Plotkowski, M.L., Kropat, J., Gingery, M., Sakane, F., Merchant, S.S., and Bowie, J.U. (2010). Zinc Binding Drives Sheet Formation by the SAM Domain of Diacylglycerol Kinase δ. Biochemistry *49*, 9667–9676.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Kugita, M., Nishii, K., Morita, M., Yoshihara, D., Kowa-Sugiyama, H., Yamada, K., Yamaguchi, T., Wallace, D.P., Calvet, J.P., Kurahashi, H., et al. (2010). Global gene expression profiling in early-stage polycystic kidney disease in the Han:SPRD Cy rat identifies a role for RXR signaling. AJP Ren. Physiol. *300*, F177–F188.

Kwan, J.J., Warner, N., Pawson, T., and Donaldson, L.W. (2004). The Solution Structure of the S.cerevisiae Ste11 MAPKKK SAM Domain and its Partnership with Ste50. J. Mol. Biol. *342*, 681–693.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lawrence, M.S., Phillips, K.J., and Liu, D.R. (2007). Supercharging Proteins Can Impart Unusual Resilience. J. Am. Chem. Soc. *129*, 10110–10112.

Leone, M., Cellitti, J., and Pellecchia, M. (2008). NMR Studies of a Heterotypic Sam−Sam Domain Association: The Interaction between the Lipid Phosphatase Ship2 and the EphA2 Receptor. Biochemistry *47*, 12721–12728.

Leone, M., Cellitti, J., and Pellecchia, M. (2009). The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. BMC Struct. Biol. *9*, 59.

Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D., and Butt, T.R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics *5*, 75–86.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). *Phaser* crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Nagao, S., Kugita, M., Yoshihara, D., and Yamaguchi, T. (2012). Animal models for human polycystic kidney disease. Exp. Anim. *61*, 477–488.

Nauli, S., Farr, S., Lee, Y.-J., Kim, H.-Y., Faham, S., and Bowie, J.U. (2007). Polymer-driven crystallization. Protein Sci. *16*, 2542–2551.

Neudecker, S., Walz, R., Menon, K., Maier, E., Bihoreau, M.-T., Obermüller, N., Kränzlin, B., Gretz, N., and Hoffmann, S.C. (2010). Transgenic Overexpression of Anks6(p.R823W) Causes Polycystic Kidney Disease in Rats. Am. J. Pathol. *177*, 3000–3009.

Pape, T., and Schneider, T.R. (2004). *HKL2MAP* : a graphical user interface for macromolecular phasing with *SHELX* programs. J. Appl. Crystallogr. *37*, 843–844.

Di Pietro, S.M., Cascio, D., Feliciano, D., Bowie, J.U., and Payne, G.S. (2010). Regulation of clathrin adaptor function in endocytosis: novel role for the SAM domain. EMBO J. *29*, 1033–1044.

Qiao, F., and Bowie, J.U. (2005). The many faces of SAM. Sci. Signal. *2005*, re7.

Qiao, F., Song, H., Kim, C.A., Sawaya, M.R., Hunter, J.B., Gingery, M., Rebay, I., Courey, A.J., and Bowie, J.U. (2004). Derepression by depolymerization: structural insights into the regulation of Yan by Mae. Cell *118*, 163–173.

Qiao, F., Harada, B., Song, H., Whitelegge, J., Courey, A.J., and Bowie, J.U. (2005). Mae inhibits Pointed-P2 transcriptional activity by blocking its MAPK docking site. EMBO J. *25*, 70–79.

Rajakulendran, T., Sahmi, M., Kurinov, I., Tyers, M., Therrien, M., and Sicheri, F. (2008). CNK and HYP form a discrete dimer by their SAM domains to mediate RAF kinase signaling. Proc. Natl. Acad. Sci. *105*, 2836–2841.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Schafer, K., Gretz, N., Bader, M., Oberbaumer, I., Eckardt, K.-U., Kriz, W., and Bachmann, S. (1994). Characterization of the Han:SPRD rat model for hereditary polycystic kidney disease. Kidney Int. *46*, 134–152.

Schatz, P.J., Cull, M.G., Martin, E.L., and Gates, C.M. (1996). Screening of Peptide Libraries Linked to lac Repressor. Methods Enzymol. *267*, 171–191.

Schrödinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. *19*, 1354–1365.

Sheldrick, G.M. (2010). Experimental phasing with *SHELXC* / *D* / *E* : combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. *66*, 479–485.

Stafford, R.L., Hinde, E., Knight, M.J., Pennella, M.A., Ear, J., Digman, M.A., Gratton, E., and Bowie, J.U. (2011). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK. Structure *19*, 1826–1836.

Stagner, E.E., Bouvrette, D.J., Cheng, J., and Bryda, E.C. (2009). The polycystic kidney disease-related proteins Bicc1 and SamCystin interact. Biochem. Biophys. Res. Commun. *383*, 16–21.

Torres, V.E., Harris, P.C., and Pirson, Y. (2007). Autosomal dominant polycystic kidney disease. The Lancet *369*, 1287–1301.

Wilson, P.D. (2004). Polycystic Kidney Disease. N. Engl. J. Med. *350*, 151–164.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the *CCP* 4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Zhang, H., Xu, Q., Krajewski, S., Krajewska, M., Xie, Z., Fuess, S., Kitada, S., Godzik, A., and Reed, J.C. (2000). BAR: An apoptosis regulator at the intersection of caspases and Bcl-2 family proteins. Sci. Signal. *97*, 2597.

Chapter 4

Characterization of the SAM domain of BICC1 and its interactions with the cystic kidney disease related proteins ANKS3 and ANKS6

Abstract

The RNA-binding protein bicaudal C homolog 1 (BICC1) is a key translational regulator. Mutations of BICC1 are associated with cystic renal dysplasia in humans and a PKD-like phenotype in several animal models. BICC1 contains a C-terminal SAM domain which is required for localization of BICC1 to the periphery of RNA-processing P-bodies and for the inhibitory effect of BICC1 on canonical Wnt signaling. To better understand the molecular importance of the BICC1 SAM domain we solved a high-resolution crystal structure and show that the BICC1 SAM domain forms a helical polymer with moderate affinity. TEM imaging of the wild-type BICC1 SAM domain reveals a propensity of individual polymers to assemble into ordered sheet-like arrays. The SAM domain-containing proteins ANKS3 and ANKS6 have recently been shown to associate with BICC1 and mutations within these proteins are associated with cystic kidney disease. Here we demonstrate that the SAM domains of BICC1 and ANKS3 bind each other through both of their conserved ML and EH interaction surfaces. Since the measured binding affinities of the two possible interfaces are nearly identical, the SAM domains of BICC1 and ANKS3 appear capable of forming the first observed alternating SAM domain co-polymer. We also demonstrate that the BICC1 SAM domain can bind the ANKS6 SAM domain with strong affinity to create a heterodimer. Furthermore, we show that the disease-associated point mutations R823W and I817N (originally identified as I747N in mouse) in ANKS6 destabilize binding to the BICC1 SAM domain. Analysis of secondary structure by circular dichroism reveals that the I817N mutation strongly destabilizes the ANKS6 SAM domain to the extent that it is primarily unfolded. The variety of possible interactions between the SAM domains of BICC1, ANKS3, and ANKS6 allow the formation of numerous different protein complexes. That BICC1, ANKS3, and ANKS6 are linked to cystic kidney diseases suggests the diverse cellular scaffolds created by this network of SAM domain interactions are important for renal development.

Introduction

Cystic diseases of the kidney including polycystic kidney disease (PKD) and nephronophthisis (NPHP) are a common genetic cause of end-stage renal failure (Lancaster and Gleeson, 2010). Both involve the formation of renal cysts, and in the case of NPHP additional disruptions to the kidney structure along with extrarenal manifestations (Hildebrandt et al., 2009; Wilson, 2004). PKD and NPHP are part of a larger class of cystic kidney diseases referred to as ciliopathies, owing to the observation that many of the gene products disrupted in these conditions localize to the primary cilium or structures at its base (Hildebrandt et al., 2011; Lancaster and Gleeson, 2010). PKD is mostly attributed to mutations in the proteins polycystin-1, polycystin-2, or fibrocystin (Wilson, 2004). Mutations in 19 different genes including *NPHP1–NPHP13* are currently known to result in NPHP and NPHP-related ciliopathies, although in roughly 50% of cases the causative mutation has not yet been identified (Chaki et al., 2012; Failler et al., 2014; Halbritter et al., 2013; Hildebrandt et al., 2009; Hoff et al., 2013; Schueler et al., 2015; The GPN Study Group et al., 2013). Accumulating evidence from animal models and human cases of PKD and NPHP implicate the sterile alpha motif (SAM) domain containing proteins bicaudal C homolog 1 (BICC1), ankyrin repeat and SAM-domain containing protein 6 (ANKS6), and ankyrin repeat and SAM-domain containing protein 3 (ANKS3) as being involved in these diseases and important for kidney development.

SAM domains are common protein-protein interaction moieties composed of approximately 70 amino acids folded into an α-helical bundle (Kim and Bowie, 2003; Qiao and Bowie, 2005). SAM domains typically interact with each other via two distinct surfaces known as the mid-loop (ML) and end-helix (EH) surfaces. Sequential association of the ML and EH surfaces of neighboring SAM domains results in the formation of helical polymers (Baron et al., 2006; Harada et al., 2008; Kim et al., 2001, 2002; Knight et al., 2011; Leettola, et al., 2014; Di Pietro et al., 2010; Stafford et al., 2011), closed oligomers (Ramachander, 2002), or

heterodimers (Grimshaw et al., 2003; Kwan et al., 2004; Leettola, et al., 2014; Leone et al., 2009; Qiao et al., 2004; Ramachander, 2002). However, SAM domains have also been shown to associate with RNA, lipids, and non-SAM-domain containing proteins (Aviv et al., 2003; Barrera, 2003; Kim and Bowie, 2003; Qiao et al., 2005; Zhang et al., 2000). SAM domains have unique cellular functions including gene regulation (Kim et al., 2001; Qiao et al., 2004), control of enzyme localization (Harada et al., 2008; Knight et al., 2010), and scaffolding (Baron et al., 2006; Gundelfinger et al., 2006; Stafford et al., 2011). Because many SAM-domain containing proteins have additional functional domains, the side-by-side assemblage of SAM domain polymers creates a platform from which these additional functional domains splay outward, thereby assisting in the organization of large protein complexes.

In the PKD/Mhm(*cy*/+) rat model of PKD a missense point mutation in the gene encoding ANKS6 results in an R823W mutation within the C-terminal SAM domain that is causative of disease (Brown et al., 2005; Neudecker et al., 2010). A newly established mouse model with an I747N point mutation within the SAM domain of ANKS6 also presents with cystic kidney disease, although with phenotypic differences that are distinct from the rat model (Bakey et al., 2015). The importance of ANKS6 is further illustrated by knockdowns of ANKS6 in zebrafish and *Xenopus* that result in renal developmental defects (Hoff et al., 2013). Mutations in ANKS6 affecting both the SAM domain and the N-terminal ankyrin repeats have also recently been identified in human cases of NPHP (Hoff et al., 2013). Efforts to understand the function of ANKS6 have revealed that ANKS6 interacts with a number of NPHP proteins (Hoff et al., 2013). In particular, ANKS6 interacts with the Nek8 kinase (NPHP9) via its ankyrin repeats and this interaction localizes ANKS6 to the inversin compartment of primary cilia (Czarnecki et al., 2015).

By mass spectrometric analysis of immunoprecipitated TAP-tagged and FLAG-tagged NPHP proteins ANKS6 has also been detected in complex with ANKS3 (Czarnecki et al., 2015; Hoff et al., 2013). In our previous work we characterized this novel interaction and showed that

it is mediated by the SAM domains of each protein. Specifically, the SAM domain of ANKS3 is weakly polymeric and forms a strong interaction with the monomeric SAM domain of ANKS6 (Leettola, et al., 2014). We also demonstrated that in the context of purified SAM domains, the R823W mutation results in partial loss of ANKS6 SAM tertiary structure and the loss of interaction with ANKS3. Recent studies using GFP-tagged ANKS3 showed that ANKS3 co-localizes with basal bodies at the base of cilia and is also found in large intracellular aggregates that are suggested to be polymers of ANKS3 (Yakulov et al., 2015). Additionally, ANKS3 has been shown to interact with NPHP proteins as well as with BICC1 and depletion of ANKS3 in zebrafish causes an NPHP-like phenotype with cystic kidneys and laterality defects (Yakulov et al., 2015).

BICC1 is an RNA-binding protein and a key developmental regulator. Mutations or altered expression of Bicaudal C in *Drosophila*, *C. elegans* and *Xenopus* result in developmental defects (Eckmann et al., 2002; Mahone et al., 1995; Maisonneuve et al., 2009; Mohler and Wieschaus, 1986). Mutations in BICC1 are also responsible for disease in the *jcpk* and *bpk* mouse models of PKD (Bryda et al., 2003). Furthermore, knockdown of BICC1 in zebrafish and *Xenopus* results in a PKD-like phenotype and human mutations in BICC1 are associated with cystic renal dysplasia (Bouvrette et al., 2010; Kraus et al., 2012; Tran et al., 2007). BICC1 contains 3 RNA-binding K-homology (KH) domains at its N-terminus and a SAM domain at its C-terminus. In several mammalian cell lines BICC1 has been show to form tube and vesicle-like structures around the periphery of RNA-processing P-bodies and this localization is dependent on the SAM domain (Maisonneuve et al., 2009). BICC1 is also involved in the post-transcriptional control of mRNA. In one instance, BICC1 promoted the translation and stability of polycystin-2 mRNA (*Pkd2* is commonly mutated in human PKD) by antagonizing a *mir-17*-containing silencing complex (Tran et al., 2010). In another instance, BICC1 has been shown to promote the silencing of adenylate cyclase 6 (AC6) mRNA and the

SAM domain in particular was required for successful transfer of the target mRNA to AGO2 to form a functional RISC complex (Piazzon et al., 2012). As a key developmental regulator, BICC1 has an inhibitory function on the canonical Wnt signaling pathway and this activity is highly dependent on the SAM domain (Maisonneuve et al., 2009). The mutation E932G within the BICC1 SAM domain is associated with human cystic renal dysplasia and while this mutation does not impact localization of BICC1 to P-bodies, it reduces the inhibitory effect of BICC1 on Wnt signaling to the same extent as a complete deletion of the SAM domain (Kraus et al., 2012).

Clearly the SAM domain of BICC1 is integral to its function. However, direct binding partners of the BICC1 SAM domain have yet to be identified. Furthermore, although prior work has established that the BICC1 SAM domain is polymeric (Knight et al., 2011), a specific understanding of the biochemical and structural basis behind polymer formation is lacking. Herein we present a high-resolution crystal structure of the polymeric form of the BICC1 SAM domain and show that the BICC1 SAM domain strongly self-associates and is capable of organizing into sheets of polymers. We also identify and characterize the SAM domain of the NPHP-related protein ANKS3 as the first direct identified binding partner of the BICC1 SAM domain. Given the high degree of sequence and structural similarity between the SAM domains of BICC1, ANKS3, and ANKS6 we further show that the BICC1 SAM domain is capable of binding the ANKS6 SAM domain and that mutations in ANKS6 that are causative of PKD perturb this binding. The ability of the ANKS3, BICC1, and ANKS6 SAM domains to bind each other creates numerous possible scaffolds which may be important for organizing signaling complexes within the kidney.

Results

Structure of the BICC1 SAM polymer

Prior research using a negGFP tagged construct of the BICC1 SAM domain indicated that the SAM domain is polymeric (Knight et al., 2011). We re-affirmed this conclusion here by repeating the negGFP native gel assay done previously. In this assay, SAM domains are fused to a green fluorescent protein (GFP) modified to have a net charge of -30 such that the degree of migration on a native gel is proportional to the extent of polymer formation. We again see that the wild-type BICC1 SAM-negGFP fusion migrates with reduced mobility and a smeared band pattern when compared to a monomeric SAM domain (ANKS6, Fig. 4-10A) and a weakly polymeric SAM domain (ANKS3, Fig. 4-8B) which is consistent with strong polymer formation. To better understand the polymeric form of the BICC1 SAM domain we solved high resolution crystal structures of the R924E mutant (Fig. 4-1, Table 4-1). This mutation prevents BICC1 SAM polymerization *in vitro* (Fig. 4-8B) but under the high concentrations required for crystal formation, polymer contacts reform as crystal contacts. This method of crystallizing soluble SAM domain point mutants has been successfully used previously to determine the polymeric structures of several SAM domains (Baron et al., 2006; Harada et al., 2008; Kim et al., 2001; Leettola, et al., 2014). We used multi-wavelength anomalous diffraction (MAD) to solve a selenomethionine (SeMet) labeled BICC1 SAM R924E mutant to 1.75 Å resolution. We also crystallized a native R924E mutant which had a slightly more compact unit cell and solved this structure to 2.00 Å resolution by molecular replacement. Both the native and SeMet constructs crystallized in space group $P2_12_12_1$ with 3 molecules in the asymmetric unit.

The BICC1 SAM polymer formed within the crystal contains 6 SAM domains per helical repeat (Fig. 4-1A). Interestingly, the SAM polymer is not symmetrical. This is because although each subunit within the asymmetric unit adopts the same fold (RMSD across all atoms is 0.430-

0.631 Å, RMSD across backbone atoms is 0.412-0.607 Å) the orientation between chains is different (Fig. 4-1B). SAM domain polymers are known to be flexible but never before has such a variety of interaction surfaces been observed within a homotypic SAM polymer.

Analysis of the intra-polymer binding interfaces shows that BICC1 SAMs associate via the sequential interaction of the mid-loop (ML) and end helix (EH) binding surfaces. The ML surface of BICC1 is composed of a shallow hydrophobic patch ringed by negatively charged residues (Fig. 4-1C, left). The EH surface complements the ML surface in both shape and electrostatics (Fig. 4-1C, right). It has a Phe protrusion which packs against the ML surface hydrophobic patch and is surrounded by positively charged residues that form ionic and hydrogen bonds with the ML surface. This structure corroborates our negGFP native gel analysis of residues critical to each interface (Fig. 4-8B). We found that the ML surface mutations E898K, D900K, T903E, D911K and the EH surface mutations K913E, F920E, R924E, K925E each destabilized polymer formation of BICC1, as indicated by the increased mobility and discrete banding pattern of these negGFP fusion constructs run on a native gel (Fig.4-8B). Analysis of the interactions formed by each of these residues in each interface show ionic interactions (D900, D911, R924E, K913) hydrogen bond formation (E898, K913, K925) hydrophobic packing (F920), and shape complementarity (T903 packs against the helical backbone of the EH surface) (Fig. 4-1C and Fig. 4-2). As such, the mutations we generated caused a loss of interactions and created charge repulsion and steric clashes at the binding interface, resulting in monomeric protein.

Inspection of these critical residues at the ML and EH surfaces also lends plausibility that the three different binding interfaces we observed may be biologically relevant. All the residues identified by point mutation to be involved in polymerization make contacts in at least one of the observed binding interfaces but there is no binding interface where all residues contribute (Fig. 4-2). However, the CB interface seems most likely since it buries the most

surface area and all critical residues make interface contacts except K913, which when mutated was the weakest polymerization destabilizer. These different interaction interfaces suggest the BICC1 polymer may be highly flexible and indeed the native R924E mutant crystallized as a slightly more compact polymer (Fig. 4-3). This flexibility could be explained by the electrostatic nature of the ML-EH binding interface, which would allow malleable docking of subunits. Flexible binding between SAM domains has been observed previously between the SAM domains of human EphA2 and SHIP2, which form a primarily polar and electrostatic ML-EH binding interface for which two different orientations are suggested to be populated (Lee et al., 2012). However, we cannot rule out that the varied interface orientations we observe may be an artifact due to the point mutation of the construct we crystallized.

The E932G mutation does not affect BICC1 SAM polymerization

To better understand the physiological relevance of BICC1 SAM polymer formation we determined the binding affinity of BICC1 SAM for itself using surface plasmon resonance (SPR). To do this, we immobilized the BICC1 SAM R924E mutant, which has a defective EH surface but retains a wild-type ML surface. We then flowed over the BICC1 SAM E898K and D911K mutants, which have defective ML surfaces but wild-type EH surfaces. With this set-up the proteins are only able to interact using their wild-type interfaces, allowing us to measure the average K_d of BICC1 SAM for itself as 1.35 ± 0.1 μM (Fig. 4-4A). No binding response was observed when the BICC1 SAM mutants F920E and K925E, each with defective EH surfaces, and the BICC1 double mutants E898K/F920E and D911K/F920E were passed over the conjugated chip, indicating that the measured binding affinity is specific for the native ML-EH interface and not an alternate binding site (Fig. 4-5).

The mutation E932G in the SAM domain of human BICC1 is associated with cystic renal dysplasia and a reduced ability of BICC1 to block canonical Wnt signaling (Kraus et al., 2012). Because the molecular basis of how this point mutation causes disease is unclear we investigated whether the E932G point mutation impacted the ability of BICC1 SAM to polymerize. A negGFP fusion of BICC1 SAM E932G migrated with the same mobility as wild-type BICC1 on a native gel, indicating this point mutation does not impact polymer formation (Fig. 4-4B). Additionally, E932G is more than 10Å away from the nearest residue of the neighboring ML surface in the crystal structure, making it too far removed to engage in any interaction between neighboring SAM domains (Fig. 4-1C).

BICC1 SAM polymers form ordered arrays

Among the different interfaces we observed in the crystal structures, the BA interfaces of both the SeMet and the native structure contain a zinc ion that bridges an interaction between residues R924E and D911 (Fig. 4-2). In the wild-type protein, the zinc ion is likely displaced by the arginine at position 924 which is expected to form a salt bridge with D911, thereby making the presence of zinc at the wild-type interface questionable. However, the zinc ion within the crystal structure also mediates inter-polymer contacts and SAM domain polymers have previously been observed to assemble into sheets of polymers in the presence of zinc (Baron et al., 2006; Knight et al., 2010). Interestingly, when a large molar excess of divalent metal is added to either purified wild-type BICC1 SAM or BICC1 SAM E932G, only zinc is able to induce precipitation of the protein (Fig 4-6A). When examined under a light microscope, this zinc-induced precipitate forms needle-like fibers which are birefringent, indicative of supramolecular ordering (Fig. 4-6B). Investigation of the precipitate by transmission electron microscopy

reveals large fiber-like sheets in both the wild-type and E932G mutant, further supporting that this mutation does not impact the ability of BICC1 SAM to oligomerize (Figs. 4-6 C-I).

These ordered arrays induced by zinc form crystalline sheets composed of polymers organized side-by-side (Figs. 4-6 C-E, G-H). Individual polymers measure approximately 7nm wide and approximately 6nm per helical repeat. These measurements agree well with the BICC1 SAM polymers formed in the crystal structure. However, polymer packing appears slightly different from that observed in the crystal structure, where polymers align side-by-side via a simple translation (Fig. 4-7). In contrast, the polymers observed by TEM appear translated and diagonally shifted relative to each other (Fig. 4-6E). Additionally, polymers within the crystal structure pack tightly and interdigitate (Fig. 4-7) while the polymers observed by TEM appear well spaced. This however could be an artifact of uranyl acetate staining which may obscure interpolymer structural details. These results suggest that although individual polymers appear similar to those observed in the crystal structure, interpolymer contacts may differ between the two conditions.

Inspection of the TEM grid also reveals thin string-like fibers which appear to be short BICC1 SAM polymers associated side-by-side (Fig. 4-6F,I). These polymers have the same dimensions as those organized into large sheets but only consist of a few helical repeats. These same fibers composed of short polymers are also observed in wild-type BICC1 SAM in the absence of zinc (Fig. 4-4C). This supports that wild-type BICC1 SAM is not only polymeric but has a natural propensity for interpolymer association. That large sheets composed of longer SAM polymers are only observed in the presence of zinc suggests that zinc enhances both side-by-side polymer association and intra-polymer contacts. However, inspection of the crystal structure does not reveal any obvious zinc binding sites and BICC1 SAM does not contain any residues such as histidine or cysteine which are typically involved in zinc binding. We therefore

suggest that BICC1 SAM is capable of forming large polymers that can assemble into sheets and that zinc enhances this natural ability.

The BICC1 and ANKS3 SAM domains associate

Close inspection of the BICC1 SAM domain structure reveals that it is highly similar to that of the ANKS3 SAM domain (average RMSD across backbone atoms = 0.754 Å, average RMSD across all atoms = 0.814 Å), a structure which we solved previously (Leettola, et al., 2014). This is not entirely surprising since a pairwise alignment of the SAM domains of ANKS3 and BICC1 shows the sequences are 50% identical and 70% similar (Fig. 4-8A). If this analysis is restricted solely to the binding interfaces, the similarity becomes even more pronounced: 63% identity and 89% similarity for the ML surface; 67% identity and 83% similarity for the EH surface. From a structural perspective, it therefore follows that the SAM domain of BICC1 should be capable of binding the SAM domain of ANKS3 and indeed these proteins have recently been shown to associate (Yakulov et al., 2015).

We first tested this possibility using our negGFP native gel screen. When negGFP tagged wild-type BICC1 SAM is mixed with wild-type ANKS3 SAM in a 1:1 ratio a slight upshift accompanied by the loss of the ANKS3 band indicates an interaction (Fig. 4-8B). To understand which interfaces of each protein were responsible for generating this hetero-interaction we mixed ANKS3 SAM in a 1:1 ratio with monomeric point mutants of BICC1 SAM, each containing a mutation at either the ML or EH surface. We were surprised to find that when the ML surface of BICC1 is mutated, the native EH surface is capable of binding ANKS3 (Fig 4-8B, left) and likewise, when the EH surface of BICC1 is mutated, the native ML surface is capable of binding ANKS3 (Fig. 4-8B, right). This suggested that BICC1 SAM is capable of binding the SAM domain of ANKS3 using both its ML and EH surfaces.

To further test this, we mixed the BICC1 ML mutants E898K and D911K with the ANKS3 EH mutant F472E in 1:1 molar ratios and looked for elution profile changes by analytical size exclusion chromatography. With this set-up, the native ML surface of ANKS3 F472E should be capable of interacting with the native EH surface of the BICC1 ML mutants to form a heterodimer. When these constructs were analyzed by size exclusion, the 1:1 mixes showed a definite elution upshift consistent with formation of a heterodimer (Fig. 4-8C). We also tested the alternate interface by mixing the BICC1 EH mutants F920E and R924E with the ANKS3 ML mutant I455E in 1:1 molar ratios and analyzed the protein by SEC-MALS. In this scenario, the native ML surface of the BICC1 SAM domains should be capable of interacting with the native ANKS3 EH surface. As predicted, the proteins formed a heterodimer and eluted as single peaks with a molecular weight of 17.4 ± 0.6 kDa for the BICC1 R924E/ANKS3 I455E heterodimer and 18.4 ± 0.1 kDa for the BICC1 F920E/ANKS3 I455E heterodimer (predicted molecular weights were 16.1 kDa) (Fig. 4-8D). This supports that the BICC1-EH surface binds the ANKS3-ML surface and also that the BICC1-ML surface binds the ANKS3-EH surface.

To further characterize these different hetero-interactions we tested the strength of the binding interfaces using SPR. To determine the binding affinity of the BICC1-ML/ANKS3-EH interface, we immobilized the BICC1 EH mutant R924E (with a native ML surface) and flowed over increasing concentrations of the ANKS3 ML mutant I455E (with a native EH surface). Since only the native interfaces are capable of interacting we determined the binding affinity for the BICC1-ML/ANKS3-EH interface as 2.2 ± 0.1 μM (Fig. 4-8E). To probe the alternate interface formed by the BICC1-EH and ANKS3-ML surfaces, we immobilized the ANKS3 EH mutant F472E (with a native ML surface) and flowed over increasing concentrations of the BICC1 ML mutants E898K and D911K (with native EH surfaces). Again, only the native interfaces are capable of interacting, allowing us to determine the binding affinity for the BICC1-EH/ANKS3-ML interface as 1.35 ± 0.1 μM (Figs. 4-8 F-G). To verify that this measured binding

affinity was specific for the BICC1-EH/ANKS3-ML surface we passed the BICC1 double mutants E898K/F920E and D911K/F920E, each with defective ML and EH surfaces, over the conjugated chip and observed minimal binding (Fig. 4-9A). Given that BICC1 and ANKS3 can interact with each other through both their ML and EH surfaces and that the affinities of these different interfaces are roughly the same, this creates the possibility that these SAM domains associate *in vivo* as an alternating co-polymer. SAM domains have previously been observed to form block co-polymers (Kim et al., 2005) but this is the first instance of a possible alternating co-polymer.

The BICC1 and ANKS6 SAM domains associate

We previously discovered and characterized the interaction between the SAM domains of ANKS3 and ANKS6 (Leettola, et al., 2014). In this interaction, the EH surface of ANKS3 binds the ML surface of ANKS6 with high affinity. Given the sequence and structural conservation between the BICC1 and ANKS3 SAM domains, it follows from a structural and biochemical standpoint that the EH surface of BICC1 should also be capable of binding ANKS6. We tested this hypothesis by checking for the predicted hetero-interaction using our negGFP native gel screen (Fig. 4-10A). Oddly, when the wild-type negGFP fusion proteins were mixed in a 1:1 molar ratio, these SAMs did not appear to interact as there was no gel shift. However, when point mutants of BICC1 were used in this assay, BICC1 ML surface mutants with a wild-type EH surface show clear evidence of a hetero-interaction with ANKS6 SAM. This is further supported by a mutation at the hypothetically involved EH interface (F920E) that results in no gel shift, indicating no interaction. It should be noted that the magnitude of the gel-shift between the BICC1 ML mutants and ANKS6 is the same as that observed when these BICC1 ML mutants were mixed with ANKS3 (Fig. 4-8B). Even though the binding affinities between the

different BICC1-ANKS3 interfaces are nearly identical, the BICC1 ML mutants exhibit less of a native gel shift than the EH mutants. We therefore suspect that the differing magnitudes of native gel-shift are an artifact of the negGFP fusions and not entirely indicative of the strength of the binding interaction.

To verify the interaction between the BICC1-EH and ANKS6-ML surfaces we mixed purified BICC1 ML mutants E898K and D911K with ANKS6 in 1:1 molar ratios and analyzed the proteins by SEC (Fig. 4-10B). As predicted, the BICC1-ANKS6 mixes eluted earlier than their monomeric counterparts, consistent with formation of a heterodimer between the ANKS6-ML and BICC1-EH surfaces. We next tested the binding affinity of this interaction by SPR using a chip with immobilized ANKS6 SAM (Figs. 4-10 C-D). We flowed over the BICC1 ML mutants E898K and D911K, each with a native EH surface capable of binding ANKS6, and measured a strong binding affinity for this interaction (average K_d = 460.5 ± 10 nM). As a control, we flowed the BICC1 double mutants E898K/F920E and D911K/F920E, each with both the ML and EH surfaces knocked out, over the ANKS6 conjugated chip and observed minimal binding (Fig. 4-9B). This indicates that the interaction between BICC1 and ANKS6 is indeed mediated by the EH surface of BICC1 and not an alternate binding site. BICC1 and ANKS6 have previously been shown to associate (Stagner et al., 2009; Yakulov et al., 2015) but their association was not predicted to involve each of their respective SAM domains, as we have shown here.

The human ANKS6 I817N mutation destabilizes the SAM domain

It was recently discovered that the point mutation I747N in the SAM domain of mouse ANKS6 produces a cystic kidney disease phenotype in mutant mice (Bakey et al., 2015). This point mutation corresponds to I817N in the human ANKS6 protein and lies in close proximity to the R823W mutation which is causative of disease in the PKD/Mhm(*cy*/+) rat model of PKD

(Brown et al., 2005). Our previous analysis showed that the R823W mutation causes localized unfolding of the ANKS6 SAM domain and a loss of interaction with the SAM domain of ANKS3 (Leettola, et al., 2014). Because the I817N point mutation lies in close proximity to the R823W mutation, we reasoned it may also have an effect on the tertiary structure of the ANKS6 SAM domain and the newly identified hetero-interaction between the BICC1 and ANKS6 SAM domains. To test this we ran negGFP fusions of the wild-type and mutant proteins on a native gel and looked for changes in gel-shifts. The BICC1 ML mutants E898K and D911K are capable of interaction with wild-type ANKS6 but are no longer able to bind the ANKS6 mutants R823W or I817N (Fig. 4-11A). Similarly, the ANKS6 I817N mutation also results in a loss of interaction with the SAM domain of ANKS3 (Fig. 4-11B).

To understand the molecular basis whereby the ANKS6 I817N mutation causes a loss of interaction with the BICC1 and ANKS3 SAM domains we used circular dichroism to assess the folded state of the ANKS6 I817N SAM domain. Strikingly, the ANKS6 I817N mutation causes a dramatic loss of secondary structure, so much so that the CD spectra of ANKS6 I817N overlays almost perfectly with that of the denatured state of the wild-type ANKS6 SAM domain (Fig. 4-11C). Furthermore, the ANKS6 I817N mutant exhibited no cooperative unfolding transition during thermal denaturation and little change in signal overall (Fig. 4-11D). This supports that ANKS6 I817N is largely unfolded and not in an alternate structural conformation. These results are consistent with the retarded migration of the negGFP fusion of ANKS6 I817N on a native gel, which would be expected from an unfolded protein with a larger hydrodynamic radius (Fig. 4-11A).

Discussion

Mutations and/or altered expression of the SAM-domain containing proteins BICC1, ANKS3, and ANKS6 result in cystic kidney disease phenotypes in various animal models, suggesting that these proteins function in the same signaling and developmental pathways. We previously characterized the interaction between the SAM domains of ANKS3 and ANKS6 (Leettola, et al., 2014). Supporting this, these proteins have been detected as part of the same complex through co-immunoprecipitation experiments involving both tagged and endogenous proteins (Czarnecki et al., 2015; Hoff et al., 2013; Yakulov et al., 2015). BICC1 has been detected in complex with both ANKS3 and ANKS6, although the exact details of these associations are unknown (Stagner et al., 2009; Yakulov et al., 2015). BICC1 is also known to be polymeric and the SAM domain is required for full functional activity and correct intracellular localization, however, details regarding polymer formation by the SAM domain remain unknown (Knight et al., 2011; Kraus et al., 2012; Maisonneuve et al., 2009; Piazzon et al., 2012).

Herein we have characterized the polymeric form of the BICC1 SAM domain. By solving a high-resolution crystal structure we have shown that the BICC1 SAM domain is capable of interacting with itself to form a helical polymer with a physiologically relevant binding affinity. Like other known SAM domain polymers, the N- and C-termini extend away from the helical axis. Because the SAM domain of BICC1 lies at the C-terminus, the large N-terminal extension containing the RNA-binding KH domains is expected to decorate the outside of the polymer. We have further demonstrated that polymers of the BICC1 SAM domain have a natural propensity to associate side-by-side into sheets. BICC1 has been shown to form tubule and vesicle-like assemblies around the periphery of P-bodies (Maisonneuve et al., 2009), structures which are consistent with polymer and sheet formation mediated by the BICC1 SAM domain. Sheet formation has previously been observed with the SAM domains of Shank3 and DGKδ (Baron et al., 2006; Knight et al., 2010). In the case of Shank3, it was experimentally shown

that a large protein segment N-terminal to the SAM domain still permitted sheet formation of polymers. Since the BICC1 SAM polymer arrays examined by TEM do not appear closely packed, it is possible that the bulk of the protein N-terminal to the SAM domain may be accommodated to allow sheet formation by the full-length protein. The SAM domains of Shank3, DGKδ, and BICC1 all lie at the extreme C-terminal end of their respective proteins. Thus, when polymers assemble into sheets only additional bulk at one terminus needs to be accommodated. In the *bpk* mutant mouse model of PKD, a two base pair insertion is predicted to alter residues C-terminal to the SAM domain and extend the protein by an additional 129 amino acids (Bryda et al., 2003). It is currently unknown how this extension affects the function of BICC1 and leads to disease. However, one possible mechanism may be that packing of BICC1 SAM domain polymers into an ordered array may be impeded by the additional bulk of extra residues at the C-terminus that cannot be easily accommodated.

We also investigated the E932G mutation of BICC1 which has been shown to reduce BICC1's ability to inhibit Wnt signaling but does not affect localization of BICC1 to the periphery of P-bodies (Kraus et al., 2012). Examination of this mutation by native gel analysis and within the crystal structure show that it has no impact on the ability of BICC1 SAM to polymerize or form sheets and lies too far out of range from the ML and EH surfaces to make any intra-polymer contacts. Since the E932G mutation still permits P-body association but is functionally equivalent to a complete loss of the SAM domain, this suggests that the SAM domain of BICC1 serves an additional functional role besides simple polymer formation and may have other important binding partners. For example, the SAM domain of BICC1 is required for the transfer of AC6 mRNA to AGO2 and perhaps the C-terminal end of the SAM domain is involved in this activity (Piazzon et al., 2012). Future work is needed to parse apart the polymerization function of the BICC1 SAM domain from other possible activities.

BICC1 and ANKS3 have recently been shown to associate (Yakulov et al., 2015). We have determined that this interaction is mediated by the SAM domains of each protein. Interestingly, the SAM domains of BICC1 and ANKS3 are capable of using both their ML and EH surfaces to interact with each other. Furthermore, the binding affinities of ANKS3 SAM homo-polymerization (Leettola, et al., 2014), BICC1 SAM homo-polymerization, and the two possible ANKS3-BICC1 SAM interfaces all roughly the same. Thus it is possible that the ANKS3 and BICC1 SAM domains form an alternating co-polymer within the cell (Fig. 4-12C). A SAM domain co-polymer has been observed previously between the *Drosophila* SAM domains of the proteins sex-comb-on-midleg (Scm) and polyhomeotic (Ph) (Kim et al., 2005). However, measurement of the binding affinities for the various possible interfaces revealed a strong preference for a single junction between the free EH surface of a Scm SAM polymer and the free ML surface of a Ph SAM polymer, thus suggesting these proteins form a block co-polymer *in vivo*. Since the binding affinities of the possible interfaces between the BICC1 and ANKS3 SAM domains are roughly identical, we expect that these proteins may assemble into either a block co-polymer or the first observed alternating SAM domain co-polymer (Fig. 4-12 B,C).

The ability of the SAM domains of BICC1, ANKS3, and ANKS6 to interact with each other allows for the creation of a variety of scaffolds, each capable of organizing different protein complexes through their additional functional domains (Fig. 4-12). For example, a BICC1 SAM domain polymer may be capped by ANKS6 (Fig. 4-12A). In support of this, BICC1 is known to bind ANKS6 (Stagner et al., 2009; Yakulov et al., 2015). We have shown that this association is mediated by the high affinity binding of the SAM domains of each protein. Contrary to our results, co-immunoprecipitation (co-IP) experiments using truncated BICC1 and ANKS6 proteins showed that this association required the SAM domain of ANKS6 but not the SAM domain of BICC1 (Stagner et al., 2009). Given that BICC1, ANKS3, and ANKS6 may all be part of the

same complex, truncation of a single protein domain may not always result in an observed loss of association. We further demonstrated that the mutations R823W and I817N in ANKS6 which are associated with cystic kidney disease in rats and mice, respectively, cause a loss of interaction with the BICC1 and ANKS3 SAM domains. As determined by CD, the ANKS6 I817N mutation causes a dramatic loss of secondary structure, consistent with the protein existing in a largely unfolded state. In the crystal structure of ANKS6 (PDB:4NL9) Ile 817 (I48 in the crystal structure numbering) packs into the hydrophobic core of the protein. Mutation to the hydrophilic residue Asn would understandably disrupt packing and destabilize the structure of the ANKS6 SAM domain. Despite our finding that the I817N mutation results in a loss of hetero-interactions, recent work by Bakey et al. (Bakey et al., 2015) suggests the opposite. Using co-IP experiments, they demonstrate that ANKS3 still associates with the ANKS6 I817N mutant and that the association between BICC1 and ANKS6 appears to increase in the presence of the I817N mutation. These discrepancies allude to the complexity of the interaction network involving these proteins.

We suggest that the different hetero-interactions between the SAM domains of BICC1, ANKS3, and ANKS6 create a variety of scaffolds responsible for the organization of different protein complexes involved in cellular development and maintenance. It will be of specific interest to better understand the intracellular co-localization of these proteins. ANKS6 has been detected in association with ANKS3 as part of a protein complex found in the inversin compartment of primary cilia (Czarnecki et al., 2015). While BICC1 was not detected in the proteomic analysis of this co-IP, BICC1 has been shown to localize to the ciliary bulb (Mohieldin et al., 2015). Alternatively, the dynamic interplay between the SAM domains of these proteins may regulate polymerization and thereby control protein function. For example, the strong binding between the SAM domains of ANKS6 and BICC1 may allow the ANKS6 SAM domain to effectively compete with and de-polymerize BICC1. Since polymerization appears important for

BICC1 function, this interaction with ANKS6 may regulate BICC1 activity. Our structural and biochemical analysis of the SAM domain of BICC1 and its interactions with SAM domains of ANKS3 and ANKS6 provides a solid foundation for future *in vivo* experiments that address the importance of the associations between these SAM domains.

Methods

Cloning and Mutagenesis

negGFP fusions of the BICC1, ANKS3, and ANKS6 human SAM domains and DNA used in subsequent cloning were as described previously (Knight et al., 2011). All human BICC1-SAM constructs contained residues 870-939 [Uniprot: Q9H694]. Human ANKS3-SAM (residues 421-490) and human ANKS6-SAM (residues 771-840) were as formerly described (Leettola, et al., 2014). For large-scale protein expression and purification, hexahistidine small ubiquitin-like modifier (SUMO) tagged constructs were generated by cloning BICC1-SAM, ANKS3-SAM and ANKS6-SAM sequences into a pHis-SUMO vector (Senturia et al., 2010). The Quickchange method (Agilent) was used for site-directed mutagenesis. The mouse ANKS6 mutation I747N (Uniprot ID:E9PUD3) corresponds to I817N in the human ANKS6 protein. All plasmid sequences were verified by DNA sequencing (Genewiz).

negGFP Native Gel Binding Assays

All negGFP native gel assays were performed as previously described (Leettola, et al., 2014). Briefly, lysates of ARI814 cells expressing negGFP-human-SAM fusions were loaded on 20% RunBlue 12-well Native gels in RunBlue Native Run Buffer (Expedeon) and developed at 90V for 16 hours at 4^{o}C. Gels were visualized with a Bio-Rad Molecular Imager FX Pro-Plus and a Bio-Rad Molecular Imager PharosFX using an excitation wavelength of 488nm and an emission wavelength of 510nm. Where hetero-SAM interactions were tested, equal amounts of each protein (as determined by fluorescence) were mixed and allowed to equilibrate at 4^{o}C for 4 hours prior to gel loading.

Protein Expression and Purification

All purified proteins were expressed as pHis-SUMO constructs in Rosetta(DE3) pLysS cells (Novagen) and processed as described previously (Leettola, et al., 2014). In short, lysate supernatant from harvested cells, clarified by centrifugation at 13,200*g* for 20 minutes at 4^{o}C and supplemented with 10mM imidazole, was bound to Ni-NTA Superflow agarose (Qiagen). Resin was subsequently washed with 8 column volumes of lysis buffer (20mM $NaHPO_4$ pH 8.0, 0.5M NaCl) containing 15mM imidazole, 8 column volumes of lysis buffer containing 20mM imidazole and eluted with lysis buffer containing 250mM imidazole. After this initial purification all proteins except for ANKS6 I747N, BICC1 wt and BICC1 E932G were diluted to 5mg/mL and dialyzed into 20mM Tris pH 8.0, 0.5M NaCl. ANKS6 I747N was dialyzed into 20mM Tris pH 8.0, 0.3M NaCl. EDTA was added to 25mM to purified BICC1 wt and BICC1 E932G to prevent possible aggregation due to Ni^{2+} ions leached from the resin and these constructs were then dialyzed extensively against 20mM $NAHPO_4$ pH 8.0, 0.5M NaCl. All constructs were digested with the catalytic domain of SUMO protease 1 (ULP1) at a 50:1 protein:protease molar ratio for 16 hours at 4^{o}C and subsequently passed back over Ni-NTA resin (with the exception of BICC1 wt and BICC1 E932G, see below) to remove the His_6-SUMO tag (Malakhov et al., 2004). Digested proteins were present in the flow-thru of the subtractive Ni-NTA and were further purified as follows.

ANKS6 I747N was dialyzed against 10mM Tris pH 8.0, 0.05M NaCl and bound to a 5mL HiTrap Q column (GE) equilibrated in the same buffer. Protein was eluted by a shallow gradient of NaCl (0.1-0.5M) over 20 column volumes. Pooled fractions containing pure ANKS6 I747N were dialyzed against 20mM Tris pH 8.0, 0.15M NaCl and concentrated to 14.1mg/mL.

BICC1 R924E and K925E were each dialyzed against 20mM Tris pH 8.0, 50mM NaCl and bound to a 5mL HiTrap Q column (GE) equilibrated in 20mM Tris pH 8.0. Proteins were

eluted using a shallow gradient of NaCl (0-1.0M) over 20 column volumes (BICC1 R924E) or 26 column volumes (BICC1 K925E). BICC1 R924E bound very weakly and eluted immediately as soon as the concentration gradient began. BICC1 K925E eluted at approximately 0.15M NaCl. Fractions containing pure protein were pooled and dialyzed against 20mm Tris pH 8.0, 0.1M NaCl. BICC1 R924E was concentrated to 9-30 mg/mL. BICC1 K925E was concentrated to 7.5-33mg/mL.

BICC1 F920E was dialyzed against 20mM $NaHPO_4$ pH 6.0, 25mM NaCl and bound to a 5mL HiTrap SP HP column (GE). Protein eluted as a single peak at approximately 0.15M NaCl using a shallow gradient of NaCl in 20mM $NaHPO_4$ pH 6.0. Pure BICC1 F920E was dialyzed against 20mM $NaHPO_4$ pH 6.0, 0.1M NaCl and concentrated to 5.7-32mg/mL.

BICC1 D911K/F920E was dialyzed against 20mM $NAHPO_4$ pH 6.0, 50mM NaCl and BICC1 E898K, D911K and E898K/F920E were dialyzed against 20mM $NaHPO_4$ pH 6.0, 100mM NaCl. Proteins were bound to a 5mL HiTrap SP HP column (GE) equilibrated in 20mM $NaHPO_4$ pH 6.0. Proteins were eluted using shallow gradient of NaCl (0-1M) over 30 column volumes. BICC1 E898K and D911K eluted at approximately 300mM NaCl while BICC1 D911K/F920E and BICC1 E898K/F920E eluted at 150mM NaCl. Fractions containing pure protein were pooled and dialyzed against 20mM NaHPO4 pH 6.0, 0.3M NaCl. BICC1 E898K/F920E and D911K/F920E were at final concentrations of 4mg/mL, BICC1 D911K was at a final concentration of 4.9mg/mL, and BICC1 E898K was concentrated to 4-42.6mg/mL.

After cleavage with ULP1, BICC1 wt and E932G were passed over a 5mL HisTrap HP column (GE) charged with Ni^{2+}. Although the His_6-tagged SUMO fusion had been cleaved, BICC1 wt and E932G still bound the HisTrap column with weak affinity and were eluted using a shallow gradient of 0-20mM imidazole over 8 column volumes. EDTA was added to a final concentration of 15mM to fractions containing pure protein and proteins were then dialyzed

extensively against 20mM Tris pH 7.5, 0.5M NaCl. The final concentration of each protein was 0.9mg/mL.

Selenomethionine (SeMet) labeled BICC1 R924E was produced using an adapted version of a previously described protocol (Van Duyne et al., 1993). An overnight culture of Rosetta(DE3)pLysS cells transformed with pHis-SUMO BICC1 R924E was grown in M9 minimal media supplemented with kanamycin (30µg/mL) and chloramphenicol (34µg/mL). 20mL of overnight culture was used to inoculate each of three 1L flasks of M9 minimal media and cells were grown at 37°C with shaking until an OD_{600} of 0.55 was reached, at which point SeMet (60mg), Lys, Phe, and Thr (100mg each) and Ile, Leu, Val (50mg each) were added. Cells were incubated an additional 15 minutes at 37°C, then cooled to 18°C, induced with 1mM IPTG and expressed overnight at 18°C with shaking. Harvested cells were resuspended in 30mL of degassed lysis buffer (20mM Tris pH 7.5, 0.5M NaCl, 10mM βME) supplemented with 1mM PMSF and 20µg/mL DNase1 and lysed as described previously (Leettola, et al., 2014). SeMet labeled BICC1 R924E was purified as described above with the exception that Tris at pH 7.5 was used in place of $NaHPO_4$ and all buffers contained 10mM βME. After a final purification using a HiTrap Q column (GE), pure protein was dialyzed against 20mM Tris pH 8, 0.1M NaCl, 10mM βME and concentrated to 30.3mg/mL.

ANKS6 wt, ANKS6 R832W, ANKS3 I455E (previously numbered as ANKS3 I36E), and ANKS3 SAM F472E (previously numbered as F53E) were purified as described previously with the exception that ANKS3 F472E was purified on a 5mL HiTrap Q column (GE) equilibrated in 20mM Tris pH 8.0 (Leettola, et al., 2014).

Crystallization and structure determination

Initial crystallization trials were carried out at the Macromolecular Crystallization Core Technology Center at UCLA using commercially available screens and a Mosquito crystallization robot (TTP Labtech) to set up hanging drop vapor diffusion experiments.

Selenomethionine derivative crystals were grown by hanging drop vapor diffusion in 1µL drops prepared by a Mosquito crystallization robot (TTP Labtech). BICC1-SAM R924E SeMet at 30.3 mg/mL was mixed with well solution (100mM MES pH 7.0, 35% PEG-550 MME, 5mM $ZnSO_4$) in a 2:1 protein:reservoir ratio. Boulder-shaped crystals grew at room temperature over two and a half weeks and were cryoprotected using well solution supplemented with an additional 10% PEG-550 MME. Single crystals were mounted with CrystalCap HT Cryoloops (Hampton Research, Aliso Viejo). A multiple wavelength anomalous dispersion (MAD) data set was collected on a single crystal cryo-cooled to 100 K at the Advanced Photon Source (Argonne National Laboratory), APS-NECAT beamline 24-ID-C on a DECTRIS-PILATUS 6 M detector. Each data set contained 720 0.2° oscillation frames. Because the MAD data set went to only 2.4 Å resolution, a single wavelength high-resolution data set was collected on a separate crystal at the peak (0.97910 Å) and contained 600 0.2^{o} oscillation frames. Data sets were processed using XDS (Kabsch, 2010). Phasing was accomplished by MAD using the HKL2MAP interface and SHELX programs (Pape and Schneider, 2004; Sheldrick, 2010). The single-wavelength high resolution data set collected at the peak was used with the inflection and high remote data sets from the MAD experiment to extend the initial phases. Three selenomethionines were detected and used for phasing. Density modification and model building were performed using DM and BUCCANEER in the CCP4 suite (Winn et al., 2011). The structure was refined in PHENIX (Adams et al., 2010) and BUSTER (Bricogne et al., 2011)

using individual sites, individual atomic displacement parameters, non-crystallographic symmetry, and TLS parameterization with 3 TLS groups per chain. TLS groups were chosen by inspection of the structural model and each chain was divided into 3 groups: 1) helices 1-2; 2) helices 3-4; and 3) helix 5. After each round of refinement, the model was visually inspected and problem areas rebuilt in COOT (Emsley et al., 2010). Data collection and refinement statistics are reported in Table 4-1. The coordinates have been deposited in the PDB with accession code 4RQM.

Native crystals were grown by hanging drop vapor diffusion in 500nL drops prepared by a Mosquito crystallization robot (TTP Labtech). BICC1-SAM R924E at 29.8 mg/mL was mixed with well solution (100mM MES pH 6.5, 30% PEG-550 MME, 5mM $ZnSO_4$) in a 2:1 protein:reservoir ratio. Boulder-shaped crystals grew at room temperature over 10 days and were cryoprotected using well solution supplemented with an additional 10% PEG-550 MME. A high resolution data set was collected on a single native crystal cryo-cooled to 100 K at the Advanced Photon Source (Argonne National Laboratory), APS-NECAT beamline 24-ID-C on a DECTRIS-PILATUS 6 M detector. A data set containing 720 0.5° oscillation frames was collected from a single large crystal at a wavelength of 0.97950 Å and processed using XDS (Kabsch, 2010). Molecular replacement was performed using PHASER with chain C of the BICC1-SAM R924E SeMet structure as a search model and searching for 3 copies in the asymmetric unit (McCoy et al., 2007). The model was refined using PHENIX with TLS parameterization including individual sites, individual atomic displacement parameters, and non-crystallographic symmetry (Adams et al., 2010). Data collection and refinement statistics are reported in Table 4-1. The coordinates have been deposited in the PDB with accession code 4RQN.

Structure and sequence analysis

PROCHECK (Laskowski et al., 1993), ERRAT (Colovos and Yeates, 1993), and VERIFY3D (Bowie, et al., 1991) algorithms were used to validate the final structural models. Figures were prepared in PyMOL (Schrödinger, LLC). Surface electrostatics were calculated using the Adaptive Poisson-Boltzmann Solver (APBS) plugin in Pymol and all surfaces were contoured at ±1 kT/e (Baker et al., 2001). Analysis of protein-protein binding interfaces was accomplished using PISA (Krissinel and Henrick, 2007). ClustalW2 was used to perform multiple sequence alignments (Larkin et al., 2007).

Surface plasmon resonance

Experiments were performed at 21°C in 20 mM HEPES pH 7.5, 0.04% IGEPAL CA-630, 0.15M NaCl (+1mM DTT for ANKS3 samples only) using a Biacore T100 (GE). To determine the binding affinity of the BICC1-SAM/BICC1-SAM interface, BICC1-SAM R924E was immobilized on a Biacore CM5 chip (GE) via EDC/NHS crosslinking. BICC1-SAM mutants E898K and D911K at varying concentrations were passed over the chip and equilibrium binding levels were measured. As controls, varying concentrations of the BICC1 SAM mutants F920E and K925E (both with defective EH interfaces) and the BICC1 double mutants E898K/F920E and D911K/F920E (each with defective ML and EH interfaces) at 10μM were passed over the chip. To determine the binding affinity of the BICC1-SAM ML/ANKS3-SAM EH interface, ANKS3 SAM I455E at varying concentrations was passed over the BICC1-SAM R924E conjugated chip and equilibrium binding levels were measured. To determine the binding affinity of the BICC1-SAM EH/ANKS3-SAM ML interface, ANKS3-SAM F472E was immobilized on a Biacore CM5 chip (GE) via EDC/NHS crosslinking. BICC1-SAM mutants E898K and

D911K at varying concentrations were passed over the chip and equilibrium binding levels were measured. As a control the BICC1 double mutants E898K/F920E and D91K/F920E were passed over the ANKS3-SAM F472E conjugated chip. To determine the binding affinity of the BICC1-SAM EH/ANKS6-SAM ML interface, ANKS6-SAM was immobilized on a Biacore CM5 chip (GE) using EDC/NHS crosslinking. BICC1 SAM mutants E898K and D911K were passed over the chip at varying concentrations and equilibrium binding levels were measured. BICC1 double mutants E898K/F920E and D911K/F920E were passed over the chip as controls and no binding was observed. All data points were taken in triplicate and binding data was fit to a 1:1 steady-state model using Biacore T100 Evaluation software. Special handling of ANKS3-SAM I455E to maintain protein stability was as described previously (Leettola, et al., 2014).

Circular Dichroism

Spectra were collected on protein samples at 0.2 mg/ml in 10mM Tris pH 8.0, 75mM NaCl at 25^{o}C using a 1mm path length quartz cuvette on a JASCO J-715 circular dichroism spectrophotometer equipped with a Peltier temperature control. The Selcon, Neural Network, and Contin algorithms available in SoftSec (Softwood Inc.) were used to analyze spectra for secondary structure content. Thermal denaturations were performed by monitoring the change in CD signal at 222nm across a temperature range of 25-80^{o}C, with ramping of 1^{o}C per minute.

Electron Microscopy

Protein samples were applied to carbon/formvar coated copper grids (Ted Pella, catalog number 01754-F) made hydrophilic by glow discharge immediately before use and allowed to bind for several minutes. Grids were rinsed with nanopure water and negatively stained with 1% uranyl acetate. Samples were analysed on a T12 transmission electron microscope (FEI Tecnai) operating at 120kV. Images were recorded using a Gatan 2k x 2k CCD camera and processed using ImageJ (NIH). FFT analysis was performed using ImageJ (NIH). Images

presented in figures 4-6D and 4-6E are mirror images of the original and have been flipped 180^{o}. This was done to present the correct handedness of the protein sheet, which randomly laid down on the grid in two different orientations.

Analytical Size-Exclusion Chromatography

For analysis of the BICC1-EH/ANKS3-ML binding 250uL of protein in 20mM HEPES pH 7.5, 0.15M NaCl, 1mM DTT was loaded on an Superdex 200 10/300 GL column (GE) equilibrated in the same buffer and run at 0.4mL/min. Proteins were loaded as follows: BICC1 E898K was loaded at 2mg/mL; ANKS3 F472E and BICC1 D911K were loaded at 4mg/mL; the 1:1 molar ratio mix of BICC1 E898K + ANKS3 F472E was loaded at 4mg/mL; the 1:1 molar ratio mix of BICC1 D911K + ANKS3 F472E was loaded at 8mg/mL. For the analysis of the BICC1-ANKS6 binding, 250uL of protein in 20mM HEPES pH 7.5, 0.15M NaCl was loaded on a Superdex 200 10/300 GL column (GE) equilibrated in the same buffer and run at 0.4mL/min. ANKS6 wt, ANKS6 R823W, ANKS6 I747N, BICC1 E898K were each loaded at 4mg/mL. The 1:1 molar ratio mixes of BICC1 E898K + ANKS6 wt, BICC1 E898K +ANSK6 R823W and BICC1 E898K + ANKS6 I747N were each loaded at 8mg/mL. Elution was monitored by absorbance at 280nm.

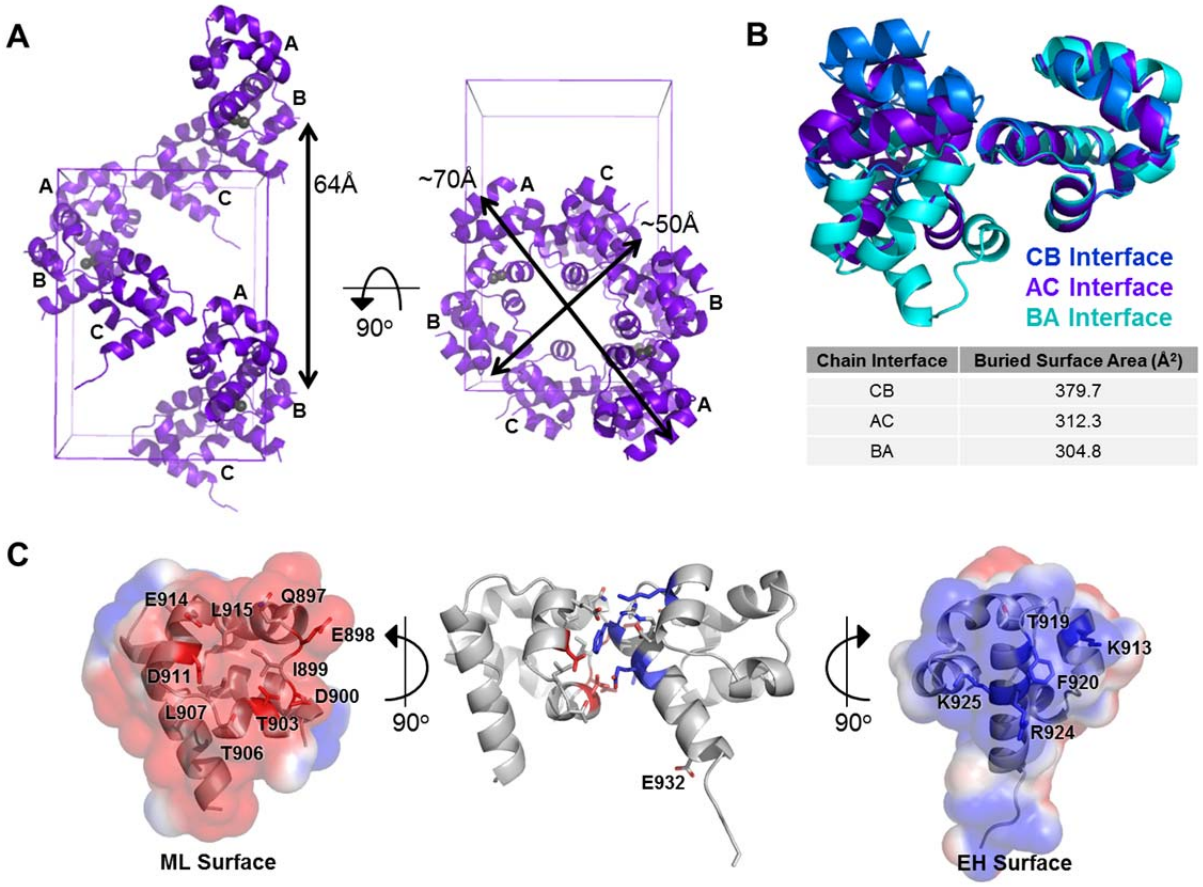
SEC-MALS

Protein samples at 10mg/mL were analysed by SEC-MALS. 50µL of BICC1 F920E + ANKS3 I455E and 100µL of BICC1 R924E + ANKS3 I455E were loaded onto a WTC-030S5 analytical size-exclusion column (Wyatt Technology Co.) equilibrated in 20mM Tris pH 8.0, 0.15M NaCl, 2mM βME using an AKTA purifier (GE) and run at 0.7mL/min on a miniDAWN TREOS (Wyatt Technology Co.). Protein peaks were analysed for calculated molecular weight and monodispersity using ASTRA software (Wyatt Technology Co.)

Divalent metal assay

Purified BICC1 SAM wild-type and E932G at 120μM and suspended in 20mM Tris pH 7.5, 0.5M NaCl had the metal salts $MnCl_2$, $ZnSO_4$, $NiSO_4$, $MgCl_2$, $CaCl_2$, and $CoCl_2$ or buffer alone added to a final concentration of 10mM. Proteins were allowed to equilibrate for 3 hours before imaging. To further test the zinc dependence of precipitate formation, $ZnSO_4$ was added a final concentration of 1.25mM to purified BICC1 SAM wild-type and E932G at 120μM suspended in 20mM Tris pH 7.5, 0.5M NaCl. Protein was allowed to equilibrate at room temperature for 4 hours before binding to EM grids.

Figures and Tables



Chain Interface	Buried Surface Area (Å²)
CB	379.7
AC	312.3
BA	304.8

Figure 4-1. Structure of the SeMet BICC1 SAM R924E mutant.

A) BICC1 SAMs form a polymer with 6 subunits per helical repeat. The polymer measures 64Å per helical repeat and is not symmetrical, measuring approximately 50Å by 70Å wide. Each asymmetric unit contains 3 subunits and 2 zinc ions, each with half occupancy and modeled as grey spheres. A single polymer is shown in the context of the unit cell. The chain ID of each subunit is labeled. B) BICC1 SAMs associate in three different orientations within the polymer. Because there are 3 chains per asymmetric unit, there are 3 possible interfaces, each labeled using the chains in the crystal structure that form the respective homodimer. Each homodimer has been aligned using the subunit containing the EH surface, allowing the varying angles of ML surface interaction to be apparent. Surface area buried at each interface was calculated using the PISA server. C) Residues critical for BICC1 SAM interaction are highlighted on the AC interface. Residues we identified in our negGFP native gel screen (Fig. 4-8B) as being crucial to the ML and EH surfaces are colored red and blue, respectively. Surface electrostatics calculated using APBS in Pymol and contoured at ±1kT/e show the charge complementarity between the negatively charged ML surface and positively charged EH surface. The dimer structure (middle) contains the R924E mutation. In the EH surface (left) residue 924 is modeled as the wild-type Arg to more clearly show the positive charge of the EH surface.

Table 4-1. Crystallographic data collection and refinement statistics

	BICC1 R924E SeMet			BICC1 R924E Native
PDB Accession #	4RQM			4RQN
Data collection				
Location	APS 24-ID-C			APS 24-ID-C
Space group	$P2_12_12_1$			$P2_12_12_1$
	Peak	High remote	Inflection	
Wavelength (Å)	0.97910	0.97140	0.97930	0.97950
Cell dimensions				
a, *b*, *c* (Å)	46.81, 64.43, 70.64	46.26, 64.03, 70.05	46.25, 64.04, 70.07	41.68, 59.22, 68.66
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	1.75	2.40	2.40	2.00
R_{svm}	0.068 (.550)	0.094 (.784)	0.096 (.866)	0.097 (0.505)
$I/\sigma I$	10.29 (2.05)	8.13 (1.93)	7.78 (1.74)	14.94 (4.38)
$CC_{1/2}$	99.5 (67.7)	99.3 (77.1)	99.4 (75.0)	99.6 (95.2)
Completeness (%)	93.0 (92.3)	97.2 (94.6)	97.3 (93.7)	98.5 (98.7)
Redundancy	2.27 (2.15)	2.78 (2.76)	2.77 (2.65)	12.65 (13.40)
Phasing Statistics				
Number of sites	3			
Mean figure of merit				
MAD/after density	0.602/0.778			
MapCC (SHELXE)	0.822			
CC (%)	65.46			
Refinement				
Resolution (Å)	1.75			2.00
No. reflections	22189			11813
R_{work}/ R_{free}	0.2297/0.2672			0.2124/0.2652
No. atoms				
Protein	1461			1397
Water	40			16
Zinc	2			1
B-factor (Å^2)	43.0			54.1
R.m.s deviations				
Bond lengths (Å)	0.008			0.008
Bond angles (º)	1.12			1.08

Highest resolution shell is shown in parenthesis.
$R_{sym} = \sum |I - < I > |/\sum < I >$, where I is the observed intensity and <I> is the average intensity from observations of symmetry-related reflections. $CC_{1/2}$ = correlation coefficient between two halves of the data (Karplus and Diederichs, 2012). $R_{work} = \sum |F_{obs} - F_{calc}|/\sum F_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. R_{free} is calculated for a set of reflections (10%) that were not included in atomic refinement.

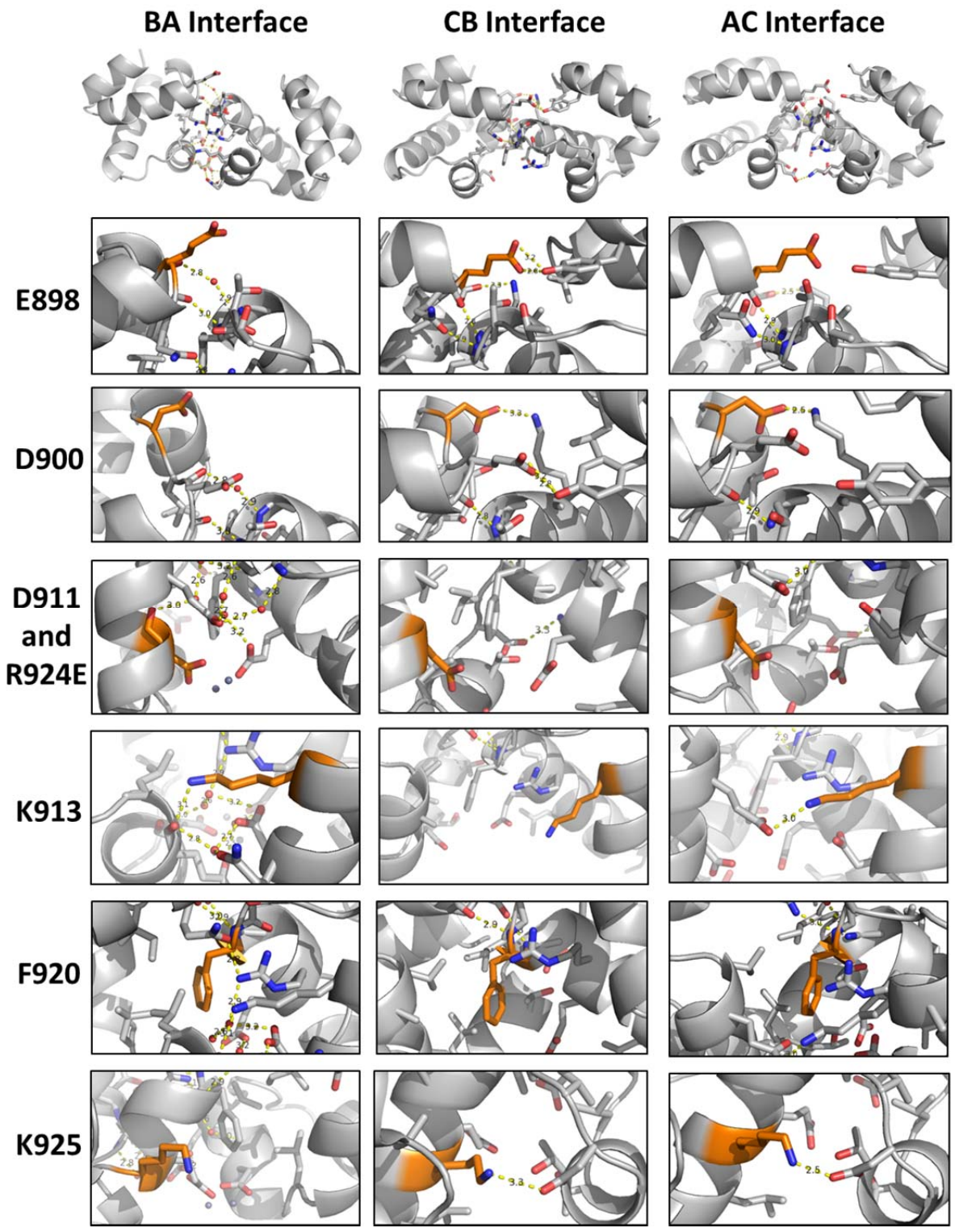
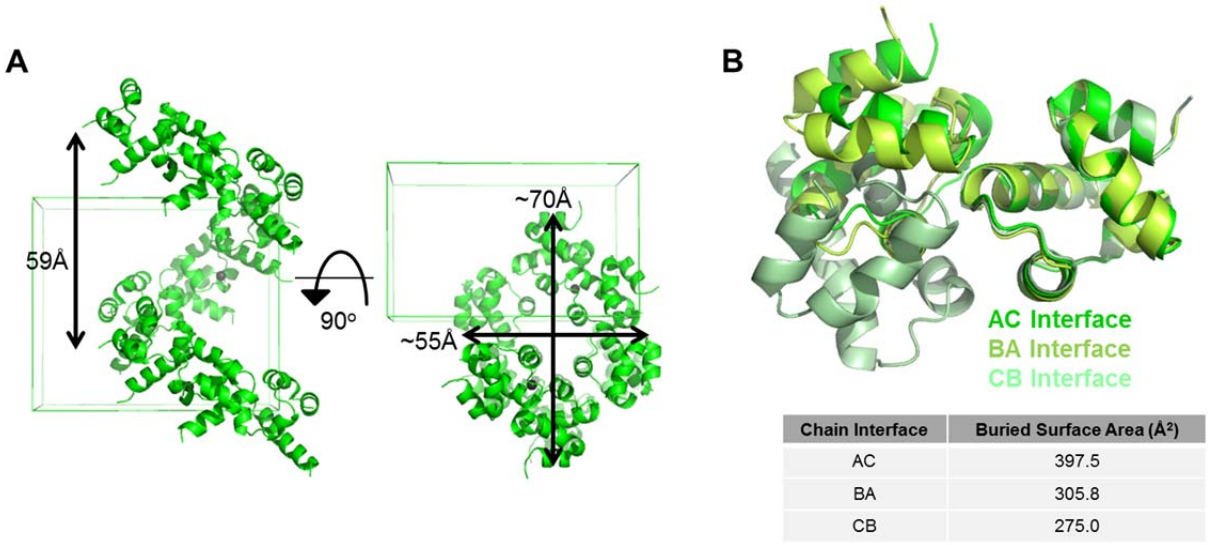


Figure 4-2. Detailed views of the SeMet BICC1 R924E interfaces.

Note the BA interface is highly solvated and contains 2 zinc ions, each with half occupancy, modeled as grey spheres. Below each interface are close up views of residues (highlighted in

orange) identified in the negGFP native gel screen as being critical for polymer formation. Hydrogen bond and ionic interactions of each residue are shown, along with distance measurements in Å. T903 on the ML surface, not shown above, was identified as being critical for polymer formation and packs against the backbone atoms of the neighboring EH helix in each interface. Not all residues identified in the negGFP native gel screen participate in interactions across each possible interface.



Chain Interface	Buried Surface Area (Å²)
AC	397.5
BA	305.8
CB	275.0

Figure 4-3. Structure of the native BICC1 SAM R924E mutant.

A) In the native structure, BICC1 SAMs form a polymer with 6 subunits per helical repeat. However, the polymer is slightly more compact than the SeMet structure and measures 59Å per helical repeat and approximately 55Å by 70Å wide. Each asymmetric unit contains 3 subunits and 1 zinc ion modeled as a grey sphere. B) BICC1 SAMs in the native structure associate in three different orientations within the polymer. Each interface has been labeled using the chains in the crystal structure that form the respective homodimer and chains are aligned by the subunit containing the EH surface, revealing the varying angles of ML surface interaction. Surface area buried at each interface was calculated using the PISA server.

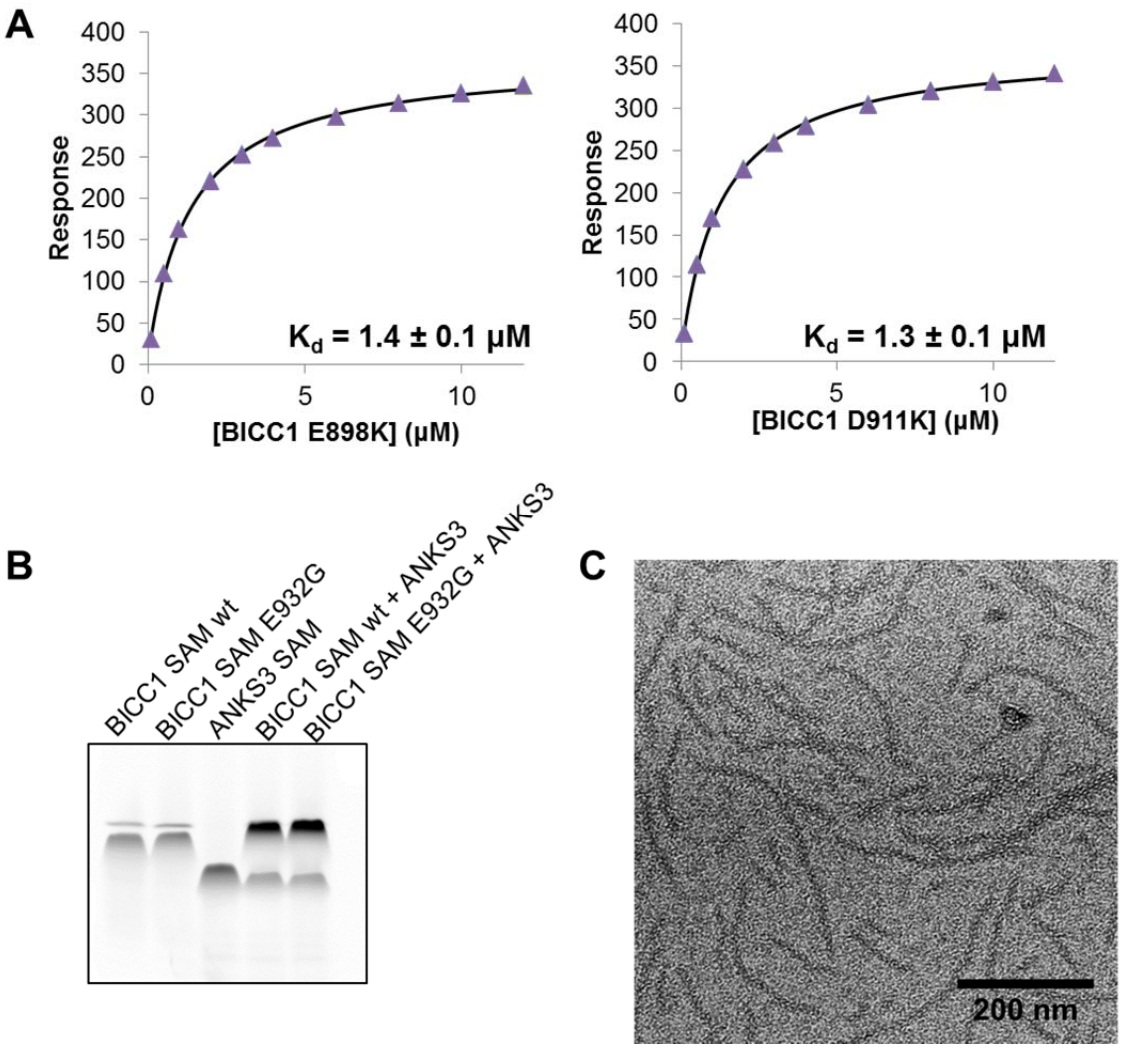
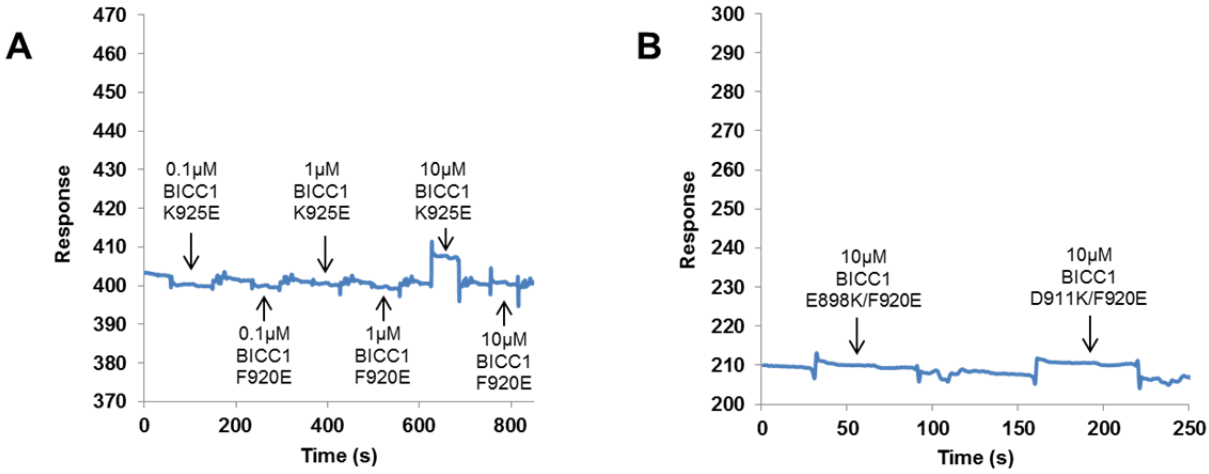


Figure 4-4. Characterization of the BICC1 SAM polymer.

A) The binding affinity of the native BICC1 SAM interface was measured using SPR (average K_d = 1.35 ± 0.1 μM). The BICC1 EH mutant R924E was immobilized on the chip surface and equilibrium binding measurements using the BICC1 ML mutants E898K (left) or D911K (right) were measured in triplicate and fit to a 1:1 steady-state model. Error bars are smaller than the data points. B) negGFP fusions of BICC1 SAM wt and E932G run on a native gel. The E932G mutation does not affect the mobility of BICC1 SAM, indicating no impact on polymerization ability. When wild-type BICC1 SAM is mixed in a 1:1 ratio with ANKS3 SAM, the distinct upshift and decrease in intensity of the ANKS3 band indicates an interaction between these SAM domains. The E932G mutation does not affect this interaction. C) Wild-type BICC1 SAM at 120μM was examined by transmission electron microscopy and found to form strands of short polymers joined side-by-side. Individual polymers consist of only a few helical repeats and measure approximately 7nm wide and 6nm per helical repeat.



Protein	Relative Response
0.1μM BICC1 K925E	-2.8
0.1μM BICC1 F920E	-1.8
1μM BICC1 K925E	-0.8
1μM BICC1 F920E	-1.4
10μM BICC1 K925E	6.8
10μM BICC1 F920E	0.0

Protein	Relative Response
10μM BICC1 E898K/F920E	-0.2
10μM BICC1 D911K/F920E	2.7

Figure 4-5. SPR negative controls for the BICC1 R924E conjugated Biacore chip.

A) Raw Biacore data and relative response units of the BICC1 EH mutants F920E and K925E flowed at various concentrations over a BICC1 R924E conjugated Biacore chip. Because the F920E and K925E mutants have disrupted EH surfaces they do not bind the native ML surface or any alternate surface of the protein on the chip. B) Raw Biacore data and relative response units of the BICC1 double mutants E898K/F920E and D911K/F920E flowed at 10μM over a BICC1 R924E conjugated Biacore chip. With mutations at both the ML and EH surfaces the double mutants are incapable of interaction, again affirming that the measured binding affinity is between a native ML and EH surface and not an alternate binding site.

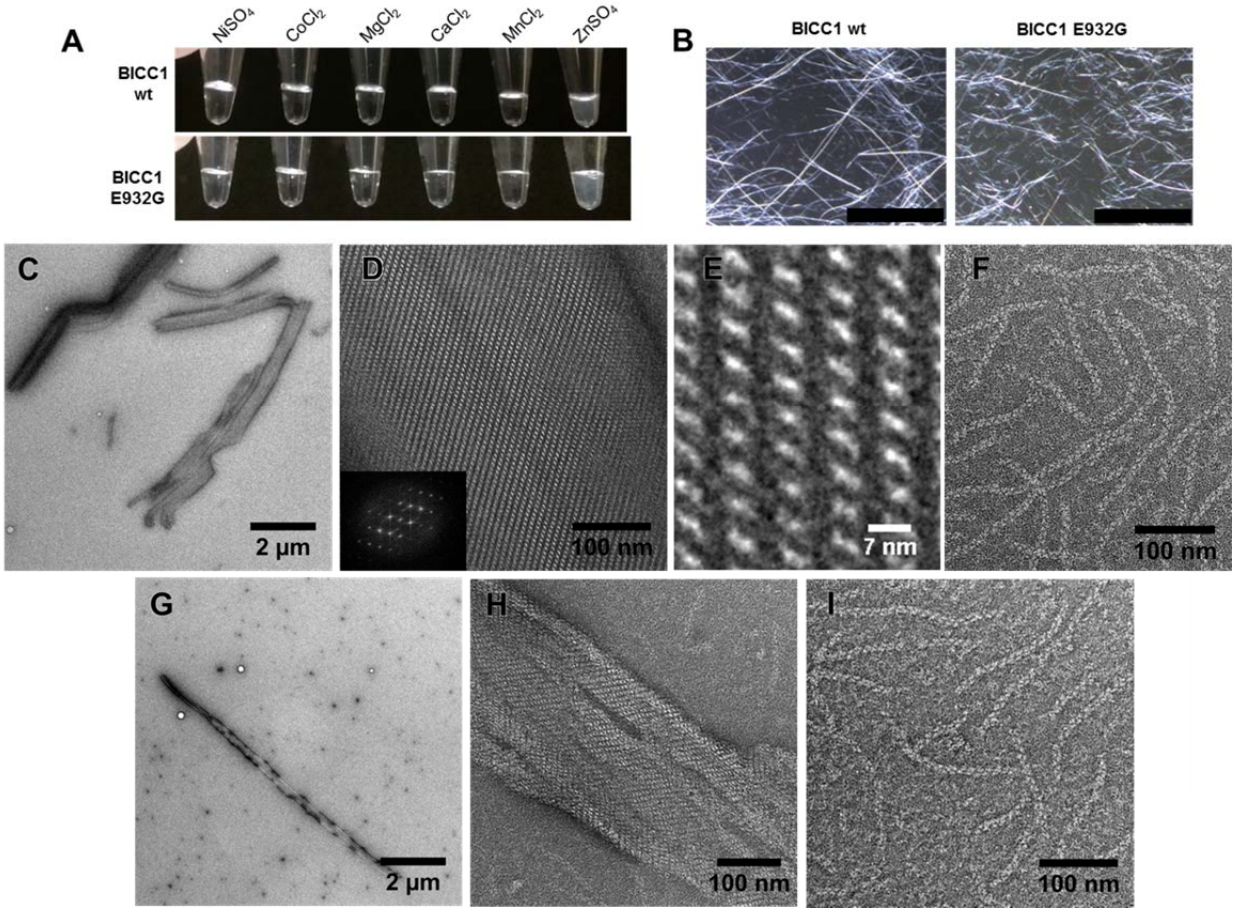


Figure 4-6. BICC1 SAMs form large polymer arrays.

A) BICC1 SAM wild-type and E932G both form precipitate in the presence of $ZnSO_4$ but not other divalent metals. Proteins at 120µM had different divalent metal salts added to a final concentration of 10mM and were allowed to equilibrate at room temperature for 3 hours. B) BICC1 SAM wt and E932G both form birefringent fibers in the presence of $ZnSO_4$. Precipitate from (A) was examined by light microscope and showed fiber-like birefringent strands, indicative of a high level of ordering. Scale bars are 500 µm. C-E) BICC1 SAM wt at 120µM in the presence of a 12.5-fold molar excess of $ZnSO_4$ examined by transmission electron microscopy. In the presence of $ZnSO_4$ large sheets of polymers are present (C). Magnification of these structures show individual polymers assembled side-by-side (D). The diffraction pattern revealed by FFT analysis (inset of D) shows the 2D crystalline nature of the polymer array. Individual polymers measure approximately 7nm wide and 6nm per helical repeat (E). Strands of short polymers assembled side-by-side are also present (F). G-I) BICC1 SAM E932G at 120 µM in the presence of a 12.5-fold molar excess of $ZnSO_4$. The E932G mutant is equally capable of forming large sheets of polymers (G) that are composed of individual polymers approximately 7nm wide and 6nm per helical repeat aligned side-by-side (H). BICC1 SAM E932G forms strands of short polymers associated side-by-side (I).

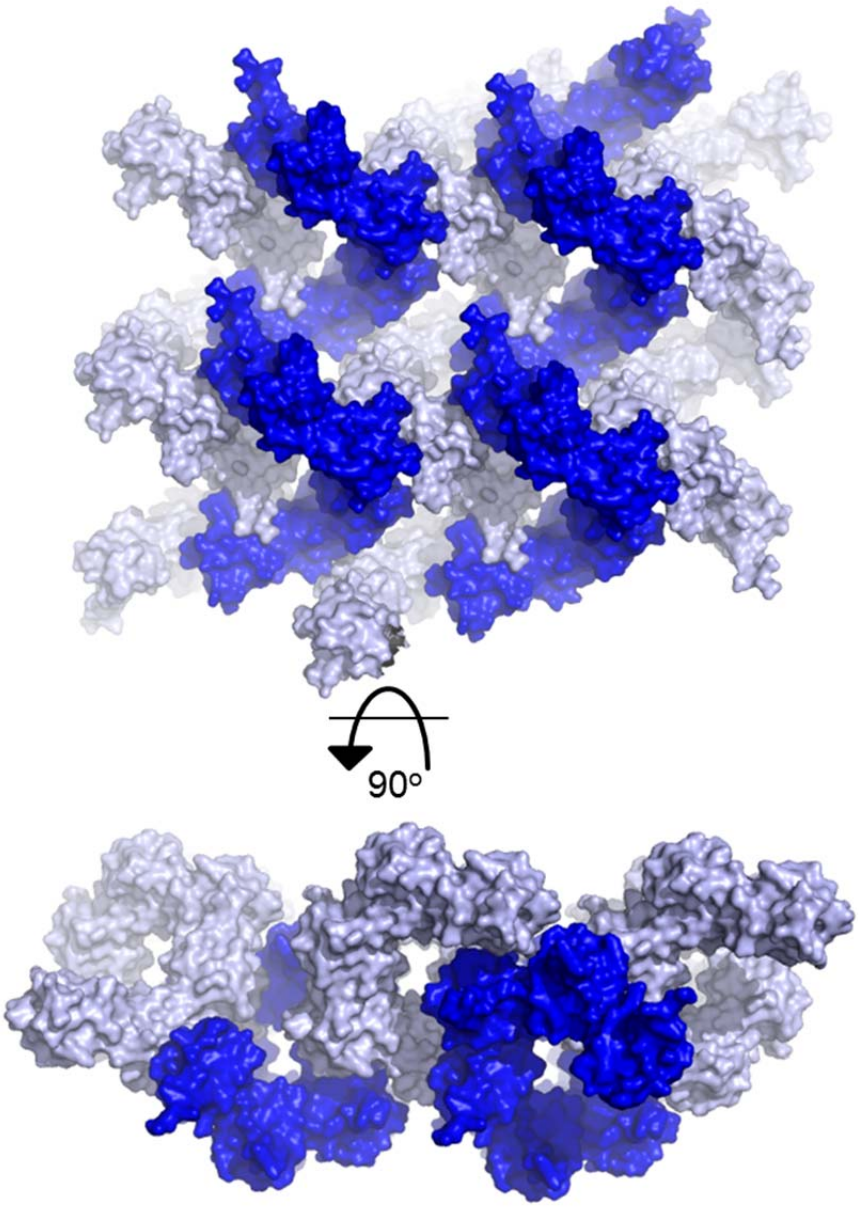


Figure 4-7. Polymer packing within the SeMet BICC1 SAM R924E crystal structure.

Individual polymers are shown as space-filled models. Polymers within the crystal structure form a left-handed helix and are closely interdigitated.

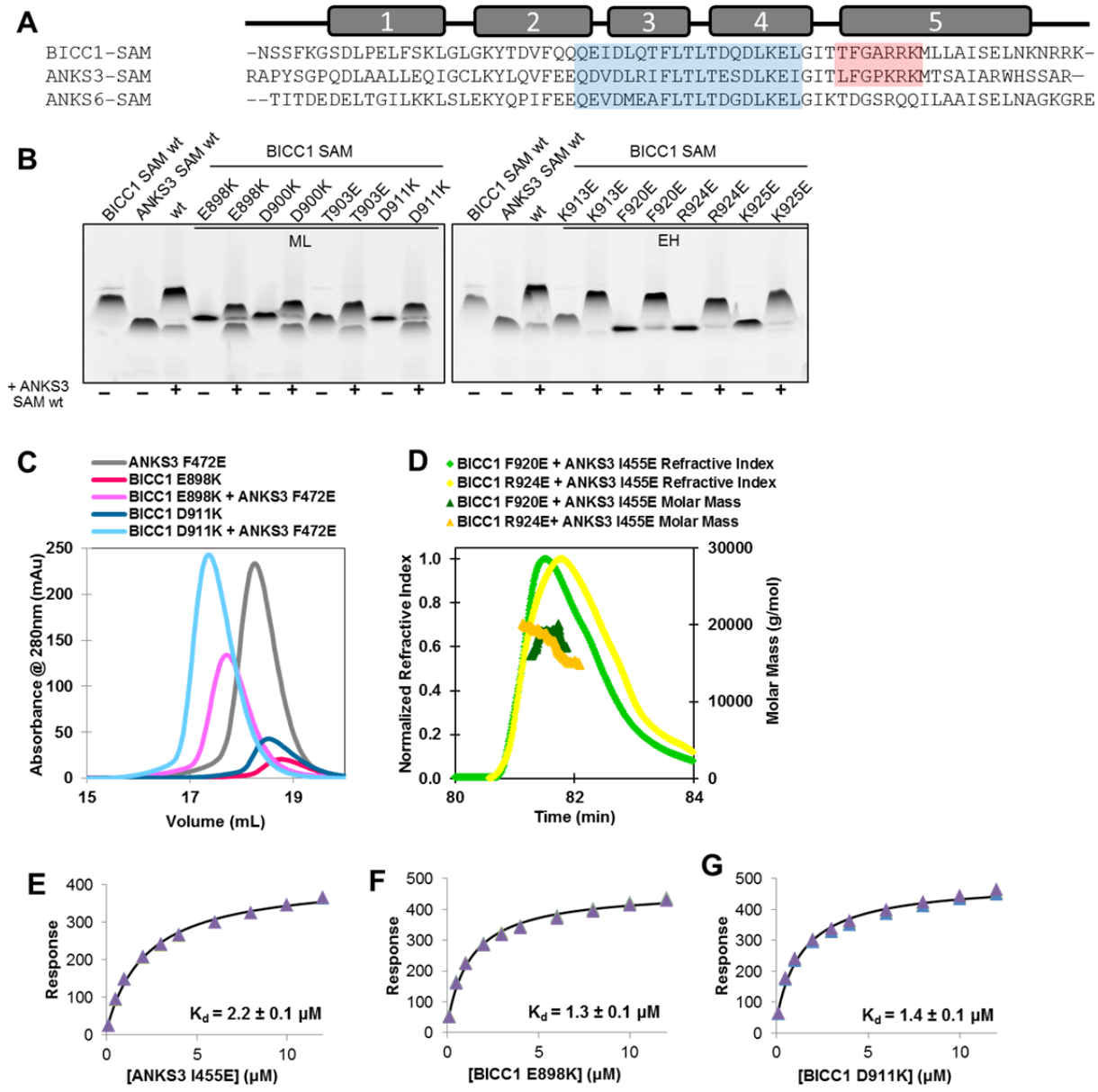
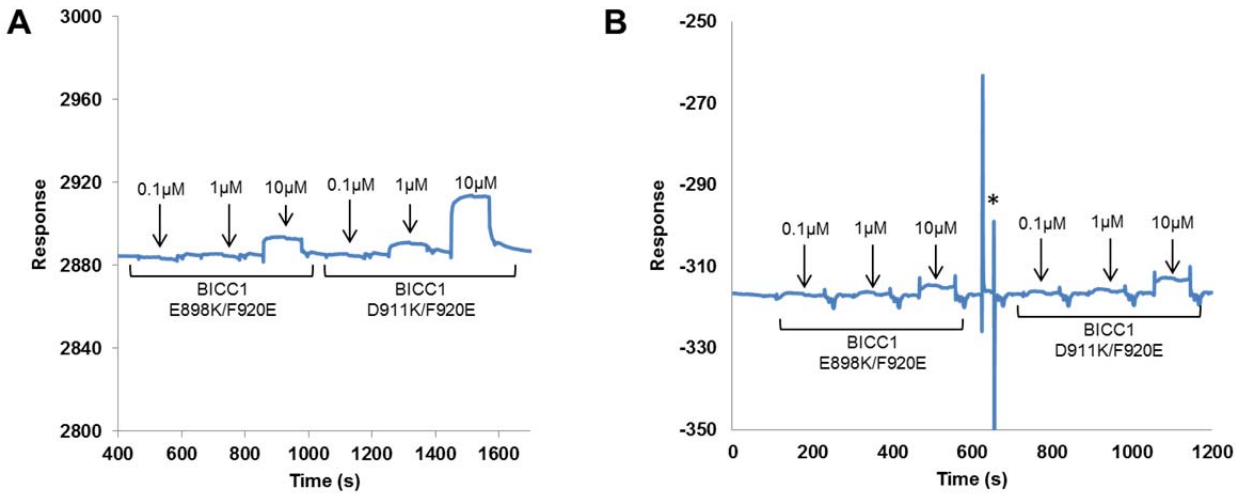


Figure 4-8. The BICC1 and ANKS3 SAM domains associate.

A) Multiple sequence alignments of the SAM domains of BICC1 (residues 877-938), ANKS3 (residues 429-490), and ANKS6 (residues 771-840). A secondary structure schematic is shown above and indicates the position and numbering of α-helices within the SAM structures. The ML surface of ANKS3 (residues 449-467) is 63% identical and 89% similar to the ML surface of BICC1 (residues 897-915). The EH surface of ANKS3 (residues 471-477) is 67% identical and 83% similar to the EH surface of BICC1 (residues 919-925). Residues encompassing the ML surface of each protein are shaded in blue. Residues that form the EH surface of BICC1 and ANKS3 are shaded in red. B) negGFP native gel analysis of the BICC1 and ANKS3 SAM interaction. When mixed in a 1:1 molar ratio, BICC1 SAM wt and ANKS3 SAM interact, as

indicated by a gel upshift and loss of the ANKS3 SAM band. ML and EH surface mutants of BICC1 cause a loss of polymerization, indicated by faster gel migration of the mutants and a discrete banding pattern. When ML and EH surface mutants are mixed in 1:1 ratios with ANKS3 each produces a gel shift, indicating an interaction. C) Analytical size exclusion chromatography analysis of the BICC1-EH/ANKS3-ML interface. ANKS3 F472E, BICC1 E898K, and BICC1 D911K each run as monomers. 1:1 molar ratio mixes of the BICC1 ML mutants with ANKS3 F472E results in earlier elution, consistent with formation of a heterodimer. D) SEC-MALS analysis of the BICC1-ML/ANKS3-EH interface. When the BICC1 EH mutants F920E and R924E are mixed in 1:1 molar ratios with ANKS3 I455E each migrates as a heterodimer (measured MW of BICC1 F920E/ANKS3 I455E = 17.4 ± 0.6 kDa, measured MW of BICC1 R924E/ANKS3 I455E = 18.4 ± 0.1 kDa, calculated MWs= 16.1 kDa). E) The binding affinity of the BICC1-ML/ANKS3-EH surface was measured by SPR (K_d = 2.2 ± 0.1 μM). BICC1 SAM R924E was immobilized on a Biacore chip and equilibrium binding measurements using ANKS3 SAM I455E were measured in triplicate and fit to a 1:1 steady-state model. Error bars are smaller than the data points. F-G) The binding affinity of the BICC1-EH/ANKS3-ML surface was measured by SPR (average K_d = 1.35 ± 0.1 μM). ANKS3 F472E was immobilized on a Biacore chip and equilibrium binding measurements using the BICC1 ML mutants E898K (F) and D911K (G) were measured in triplicate and fit to a 1:1 steady-state model. Error bars are smaller than the data points.



Protein	Relative Response
0.1μM BICC1 E898K/F920E	-1.7
1μM BICC1 E898K/F920E	-1.1
10μM BICC1 E898K/F920E	7.4
0.1μM BICC1 D911K/F920E	-1.3
1μM BICC1 D911K/F920E	4.7
10μM BICC1 D911K/F920E	26.6

Protein	Relative Response
0.1μM BICC1 E898K/F920E	5.4
1μM BICC1 E898K/F920E	7.6
10μM BICC1 E898K/F920E	27.3
0.1μM BICC1 D911K/F920E	3.2
1μM BICC1 D911K/F920E	5.1
10μM BICC1 D911K/F920E	25.4

Figure 4-9. SPR negative controls for the ANKS3 F472E and ANKS6 conjugated Biacore chips.

A) Raw Biacore data and relative response units of the BICC1 double mutants E898K/F920E and D911K/F920E flowed at various concentrations over an ANKS3 F472E conjugated Biacore chip. The E898K/F920E and D911K/F920E double mutants have disrupted ML and EH surfaces and do not bind the protein on the chip. This demonstrates that ANKS3 and BICC1 associate via their native ML and EH surfaces, respectively, and not an alternate binding site. B) Raw Biacore data and relative response units of the BICC1 double mutants E898K/F920E and D911K/F920E flowed at varying concentrations over an ANKS6 conjugated Biacore chip. With mutations at both the ML and EH surfaces the BICC1 double mutants exhibit dramatically reduced binding, demonstrating that ANKS6 and BICC1 associate via their ML and EH surfaces, respectively, as opposed to an alternate binding site. * Binding buffer + 0.5M NaCl was flowed over the chip in a regeneration step.

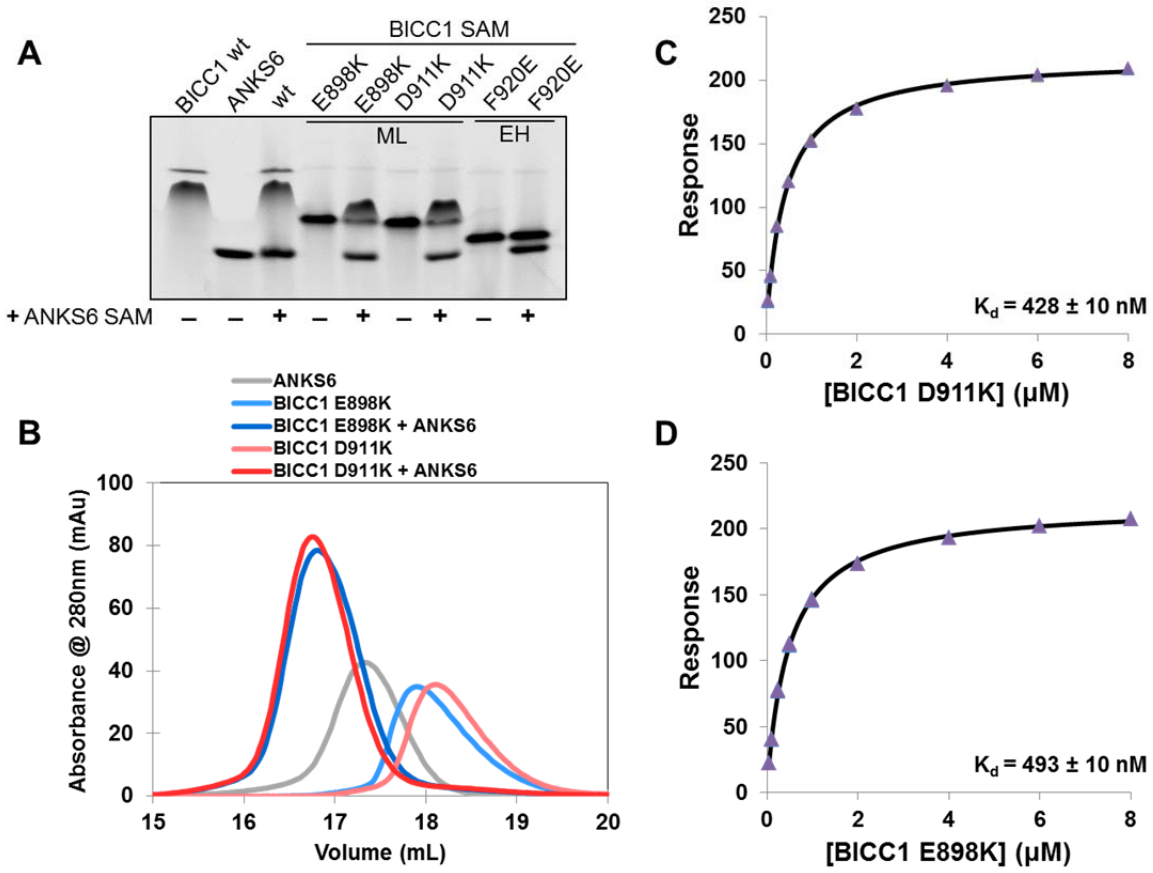


Figure 4-10. The BICC1 and ANKS6 SAM domains associate.

A) negGFP native gel analysis of the BICC1 and ANKS6 SAM domains. The negGFP fusion of BICC1 SAM wt does not appear to interact with ANKS6 as there are no apparent gel shifts. However, when the BICC1 ML mutants E898K and D911K are mixed in 1:1 molar ratios with ANKS6 SAM the presence of a gel shift indicates an interaction. The BICC1 EH mutant F920E is not capable of interaction with ANSK6. B) Analytical size exclusion chromatography shows the BICC1 SAM EH surface and ANKS6 SAM interact. When the BICC1 ML mutants E898K and D911K are mixed in 1:1 molar ratios with ANKS6, the protein elutes earlier than each of the monomeric counterparts on their own, indicating the presence of a heterodimer. C-D). The binding affinity of the BICC1-EH/ANKS6-ML surface was measured (average K_d = 460.5 ± 10nM). ANKS6 SAM was immobilized on a Biacore chip and equilibrium binding using the BICC1 ML mutants D911K (C) and E898K (D) was measured in triplicate and fit to a 1:1 steady-state model. Error bars are smaller than the data points.

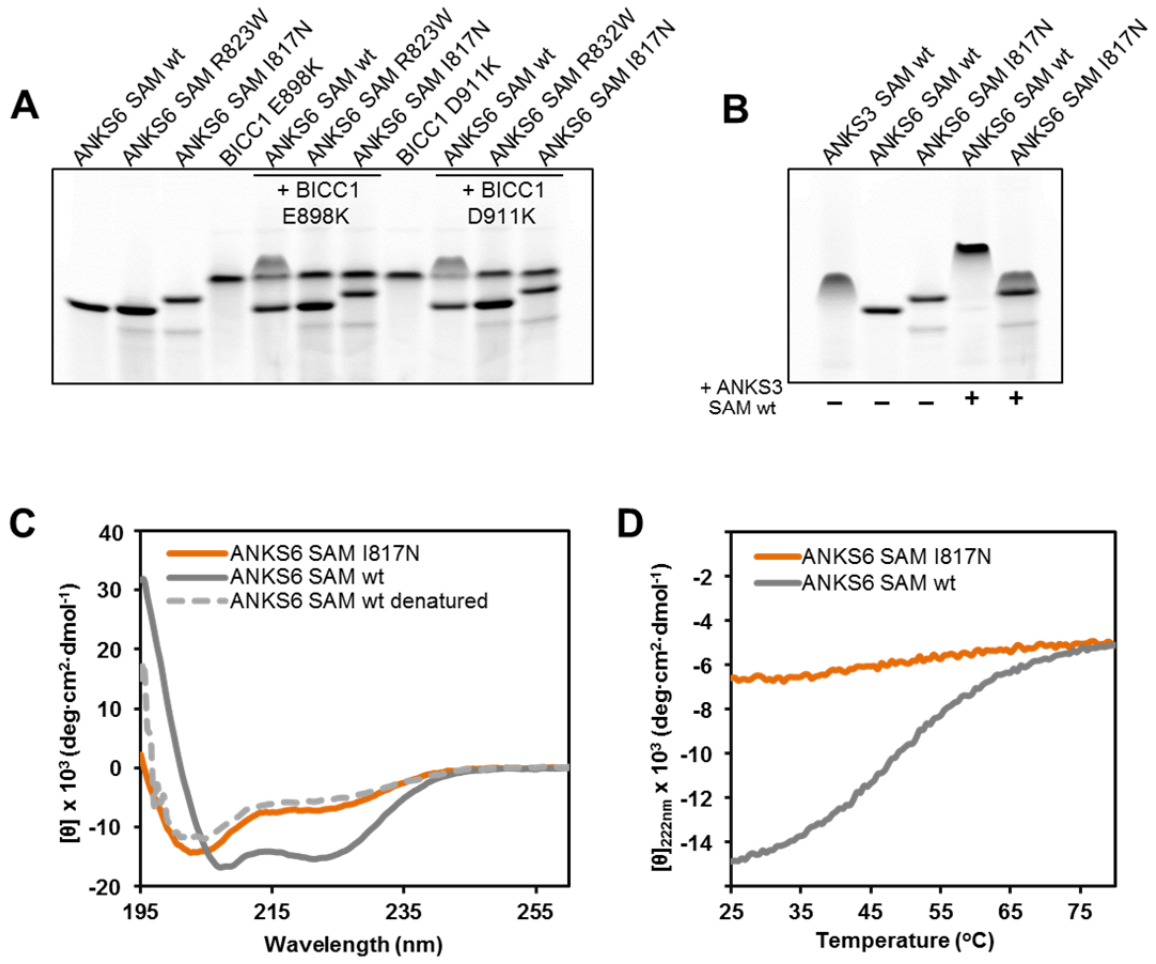


Figure 4-11. Characterizing the I817N mutation of ANKS6.

A) negGFP native gel analysis of the interaction between the SAM domains of BICC1 and ANKS6. The BICC1 ML mutants E898K and D911K interact with wild-type ANKS6 but not with the ANKS6 mutants R823W or I817N. B) The I817N mutation in ANKS6 causes a loss of interaction with the ANKS3 SAM domain, as assessed by negGFP native gel analysis. C) The CD spectra of ANKS6 SAM I817N, measured at 25°C, overlays well with the CD spectra of denatured wild-type ANKS6 SAM, measured at 80°C. Compared to the wild-type ANKS6 SAM domain, the I817N mutation causes a 16-30% loss of α-helicity accompanied by an increase in random coil and β-sheet, as calculated using the Selcon, Neural Network, and Contin algorithms. E) Thermal denaturation of ANKS6 SAM I817N compared to the wild-type ANKS6 SAM domain. The lack of an unfolding transition shows that ANKS6 SAM I817N has not adopted an alternate structure and exists in a largely unfolded state.

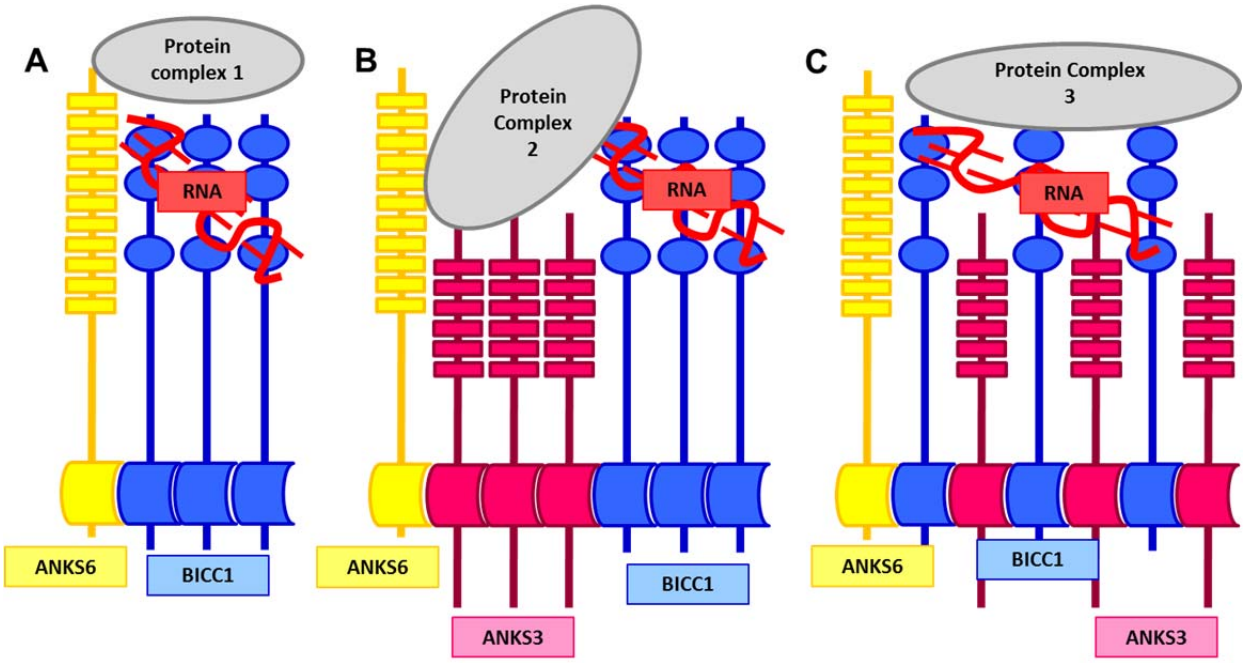


Figure 4-12. Possible scaffolding complexes mediated by the BICC1 SAM domain.

A) A BICC1 polymer may be capped on the free EH surface by an interaction with the ML surface of the SAM domain of ANKS6. The ANK repeats of ANKS6 and the RNA-binding KH domains of BICC1 may associate with an additional complex of proteins. B) The SAM domains of BICC1 and ANKS3 may associate as a block co-polymer. The SAM domain of ANSK6 can associate with the free EH surface of either ANKS3 (shown here) or BICC1 and cap this complex. The ANK repeats of ANKS6 and ANKS3 and the RNA-binding KH domains of BICC1 may organize a complex of proteins. It is also possible that only ANKS3 and BICC1 form a scaffold, without ANKS6 present. C) The SAM domains of BICC1 and ANKS3 may associate as an alternating co-polymer, which may or may not be capped on the free EH surface end by ANKS6. The alternating pattern of ANKS3 and BICC1 may allow binding of a unique complex of proteins.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). *PHENIX* : a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Aviv, T., Lin, Z., Rendl, L., M., Sicheri, F., and Smibert, C.A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. Nat. Struct. Biol. *10*, 614–621.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. *98*, 10037–10041.

Bakey, Z., Bihoreau, M.-T., Piedagnel, R., Delestré, L., Arnould, C., de Villiers, A. d'Hotman, Devuyst, O., Hoffmann, S., Ronco, P., Gauguier, D., et al. (2015). The SAM domain of ANKS6 has different interacting partners and mutations can induce different cystic phenotypes. Kidney Int.

Baron, M.K., Boeckers, T.M., Vaida, B., Faham, S., Gingery, M., Sawaya, M.R., Salyer, D., Gundelfinger, E.D., and Bowie, J.U. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. Science *311*, 531–535.

Barrera, F.N. (2003). Binding of the C-terminal Sterile Motif (SAM) Domain of Human p73 to Lipid Membranes. J. Biol. Chem. *278*, 46878–46885.

Bouvrette, D.J., Sittaramane, V., Heidel, J.R., Chandrasekhar, A., and Bryda, E.C. (2010). Knockdown of bicaudal C in zebrafish (Danio rerio) causes cystic kidneys: a nonmammalian model of polycystic kidney disease. Comp. Med. *60*, 96.

Bowie,, J.U., Lüthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. Science *253*, 164–170.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., et al. (2011). BUSTER version 2.10.0.

Brown, J.H., Bihoreau, M.-T., Hoffmann, S., Kranzlin, B., Tychinskaya, I., Obermuller, N., Podlich, D., Boehn, S.N., Kaisaki, P.J., Megel, N., et al. (2005). Missense Mutation in Sterile Motif of Novel Protein SamCystin is Associated with Polycystic Kidney Disease in (cy/+) Rat. J. Am. Soc. Nephrol. *16*, 3517–3526.

Bryda, E.C., Flaherty, L., Cogswell, C., Price, S.J., Hou, X., and Guay-Woodford, L.M. (2003). Positional cloning of jcpk/bpk locus of the mouse. Mamm. Genome *14*, 242–249.

Chaki, M., Airik, R., Ghosh, A.K., Giles, R.H., Chen, R., Slaats, G.G., Wang, H., Hurd, T.W., Zhou, W., Cluckey, A., et al. (2012). Exome Capture Reveals ZNF423 and CEP164 Mutations, Linking Renal Ciliopathies to DNA Damage Response Signaling. Cell *150*, 533–548.

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. *2*, 1511–1519.

Czarnecki, P.G., Gabriel, G.C., Manning, D.K., Sergeev, M., Lemke, K., Klena, N.T., Liu, X., Chen, Y., Li, Y., San Agustin, J.T., et al. (2015). ANKS6 is the critical activator of NEK8 kinase in embryonic situs determination and organ patterning. Nat. Commun. *6*, 6023.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic Structures of the Human Immunophilin FKBP-12 Complexes with FK506 and Rapamycin. J. Mol. Biol. *229*, 105–124.

Eckmann, C.R., Kraemer, B., Wickens, M., and Kimble, J. (2002). GLD-3, a bicaudal-C homolog that inhibits FBF to control germline sex determination in C. elegans. Dev. Cell *3*, 697–710.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of *Coot*. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Failler, M., Gee, H.Y., Krug, P., Joo, K., Halbritter, J., Belkacem, L., Filhol, E., Porath, J.D., Braun, D.A., Schueler, M., et al. (2014). Mutations of CEP83 Cause Infantile Nephronophthisis and Intellectual Disability. Am. J. Hum. Genet. *94*, 905–914.

Grimshaw, S.J., Mott, H.R., Stott, K.M., Nielsen, P.R., Evetts, K.A., Hopkins, L.J., Nietlispach, D., and Owen, D. (2003). Structure of the Sterile Motif (SAM) Domain of the Saccharomyces cerevisiae Mitogen-activated Protein Kinase Pathway-modulating Protein STE50 and Analysis of Its Interaction with the STE11 SAM. J. Biol. Chem. *279*, 2192–2201.

Gundelfinger, E.D., Boeckers, T.M., Baron, M.K., and Bowie, J.U. (2006). A role for zinc in postsynaptic density asSAMbly and plasticity? Trends Biochem. Sci. *31*, 366–373.

Halbritter, J., Bizet, A.A., Schmidts, M., Porath, J.D., Braun, D.A., Gee, H.Y., McInerney-Leo, A.M., Krug, P., Filhol, E., Davis, E.E., et al. (2013). Defects in the IFT-B Component IFT172 Cause Jeune and Mainzer-Saldino Syndromes in Humans. Am. J. Hum. Genet. *93*, 915–925.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Hildebrandt, F., Attanasio, M., and Otto, E. (2009). Nephronophthisis: Disease Mechanisms of a Ciliopathy. J. Am. Soc. Nephrol. *20*, 23–35.

Hildebrandt, F., Benzing, T., and Katsanis, N. (2011). Ciliopathies. N. Engl. J. Med. *364*, 1533–1543.

Hoff, S., Halbritter, J., Epting, D., Frank, V., Nguyen, T.-M.T., van Reeuwijk, J., Boehlke, C., Schell, C., Yasunaga, T., Helmstädter, M., et al. (2013). ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3. Nat. Genet. *45*, 951–956.

Kabsch, W. (2010). *XDS*. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Karplus, P.A., and Diederichs, K. (2012). Linking Crystallographic Model and Data Quality. Science *336*, 1030–1033.

Kim, C.A., and Bowie, J.U. (2003). SAM domains: uniform structure, diversity of function. Trends Biochem. Sci. *28*, 625–628.

Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S., and Bowie, J.U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J. *20*, 4173–4182.

Kim, C.A., Gingery, M., Pilpa, R.M., and Bowie, J.U. (2002). The SAM domain of polyhomeotic forms a helical polymer. Nat. Struct. Biol. *9*, 453–457.

Kim, C.A., Sawaya, M.R., Cascio, D., Kim, W., and Bowie, J.U. (2005). Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. J. Biol. Chem. *280*, 27769–27775.

Knight, M.J., Joubert, M.K., Plotkowski, M.L., Kropat, J., Gingery, M., Sakane, F., Merchant, S.S., and Bowie, J.U. (2010). Zinc Binding Drives Sheet Formation by the SAM Domain of Diacylglycerol Kinase δ. Biochemistry *49*, 9667–9676.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Kraus, M.R.-C., Clauin, S., Pfister, Y., Di Maïo, M., Ulinski, T., Constam, D., Bellanné-Chantelot, C., and Grapin-Botton, A. (2012). Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. Hum. Mutat. *33*, 86–90.

Krissinel, E., and Henrick, K. (2007). Inference of Macromolecular Assemblies from Crystalline State. J. Mol. Biol. *372*, 774–797.

Kwan, J.J., Warner, N., Pawson, T., and Donaldson, L.W. (2004). The Solution Structure of the S.cerevisiae Ste11 MAPKKK SAM Domain and its Partnership with Ste50. J. Mol. Biol. *342*, 681–693.

Lancaster, M.A., and Gleeson, J.G. (2010). Cystic kidney disease: the role of Wnt signaling. Trends Mol. Med. *16*, 349–360.

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.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Lee, H.J., Hota, P.K., Chugha, P., Guo, H., Miao, H., Zhang, L., Kim, S.-J., Stetzik, L., Wang, B.-C., and Buck, M. (2012). NMR Structure of a Heterodimeric SAM:SAM Complex: Characterization and Manipulation of EphA2 Binding Reveal New Cellular Functions of SHIP2. Structure *20*, 41–55.

Leettola,, C.N., Knight, M.J., Cascio, D., Hoffmann, S., and Bowie, J.U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. BMC Struct. Biol. *14*.

Leone, M., Cellitti, J., and Pellecchia, M. (2009). The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. BMC Struct. Biol. *9*, 59.

Mahone, M., Saffman, E.E., and Lasko, P.F. (1995). Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA binding motif of FMR1. EMBO J. *14*, 2043.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Blum, M., and Constam, D.B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. Development *136*, 3019–3030.

Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D., and Butt, T.R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics *5*, 75–86.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). *Phaser* crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Mohieldin, A.M., Haymour, H.S., Lo, S.T., AbouAlaiwi, W.A., Atkinson, K.F., Ward, C.J., Gao, M., Wessely, O., and Nauli, S.M. (2015). Protein composition and movements of membrane swellings associated with primary cilia. Cell. Mol. Life Sci. *72*, 2415–2429.

Mohler, J., and Wieschaus, E.F. (1986). Dominant maternal-effect mutations of Drosophila melanogaster causing the production of double-abdomen embryos. Genetics *112*, 803–822.

Neudecker, S., Walz, R., Menon, K., Maier, E., Bihoreau, M.-T., Obermüller, N., Kränzlin, B., Gretz, N., and Hoffmann, S.C. (2010). Transgenic Overexpression of Anks6(p.R823W) Causes Polycystic Kidney Disease in Rats. Am. J. Pathol. *177*, 3000–3009.

Pape, T., and Schneider, T.R. (2004). *HKL2MAP* : a graphical user interface for macromolecular phasing with *SHELX* programs. J. Appl. Crystallogr. *37*, 843–844.

Piazzon, N., Maisonneuve, C., Guilleret, I., Rotman, S., and Constam, D.B. (2012). Bicc1 links the regulation of cAMP signaling in polycystic kidneys to microRNA-induced gene silencing. J. Mol. Cell Biol. *4*, 398–408.

Di Pietro, S.M., Cascio, D., Feliciano, D., Bowie, J.U., and Payne, G.S. (2010). Regulation of clathrin adaptor function in endocytosis: novel role for the SAM domain. EMBO J. *29*, 1033–1044.

Qiao, F., and Bowie, J.U. (2005). The many faces of SAM. Sci. Signal. *2005*, re7.

Qiao, F., Song, H., Kim, C.A., Sawaya, M.R., Hunter, J.B., Gingery, M., Rebay, I., Courey, A.J., and Bowie, J.U. (2004). Derepression by depolymerization: structural insights into the regulation of Yan by Mae. Cell *118*, 163–173.

Qiao, F., Harada, B., Song, H., Whitelegge, J., Courey, A.J., and Bowie, J.U. (2005). Mae inhibits Pointed-P2 transcriptional activity by blocking its MAPK docking site. EMBO J. *25*, 70–79.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Schueler, M., Braun, D.A., Chandrasekar, G., Gee, H.Y., Klasson, T.D., Halbritter, J., Bieder, A., Porath, J.D., Airik, R., Zhou, W., et al. (2015). DCDC2 Mutations Cause a Renal-Hepatic Ciliopathy by Disrupting Wnt Signaling. Am. J. Hum. Genet. *96*, 81–92.

Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.7.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. *19*, 1354–1365.

Sheldrick, G.M. (2010). Experimental phasing with *SHELXC* / *D* / *E* : combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. *66*, 479–485.

Stafford, R.L., Hinde, E., Knight, M.J., Pennella, M.A., Ear, J., Digman, M.A., Gratton, E., and Bowie, J.U. (2011). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK. Structure *19*, 1826–1836.

Stagner, E.E., Bouvrette, D.J., Cheng, J., and Bryda, E.C. (2009). The polycystic kidney disease-related proteins Bicc1 and SamCystin interact. Biochem. Biophys. Res. Commun. *383*, 16–21.

The GPN Study Group, Halbritter, J., Porath, J.D., Diaz, K.A., Braun, D.A., Kohl, S., Chaki, M., Allen, S.J., Soliman, N.A., Hildebrandt, F., et al. (2013). Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. Hum. Genet. *132*, 865–884.

Tran, U., Pickney, L.M., Özpolat, B.D., and Wessely, O. (2007). Xenopus Bicaudal-C is required for the differentiation of the amphibian pronephros. Dev. Biol. *307*, 152–164.

Tran, U., Zakin, L., Schweickert, A., Agrawal, R., Doger, R., Blum, M., De Robertis, E.M., and Wessely, O. (2010). The RNA-binding protein bicaudal C regulates polycystin 2 in the kidney by antagonizing miR-17 activity. Development *137*, 1107–1116.

Wilson, P.D. (2004). Polycystic Kidney Disease. N. Engl. J. Med. *350*, 151–164.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the *CCP* 4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Yakulov, T.A., Yasunaga, T., Ramachandran, H., Engel, C., Müller, B., Hoff, S., Dengjel, J., Lienkamp, S.S., and Walz, G. (2015). Anks3 interacts with nephronophthisis proteins and is required for normal renal development. Kidney Int.

Zhang, H., Xu, Q., Krajewski, S., Krajewska, M., Xie, Z., Fuess, S., Kitada, S., Godzik, A., and Reed, J.C. (2000). BAR: An apoptosis regulator at the intersection of caspases and Bcl-2 family proteins. Sci. Signal. *97*, 2597.

Chapter 5

A helical region flanking the ANKS3 SAM domain affects homo- and heterotypic SAM domain interactions

Abstract

Stability and function of SAM domains can be impacted by N- and C-terminal flanking residues. The ANKS3 SAM domain, implicated as being involved in the cystic kidney diseases PKD and NPHP, has a coiled coil region immediately adjacent to its C-terminus. Here we demonstrate that this coiled coil region appears to inhibit ANKS3 SAM polymer formation and prevent a hetero-interaction with the EH surface of the BICC1 SAM domain. This coiled coil extension does not affect binding of the ANKS3 SAM domain EH surface with the ML surfaces of the ANKS6 and BICC1 SAM domains. This suggests that the coiled coil selectively interferes with the ANKS3 SAM domain ML surface and indicates the need for future research into how this coiled coil modulates the function of the ANKS3 SAM domain.

Introduction

Nephronophthisis (NPHP) is a cystic kidney disease that is the most common genetic cause of juvenile end stage renal failure (Hildebrandt et al., 2009). Currently mutations in 19 different genes are associated with NPHP, however the underlying causative mutations are still unknown in roughly 50% of cases (Chaki et al., 2012; Failler et al., 2014; Halbritter et al., 2013; Hoff et al., 2013; Schueler et al., 2015; The GPN Study Group et al., 2013). The *anks3* gene, which produces a protein called ankyrin repeat and SAM-domain containing protein 3 (ANKS3) has recently been implicated in NPHP and the related ciliopathy known as polycystic kidney disease (PKD) (Leettola, et al., 2014; Yakulov et al., 2015).

Human ANKS3 has 6 N-terminal ankyrin repeats and a C-terminal sterile alpha motif (SAM) domain. SAM domains are common protein-protein interaction domains capable of binding each other in a homotypic manner to form helical polymers (Baron et al., 2006; Harada et al., 2008; Kim and Bowie, 2003; Kim et al., 2001; Qiao and Bowie, 2005; Stafford et al., 2011) or in a heterotypic manner with the SAM domains of other proteins to form heterodimers and co-polymers (Kim, 2005; Kwan et al., 2004; Leone et al., 2009; Qiao et al., 2004; Ramachander, 2002). Prior characterization of the SAM domain of ANKS3 (see Chapter 3) showed that it forms helical polymers by sequential association of the mid-loop (ML) and end-helix (EH) surfaces. When expressed in *Xenopus* embryos, GFP-tagged ANKS3 localizes to large cytosolic aggregates which are thought to be ANKS3 polymers observed *in vivo* (Yakulov et al., 2015). ANKS3 has also been shown to associate with the SAM-domain containing proteins ANKS6 and BICC1 (Czarnecki et al., 2015; Hoff et al., 2013; Yakulov et al., 2015). Our analysis of these interactions revealed that these proteins associate through their SAM domains. The EH surface of ANKS3 SAM binds the ML surface of ANKS6 SAM with high affinity to create a heterodimer (Chapter 3). Both the ML and EH surfaces of ANKS3 SAM bind the ML and EH

surfaces of BICC1 SAM with similar affinities, suggesting ANKS3 SAM and BICC1 SAM may form a co-polymer (Chapter 4).

ANKS3 contains a predicted coiled coil immediately C-terminal to the SAM domain. Protein regions flanking SAM domains have been shown to affect SAM domain polymerization. For example, the *Drosophila* protein polyhomeotic (Ph) contains an unstructured linker sequence N-terminal to the SAM domain which hinders Ph-SAM polymerization (Robinson et al., 2012a). The human homolog, PHC3, also contains an unstructured region N-terminal to the SAM domain but in contrast, this linker positively influences polymer formation (Robinson et al., 2012b). Unlike the PHC3 linker region, the linker of Ph-SAM has been observed to directly interact with the SAM domain (Robinson et al., 2012a). However, the exact mechanisms whereby these unstructured linkers influence polymerization remain unknown.

SAM domain flanking regions can also affect SAM domain stability. In human liprin-α2 a short helical region flanking the N-terminus of the first SAM domain forms extensive hydrophobic contacts with the SAM domain and is required to produce stably folded protein *in vitro* (Wei et al., 2011). In a newly identified motif found in the N-terminal region of kinase suppressor of Ras-1, a coiled coil (CC) domain is found immediately N-terminal to a SAM domain (Koveal et al., 2012). This motif, termed CC-SAM forms a single modular domain where the SAM domain is likely needed for stability of the coiled coil, which was shown to bind membranes. Flanking regions may also impact SAM domain heterotypic interactions. The yeast protein Ste4 contains a leucine zipper N-terminal to the SAM domain. Trimerization mediated by the leucine zipper allows the Ste4 SAM domain to bind the Byr2 SAM domain with a more than 2,000-fold higher binding affinity (Ramachander, 2002).

Since flanking regions may affect SAM domain function we chose to investigate whether the coiled coil C-terminal to the ANKS3 SAM domain had any effect on ANKS3 SAM polymer

formation or heterotypic interactions. We find that inclusion of the coiled coil region causes ANKS3 SAM to become monomeric and lose interaction with the EH surface of BICC1 SAM. However, the coiled coil does not appear to impact interactions with the ML surfaces of the ANKS6 and BICC1 SAM domains. This suggests the coiled coil extension C-terminal to the ANKS3 SAM domain is interfering with the ML surface of ANKS3 SAM.

Results

A helical flanking region inhibits ANKS3 SAM polymerization

The SAM domain of human ANKS3 (residues 421-490 of the full length protein) is flanked at the C-terminal end by a region (residues 501-526) predicted to be a coiled coil based on sequence analysis (Fig. 5-1A). To determine the likelihood that this segment truly is a coiled coil, we analyzed the C-terminal flanking sequence in both the COILS and MARCOIL servers (Fig. 5-1B). Both algorithms strongly predict the presence of a coiled coil.

To determine if the helical region C-terminal to the SAM domain of ANKS3 impacted the ability of the SAM domain to polymerize, we generated an ANKS3 SAM-CC construct, which encompasses the SAM domain and the predicted coiled coil and spans residues 421-529 of the full length protein. When this construct was expressed as a negGFP fusion and run on a native gel, it exhibited increased mobility and a discrete band character when compared to ANKS3 SAM without the coiled coil extension (Fig. 5-2). This negGFP native gel assay has been used previously to assess the polymeric character of SAM domains (Knight et al., 2011). negGFP fusions of monomeric SAM domains are observed to run as discrete bands while polymeric SAM domains run with a smeared character and reduced mobility that is roughly inversely proportional to polymer size. While ANKS3 SAM had migration consistent with a weak polymer, addition of the coiled coil in ANKS3 SAM-CC resulted in migration similar to the monomeric SAM domain of ANKS6, suggesting that the coiled coil inhibits ANKS3 SAM polymerization.

The EH surface of ANKS3 SAM is known to bind the ML surface of ANKS6 SAM with high affinity (Leettola, et al., 2014). We used the negGFP native gel assay to discern whether the C-terminal coiled coil had any impact on this interaction. As observed previously, a 1:1 mix of the negGFP fusions of ANKS3 SAM and ANKS6 SAM produces a gel shift, indicative of interaction. ANKS3 SAM-CC produces the same gel shift when mixed with ANKS6, indicating

that the coiled coil extension does not appear to affect this hetero-interaction (Fig. 5-2). The ANKS3 SAM-CC + ANKS6 SAM band does migrate slightly faster than the ANKS3 SAM + ANKS6 SAM band; this is likely due to the increased negative charge of the protein due to the coiled coil addition which is rich in negatively charged residues (pI = 4.48). Since the coiled coil extension does not affect the interaction with ANKS6 SAM, this suggests that it does not interfere with the ANKS3 SAM EH surface.

The coiled coil extension affects the interaction between the ANKS3 and BICC1 SAM domains

When the coiled coil extension is appended to the ANKS3 SAM domain it causes an apparent loss of interaction with the BICC1 SAM domain, as indicated by the loss of a gel shift when these negGFP fusion constructs are mixed in a 1:1 ratio (Fig. 5-3). Both the ML and EH surfaces of the ANKS3 SAM domain bind the BICC1 SAM domain through the formation of two distinct interfaces with similar binding affinities. To determine which binding interface was affected by the coiled coil extension we mixed BICC1 SAM ML and EH surface point mutants with ANKS3 SAM-CC. The BICC1 SAM EH mutants F920E and R924E still interact with ANKS3 SAM-CC, as indicated by a gel shift (Fig. 5-3). However, no gel shift is detected when the BICC1 SAM ML mutants E898K and D911K are mixed with ANKS3 SAM-CC (Fig. 5-3). The BICC1 SAM ML mutants have a wild-type EH surface which binds the ML surface of ANKS3 SAM. Since this interaction does not occur in the presence of the ANKS3 SAM-CC construct, it suggests the coiled coil extension is interfering with the ML surface of the ANKS3 SAM domain.

Discussion

Sequences flanking SAM domains have been shown to impact their stability and function. We have demonstrated that the coiled coil region immediately C-terminal to the ANKS3 SAM domain appears to inhibit ANKS3 SAM polymerization and prevent formation of the BICC1 SAM-EH/ANKS3 SAM-ML interface. This coiled coil extension does not seem to affect interactions involving the ANKS3 SAM EH surface since ANKS3 SAM-CC is still able to bind the ML surfaces of both ANKS6 SAM and BICC1 SAM. These results therefore suggest that the coiled coil extension of the ANKS3 SAM domain interferes with or blocks the ANKS3 SAM ML surface.

These results provide a basis for future experiments to better understand how this coiled coil extension is impacting the ANKS3 SAM domain. Oftentimes coiled coils bind each other, forming dimers and trimers. Although the negGFP fusion of ANKS3 SAM-CC suggests the construct is monomeric, the charge repulsion from neighboring negGFPs has been shown to weaken protein-protein interactions (Knight et al., 2011). Removal of this charge repulsion by removal of the negGFP fusion may reveal a weakly associating oligomeric state, which may involve SAM domain polymer formation or coiled coil interactions. Purified ANKS3 SAM-CC should be assessed by SEC-MALS to determine whether it is monomer or oligomer.

Our results suggest that the coiled coil extension is interfering with the ANKS3 SAM ML surface. It is possible that this coiled coil is binding the SAM domain either on or nearby the ML surface and this is precluding ANKS3 SAM polymerization and interaction with BICC1. Flanking regions have previously been shown to directly bind SAM domains (Koveal et al., 2012; Robinson et al., 2012a). In the SAM domain of the *Drosophila* protein Ph, an N-terminal linker region directly binds the SAM domain and limits polymerization. Although the mechanistic basis behind this effect remains unknown, one theory is that charged patches on the polymer surface

influence linker positioning and this impacts polymerization (Robinson et al., 2012b). An electrostatic interaction directly between the ANKS3 SAM ML surface and coiled coil is not likely since each is negatively charged. However, the ANKS3 SAM polymer does display alternating electrostatic surfaces (Leettola, et al., 2014) and the negatively charged coiled coil may be interacting with the positive surface to limit polymerization in this manner. To better understand if the coiled coil is interacting with the SAM domain, protein stability could be measured by thermal denaturation. Determining whether the unfolding transition is monophasic or biphasic would yield insight into whether the coiled coil is accessory to the SAM domain or forms a single structural unit. If the coiled coil directly bound the SAM domain, these additional interactions would be expected to stabilize the protein and manifest as a higher Tm. Ideally a crystal structure of ANKS3 SAM-CC would show exactly how the coiled coil associates with the SAM domain and initial sparse matrix screens set up at 4^{o}C produced a few crystal hits to be followed up on.

In the biological context of ANKS3, our results suggest that the coiled coil region may be affecting polymerization *in vivo*. However, full-length ANKS3 has been observed to self-associate, supporting that oligomerization still occurs in the presence of the coiled coil (Yakulov et al., 2015). Future work including assessments of how the coiled coil alters the affinity of ANKS3 SAM homotypic and heterotypic interactions should be pursued to investigate how this C-terminal segment impacts ANKS3 polymerization and protein function.

Methods

Cloning

The negGFP human ANKS3 SAM fusion construct is as described previously (Knight et al., 2011). Three rounds of extension PCR using the following codon-optimized reverse primers were used to add residues 491-529 of the full-length protein to the C-terminal end of the ANKS3 SAM domain: P1 = 5'- CCAGACGGTCCGCGTACGCCAGTTCCAGCGCGTCACCCGGCGGGC GGGCACTGCTGTGCC-3', P2 =5'-CGCAACGTTTGTGCAGCTGGATCGCCAGTTCCTGCATT TCCGCTTCCAGACGGTCCGCG-3', P3 = 5'-GCGACTGGCCGGCCTTAAACCTGACCACGGG TCGCTTCAACTTCTTCGCAACGTTTGTGCAGCTG-3'. A negGFP fusion of the ANKS3 SAM domain with coiled coil extension (ANKS3 SAM-CC) was generated by cloning into a pBAD-HisA vector with negGFP as described previously (Knight et al., 2011). Cloned sequences were verified by DNA sequencing (Genewiz).

negGFP native gel assays

Preparation of negGFP fusion protein lysates was as described in prior work (Leettola, et al., 2014). The fluorescence intensity of lysates was measured as described previously (Knight et al., 2011). negGFP fusion proteins were loaded at equal amounts or the desired ratios, based on fluorescence. To control for concentration effects, gel samples were diluted using 20mM Tris pH 7.5, 0.3M NaCl, 2mM TCEP, 5mM DTT to the same final volumes. All samples were allowed to equilibrate at 4^{o}C for 4 hours prior to the addition of 4X RunBlue Native Sample Buffer (Expedeon). Samples were loaded on a 20% RunBlue 12-well native gel (Expedeon), run at 90V at 4^{o}C on ice for 15 hours, and visualized as described previously (Knight et al., 2011).

Figures

A.

```
421       430        440        450        460        470        480
  RAPYSGPQDL AALLEQIGCL KYLQVFEEQD VDLRIFLTLT ESDLKEIGIT LFGPKRKMTS
          490        500        510        520        530
  AIARWHSSAR PPGDALELAY ADRLEAEMQE LAIQLHKRCE EVEATRGQVC QEQ
```

B.

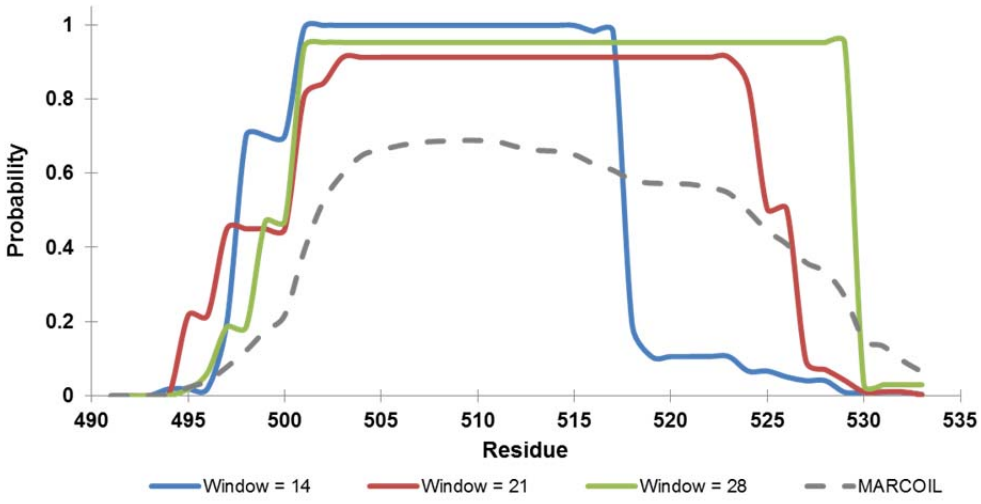


Figure 5-1. A coiled coil is predicted C-terminal to the SAM domain of ANKS3.

A) Sequence of the ANKS3 SAM domain (bold) and C-terminal flanking region. Residues are numbered according to the full length sequence (Uniprot:Q6ZW76). B) Probability of the C-terminal flanking region being a coiled-coil. Predictions were performed using the MARCOIL server (dashed line) and in three different frame windows using the COILS server (solid lines) (Delorenzi and Speed, 2002; Lupas et al., 1991). COILS predictions used the MTIDK matrix derived from myosins, paramyosins, tropomyosins, intermediate filaments type I - V, desmosomal proteins and kinesins. The MARCOIL prediction used the 9FAM matrix, derived from 9 families of proteins.

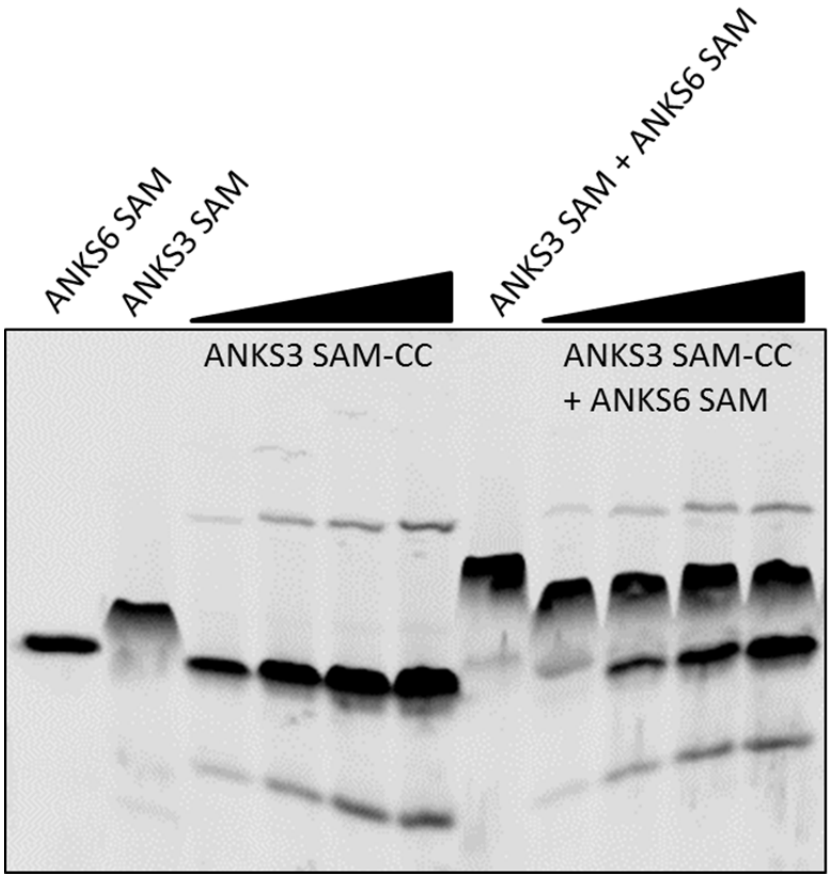


Figure 5-2. The C-terminal coiled coil prevents ANKS3 SAM polymerization.

negGFP fusions of ANKS3 SAM, ANKS6 SAM and ANKS3 SAM-CC were run on a native gel. In lanes containing only ANKS3 SAM-CC (lanes 3-6), fusion protein was loaded from left to right at 1X, 2X, 3X, and 4X the amount of fusion protein compared to ANKS6 SAM (lane 1) and ANKS3 SAM (lane 2), based on fluorescence. The negGFP fusions of ANKS3 SAM and ANKS6 SAM mixed in a 1:1 ratio based on fluorescence results in a gel shift (lane 7) indicative of interaction. The negGFP fusions of ANKS3 SAM-CC and ANKS6 SAM mixed from left to right in 1:1, 2:1, 3:1, and 4:1 ratios based on fluorescence (lanes 8-11) results in a gel shift indicating interaction.

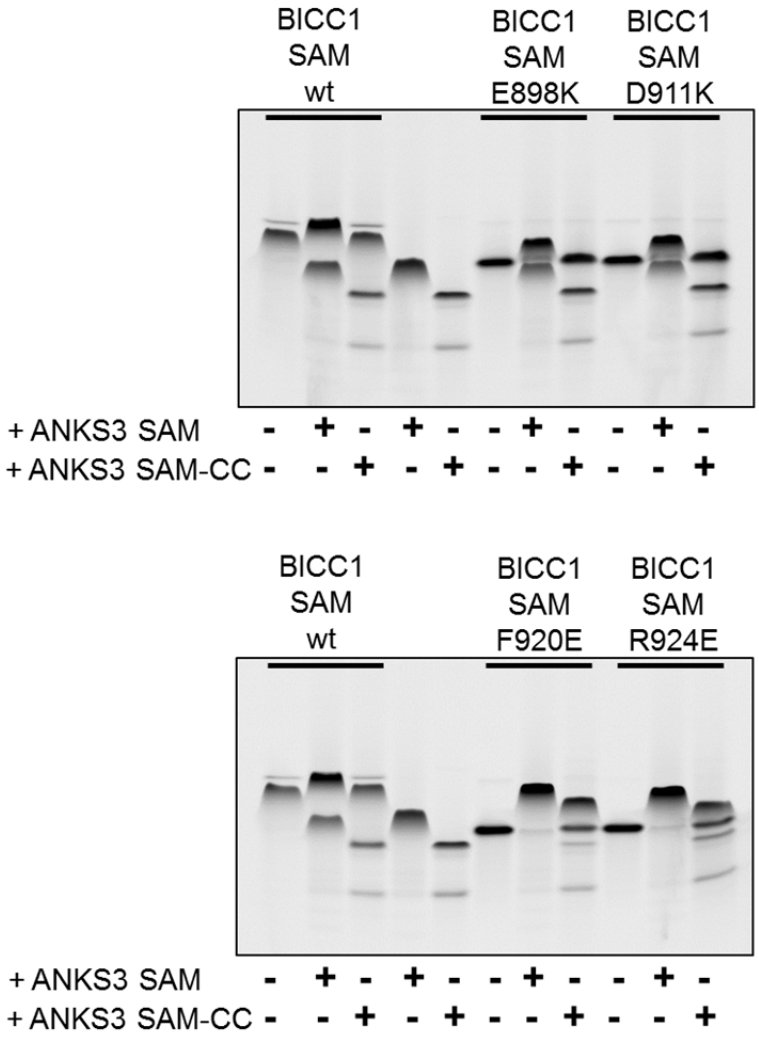


Figure 5-3. The C-terminal coiled coil prevents ANKS3 SAM interaction with the BICC1 EH surface.

negGFP fusions of BICC1 SAM constructs, ANKS3 SAM, and ANKS3 SAM-CC were run on native gels. In lanes containing two constructs, the negGFP fusion proteins were mixed in a 1:1 ratio based on fluorescence. Wild-type BICC1 SAM interacts with ANKS3 SAM but not ANKS3 SAM-CC. The BICC1 ML surface mutants E898K and D911K (with native EH surfaces) interact with ANKS3 SAM but not ANKS3 SAM-CC. The BICC1 EH surface mutants F920E and R924E (with native ML surfaces) interact with both ANKS3 SAM and ANKS3 SAM-CC. Interactions are detected as gel shifts.

References

Baron, M.K., Boeckers, T.M., Vaida, B., Faham, S., Gingery, M., Sawaya, M.R., Salyer, D., Gundelfinger, E.D., and Bowie, J.U. (2006). An Architectural Framework That May Lie at the Core of the Postsynaptic Density. Science *311*, 531–535.

Chaki, M., Airik, R., Ghosh, A.K., Giles, R.H., Chen, R., Slaats, G.G., Wang, H., Hurd, T.W., Zhou, W., Cluckey, A., et al. (2012). Exome Capture Reveals ZNF423 and CEP164 Mutations, Linking Renal Ciliopathies to DNA Damage Response Signaling. Cell *150*, 533–548.

Czarnecki, P.G., Gabriel, G.C., Manning, D.K., Sergeev, M., Lemke, K., Klena, N.T., Liu, X., Chen, Y., Li, Y., San Agustin, J.T., et al. (2015). ANKS6 is the critical activator of NEK8 kinase in embryonic situs determination and organ patterning. Nat. Commun. *6*, 6023.

Delorenzi, M., and Speed, T. (2002). An HHM model for coiled-coil domains and a comparison with PSSM-based predictions. Bioinformatics *18*, 617–625.

Failler, M., Gee, H.Y., Krug, P., Joo, K., Halbritter, J., Belkacem, L., Filhol, E., Porath, J.D., Braun, D.A., Schueler, M., et al. (2014). Mutations of CEP83 Cause Infantile Nephronophthisis and Intellectual Disability. Am. J. Hum. Genet. *94*, 905–914.

Halbritter, J., Bizet, A.A., Schmidts, M., Porath, J.D., Braun, D.A., Gee, H.Y., McInerney-Leo, A.M., Krug, P., Filhol, E., Davis, E.E., et al. (2013). Defects in the IFT-B Component IFT172 Cause Jeune and Mainzer-Saldino Syndromes in Humans. Am. J. Hum. Genet. *93*, 915–925.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Hildebrandt, F., Attanasio, M., and Otto, E. (2009). Nephronophthisis: Disease Mechanisms of a Ciliopathy. J. Am. Soc. Nephrol. *20*, 23–35.

Hoff, S., Halbritter, J., Epting, D., Frank, V., Nguyen, T.-M.T., van Reeuwijk, J., Boehlke, C., Schell, C., Yasunaga, T., Helmstädter, M., et al. (2013). ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3. Nat. Genet. *45*, 951–956.

Kim, C.A. (2005). Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. J. Biol. Chem. *280*, 27769–27775.

Kim, C.A., and Bowie, J.U. (2003). SAM domains: uniform structure, diversity of function. Trends Biochem. Sci. *28*, 625–628.

Kim, C.A., Phillips, M.L., Kim, W., Gingery, M., Tran, H.H., Robinson, M.A., Faham, S., and Bowie, J.U. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. EMBO J. *20*, 4173–4182.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Koveal, D., Schuh-Nuhfer, N., Ritt, D., Page, R., Morrison, D.K., and Peti, W. (2012). A CC-SAM, for Coiled Coil–Sterile α Motif, Domain Targets the Scaffold KSR-1 to Specific Sites in the Plasma Membrane. Sci. Signal. *5*, ra94.

Kwan, J.J., Warner, N., Pawson, T., and Donaldson, L.W. (2004). The Solution Structure of the S.cerevisiae Ste11 MAPKKK SAM Domain and its Partnership with Ste50. J. Mol. Biol. *342*, 681–693.

Leettola,, C.N., Knight, M.J., Cascio, D., Hoffmann, S., and Bowie, J.U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. BMC Struct. Biol. *14*.

Leone, M., Cellitti, J., and Pellecchia, M. (2009). The Sam domain of the lipid phosphatase Ship2 adopts a common model to interact with Arap3-Sam and EphA2-Sam. BMC Struct. Biol. *9*, 59.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting Coiled Coils from Protein Sequences. Science *252*, 1162–1164.

Qiao, F., and Bowie, J.U. (2005). The many faces of SAM. Sci. Signal. *2005*, re7.

Qiao, F., Song, H., Kim, C.A., Sawaya, M.R., Hunter, J.B., Gingery, M., Rebay, I., Courey, A.J., and Bowie, J.U. (2004). Derepression by depolymerization: structural insights into the regulation of Yan by Mae. Cell *118*, 163–173.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Robinson, A.K., Leal, B.Z., Chadwell, L.V., Wang, R., Ilangovan, U., Kaur, Y., Junco, S.E., Schirf, V., Osmulski, P.A., Gaczynska, M., et al. (2012a). The Growth-Suppressive Function of the Polycomb Group Protein Polyhomeotic Is Mediated by Polymerization of Its Sterile Alpha Motif (SAM) Domain. J. Biol. Chem. *287*, 8702–8713.

Robinson, A.K., Leal, B.Z., Nanyes, D.R., Kaur, Y., Ilangovan, U., Schirf, V., Hinck, A.P., Demeler, B., and Kim, C.A. (2012b). Human Polyhomeotic Homolog 3 (PHC3) Sterile Alpha Motif (SAM) Linker Allows Open-Ended Polymerization of PHC3 SAM. Biochemistry *51*, 5379–5386.

Schueler, M., Braun, D.A., Chandrasekar, G., Gee, H.Y., Klasson, T.D., Halbritter, J., Bieder, A., Porath, J.D., Airik, R., Zhou, W., et al. (2015). DCDC2 Mutations Cause a Renal-Hepatic Ciliopathy by Disrupting Wnt Signaling. Am. J. Hum. Genet. *96*, 81–92.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. *19*, 1354–1365.

Stafford, R.L., Hinde, E., Knight, M.J., Pennella, M.A., Ear, J., Digman, M.A., Gratton, E., and Bowie, J.U. (2011). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK. Structure *19*, 1826–1836.

The GPN Study Group, Halbritter, J., Porath, J.D., Diaz, K.A., Braun, D.A., Kohl, S., Chaki, M., Allen, S.J., Soliman, N.A., Hildebrandt, F., et al. (2013). Identification of 99 novel mutations in a worldwide cohort of 1,056 patients with a nephronophthisis-related ciliopathy. Hum. Genet. *132*, 865–884.

Wei, Z., Zheng, S., Spangler, S.A., Yu, C., Hoogenraad, C.C., and Zhang, M. (2011). Liprin-Mediated Large Signaling Complex Organization Revealed by the Liprin-α/CASK and Liprin-α/Liprin-β Complex Structures. Mol. Cell *43*, 586–598.

Yakulov, T.A., Yasunaga, T., Ramachandran, H., Engel, C., Müller, B., Hoff, S., Dengjel, J., Lienkamp, S.S., and Walz, G. (2015). Anks3 interacts with nephronophthisis proteins and is required for normal renal development. Kidney Int.

Chapter 6

Conclusion and Future Directions

Conclusion

At the initiation of this research, fairly little was known about the SAM domains of BICC1 and ANKS6. Genetic analysis of the PKD/Mhm(cy/+) rat had revealed the R823W point mutation within the SAM domain of ANKS6 was causative of polycystic kidney disease (Brown et al., 2005; Neudecker et al., 2010). However, a mechanistic basis explaining how this mutation led to disease was unknown. The developmental regulator BICC1 also had an established role in renal development as mutation or altered expression of this protein resulted in cystic kidney phenotypes in humans and several animal models (Bouvrette et al., 2010; Bryda et al., 2003; Kraus et al., 2012). The SAM domain of BICC1 had also been demonstrated as being able to self-associate, required for cellular localization to the periphery of P-bodies, and required for the activity of BICC1 as an inhibitor of canonical Wnt signaling (Knight et al., 2011; Maisonneuve et al., 2009). However, a molecular understanding of the BICC1 SAM domain and its potential interactors was lacking.

Here we have identified the SAM domain of the previously uncharacterized protein ANKS3 as an interactant of the SAM domains of ANKS6 and BICC1. We started by solving a crystal structure of the ANKS3 SAM domain and showed that it polymerized with moderate affinity. Interestingly, packing within the crystal structure revealed three individual polymers intertwined as a triple helix. Untagged SAM domains have not previously been observed to intertwine as triple helices, so this may represent a new form of SAM domain inter-polymer packing. TEM imaging of the His_6-tagged wild-type ANKS3 SAM domain also revealed that individual polymers can associate side-by-side into sheets.

We then characterized the interaction between the SAM domains of ANKS3 and ANKS6. We demonstrated that these SAM domains tightly bind each other and solved a crystal structure of the heterodimer which revealed an interaction between the EH surface of ANKS3 and the ML

surface of ANKS6. Using a negGFP native gel assay we showed that the disease-related R823W and I817N mutations in ANKS6 result in a disrupted ability to bind the SAM domain of ANKS3. This is because the R823W and I817N point mutations destabilize and disrupt the tertiary structure of the ANKS6 SAM domain. Thus, we identified ANSK3 as the first direct biding partner of the ANKS6 SAM domain and suggested that the perturbation of this interaction may lead to disease.

We next turned our attention to the BICC1 SAM domain and with a high-resolution crystal structure examined atomic details of the BICC1 SAM domain polymer. The heterogeneity of the SAM domain interfaces observed in the crystal structure suggest the BICC1 SAM domain polymer may be highly flexible. Analysis of the BICC1 SAM domain by TEM revealed a natural propensity for individual SAM domain polymers to form an ordered 2D-crystalline-like array. We also discovered a novel hetero-interaction between the SAM domains of BICC1 and ANKS3 and determined that these SAM domains bind each other with moderate affinity using both their ML and EH surfaces, allowing the possibility of the first observed alternating SAM domain co-polymer. Finally, we demonstrated a high affinity novel hetero-interaction between the SAM domains of BICC1 and ANKS6, which makes logical sense in light of the structural and sequence homology between the SAM domains of BICC1 and ANKS3.

Cumulatively, this work provides structural and biochemical evidence for the varied interactions between the SAM domains of ANKS3, BICC1, and ANKS6. Current research suggests these proteins function in large protein complexes. Thus, this network of SAM domain interactions may create scaffolding backbones for protein complex organization (Fig. 6-1).

In support of this, ANKS3 and ANKS6 were recently shown to associate with the NPHP2-3-9 module (Yakulov et al., 2015). This interaction is likely mediated by the ankyrin repeats of ANKS6, which were demonstrated to be required for interaction with NPHP9

(Czarnecki et al., 2015) (Fig. 6-1A top). ANKS3 has also been shown to associate with the NPHP1-4-8 module, although this is in the absence of ANKS6 (Yakulov et al., 2015) (Fig. 6-1A bottom). The NPHP2-3-9 module localizes to the inversin compartment of primary cilia while the NPHP1-4-8 module localizes to the transition zone of primary cilia (Szymanska and Johnson, 2012). Thus, these protein complexes containing ANKS3 and ANKS6 are sequestered within distinct regions of the primary cilia and may be important for ciliary assembly, signaling, or ciliary trafficking.

Through co-IP experiments BICC1 has recently been shown to associate with ANKS3 and ANKS6, supporting our findings and the plausibility of the scaffolds diagrammed in Fig. 6-1B (Yakulov et al., 2015). Unfortunately, the lack of additional characterization of these protein complexes means cellular localization and the identity of other involved proteins remains unknown. However, BICC1 has been detected in the bulb of primary cilia, suggesting a possible location for these complexes (Mohieldin et al., 2015). It is also possible that a polymer of BICC1 interacts solely with ANKS6 or exists in isolation (Fig. 6-1C). The BICC1 and ANSK6 proteins were previously shown to bind each other and domain truncations suggested this interaction required the SAM domain of ANKS6 but not the SAM domain of BICC1 (Stagner et al., 2009). Our results show the SAM domains of BICC1 and ANKS6 can bind each other. Due to the probable existence of these proteins within large complexes, the effect of single domain truncations may be masked by numerous other interactions holding the complex together.

Alternatively, the dynamic associations between the SAM domains of BICC1, ANKS3, and ANKS6 may serve a regulatory role by controlling the polymeric state of the different proteins (Fig. 6-2). Since the SAM domains of ANKS3 and BICC1 bind the SAM domain of ANKS6 with higher affinity then they bind themselves, ANKS6 may effectively depolymerize these polymers. This could serve to disassemble the protein scaffolds and disrupt the associated complexes. On the other hand, functional roles of the ANKS3 and BICC1 SAM

domain polymers may be disrupted by ANKS6 mediated depolymerization. For example, if polymerization by the BICC1 SAM domain is required for localization to P-bodies and for BICC1-mediated translational regulation, then depolymerization by ANKS6 would inhibit this cellular localization and function. Future work described below will be aimed at addressing several of these possibilities.

Functional relationships between BICC1, ANKS6, and ANKS3

As mentioned previously, BICC1 is a key developmental regulator and has been shown to attenuate canonical Wnt signaling. Although the exact mechanism is still unknown, BICC1 acts at the level of dishevelled (Dvl) and has been shown to inhibit Dvl2, which is a positive regulator of the canonical Wnt pathway (Maisonneuve et al., 2009). The SAM domain of BICC1 is important for this function as the inhibitory effect of BICC1 on Dvl2 is reduced when the SAM domain is deleted (Maisonneuve et al., 2009). The SAM domain of BICC1 is also required for localization to the periphery of RNA-processing P-bodies (Kraus et al., 2012; Maisonneuve et al., 2009). Given the importance of the BICC1 SAM domain for protein function and cellular localization we reasoned that the novel interactions we identified between the SAM domains of BICC1, ANKS3, and ANKS6 may impact the function and localization of BICC1.

In an effort to determine the functional significance of these SAM domain hetero-interactions we have collaborated with the lab of Dr. Michael Sheets at the University of Wisconsin. The Sheets lab has recently identified that BICC1 translationally represses *Xenopus* Cripto-1 (xCR1) mRNA in the vegetal cells of *Xenopus* embryos by binding specific elements in the 3' UTR. They have developed two new translational repression assays for *Xenopus* BICC1 function (Zhang et al., 2013). In a tethered repression assay, they fuse segments of BICC1 to the MS2 coat protein which binds a reporter mRNA. With this system, multiple copies of BICC1

bind the mRNA and the SAM domain contributes to translational repression but is not required. However, in their ectopic repressor assay, the SAM domain of BICC1 is required for translational repression. In this assay, a reporter mRNA with the xCR1 3' UTR is injected into *Xenopus* embryonic cells which do not normally express BICC1. The reporter mRNA is only translationally repressed when a BICC1 construct containing the SAM domain is expressed.

Since most of the prior work involving the SAM domain of BICC1 was conducted in mammalian cells, the Sheets lab has since transitioned to tissue culture to study BICC1 SAM domain function. Using GFP-tagged constructs in retinal pigment epithelial (RPE) cells, they have confirmed prior findings that human BICC1 localizes to cytosolic foci and that this localization is lost upon deletion of the SAM domain. The lab is still trying to get their translational repression assays functioning in human tissue culture. We have recommended several point mutations which abrogate BICC1 SAM domain polymerization and the Sheets lab is currently investigating what effect these mutations have on BICC1 localization and function. We have also provided the Sheets lab with full-length ANKS6 and ANKS3 constructs. They are currently investigating whether the full length proteins or the SAM domains alone affect the cellular localization and function of BICC1. Given that the SAM domain of BICC1 is critical for localization and function, we expect that expression of proteins which interact specifically with this domain will in some way alter localization and function and we are eagerly awaiting these results.

Effect of ANKS3 SAM domain point mutations *in vivo*

Mutations within the SAM domain of ANKS6 are responsible for NPHP in humans and PKD in rat and mouse (Bakey et al., 2015; Brown et al., 2005; Hoff et al., 2013). We now know that the SAM domains of ANKS6, ANKS3, and BICC1 are capable of interacting with each other

and that disease related point mutations disturb these interactions. However, animal models containing mutations in ANKS3 have not yet been studied. Since mutations in the SAM domains of ANKS6 and BICC1 result in developmental defects and cystic kidney disease in animal models, we reason that mutations in the SAM domain of ANKS3 may produce a similar phenotype.

To understand the functional significance of the SAM domain of ANKS3 we have collaborated with Dr. Sigrid Hoffman at the University of Heidelberg in Mannheim, Germany, with Dr. Dominique Gauguier at the University Pierre & Marie Curie in Paris, France, and with Dr. Brigitte Lelongt at Sorbonne Universités in Paris, France. These researchers are in the process of creating ANKS3 mutant rats and will then fully establish and characterize these mutant lines. Currently ANKS3 knockout rats and rats which overexpress ANKS3 are in progress. Rats containing specific point mutations within the SAM domain of ANSK3 are also in progress. Specifically, the mutations I36E and F53E within the ANKS3 SAM domain (numbering according to that in Leettola, et al., 2014) are being generated. Both point mutations limit ANKS3 SAM polymerization and limit formation of an alternating BICC1-ANKS3 SAM domain co-polymer. The F53E mutation also prevents the ANKS3 SAM domain from binding the ANKS6 SAM domain. With these specific point mutations it may be possible to detect phenotypic subtleties associated with the ANKS3 and ANKS6 SAM domain interaction, which should be preserved in the context of the I36E mutation but prevented in the presence of the F53E mutant. A recent knockdown of *anks3* in zebrafish resulted in cystic kidneys and laterality defects, supporting that this protein is essential for proper development (Yakulov et al., 2015). We are currently awaiting the final development and characterization of these mutant rats and expect to learn a wealth of information regarding the functional significance of ANKS3 and its SAM domain.

Ankyrin repeat domains of ANKS3 and ANKS6

Both human ANKS3 and ANKS6 each contain ankyrin repeats at their N-termini. Human ANKS3 contains 6 ankyrin repeats and human ANKS6 has 11 ankyrin repeats. The ankyrin repeat motif is defined as a β-hairpin-helix-loop-helix fold (Sedgwick and Smerdon, 1999). Within proteins ankyrin repeats are clustered and stack together into a structure that resembles a cupped hand: the antiparallel α-helices form the concave palm of the hand and the extended β-hairpins resemble fingers (Sedgwick and Smerdon, 1999). Ankyrin repeats are protein-protein interaction domains and typically surface residues of the "cupped hand" comprise the binding interface (Mosavi et al., 2004). The ability to create diverse and unique binding surfaces on the ankyrin repeat scaffold allows ankyrin repeats to interact with a myriad of different proteins found in all cellular compartments and involved in a variety of biological functions (Mosavi et al., 2004; Sedgwick and Smerdon, 1999).

The ankyrin repeats of ANKS6 have been observed to bind the Nek8 kinase and are responsible for targeting ANKS6 to the inversin compartment of primary cilia (Czarnecki et al., 2015; Hoff et al., 2013). It has also been suggested that the ankyrin repeats of ANKS6 are capable of binding the SAM domain of ANKS6, thus allowing ANKS6 oligomerization (Stagner et al., 2009). Mutations within the ankyrin repeats of ANKS6 are associated with nephronophthisis in humans and with cystic kidneys accompanied by laterality defects in mice (Czarnecki et al., 2015; Hoff et al., 2013; Taskiran et al., 2014). No direct binding partner of the ANKS3 ankyrin repeats has yet been identified. However, ANKS3 does co-immunoprecipitate with the NPHP1-4-8 module, suggesting that one of these proteins may interact with the ankyrin repeats (Yakulov et al., 2015). Additionally, both ANKS3 and ANKS6 are hydroxylated by HIF1AN on conserved asparagine residues within their ankyrin repeats (Hoff et al., 2013; Yakulov et al., 2015).

Given the importance of the ankyrin repeat domains we have taken initial steps toward structural and biochemical characterization. Since cDNA for human ANKS3 was unavailable, we synthesized the full-length codon-optimized cDNA (Genewiz). Rare codons were swapped out individually as follows: CCC was changed to CCG; ATA was changed to ATT; CTA was changed to CTG; and AGG, AGA, CGA were changed to CGT, CGC, and CGG respectively. We generated His_6-SUMO fusion constructs containing solely the ankyrin repeats as well as including N- and C-terminal linker regions for both ANKS3 and ANKS6 (Fig. 6-3). ANKS6 cDNA was generously provided by Soeren Lienkamp at the University of Freiburg Medical Center, Freiburg, Germany. With the help of Tyler Luu, Christopher Koo, and Mark Arbing at the UCLA Protein Expression Technology Center, a protocol for the purification of the ankyrin repeats from both ANKS3 and ANKS6 was developed. Constructs containing additional residues N- and C-terminal to the ankyrin repeats did not express well or were extensively degraded. Fortunately, conditions were found to successfully purify the ankyrin repeat domains of both ANKS3 and ANKS6. His_6-SUMO fusions of both proteins were purified under denaturing conditions in the presence of 8M urea and proteins were refolded by serial dialysis into buffer containing no urea. Upon removal of the His_6-SUMO fusion by cleavage, the remaining proteins were stable and eluted as non-aggregated peaks by size exclusion chromatography. From this trial purification 300μL of purified ANKS3 ankyrin repeats at 10mg/mL and 1mL of purified ANKS6 ankyrin repeats at 20mg/mL were obtained. Purified proteins will be used for crystallography screens and will be tested for possible interactions with the SAM domains of ANSK3, ANKS6, and BICC1.

Figures

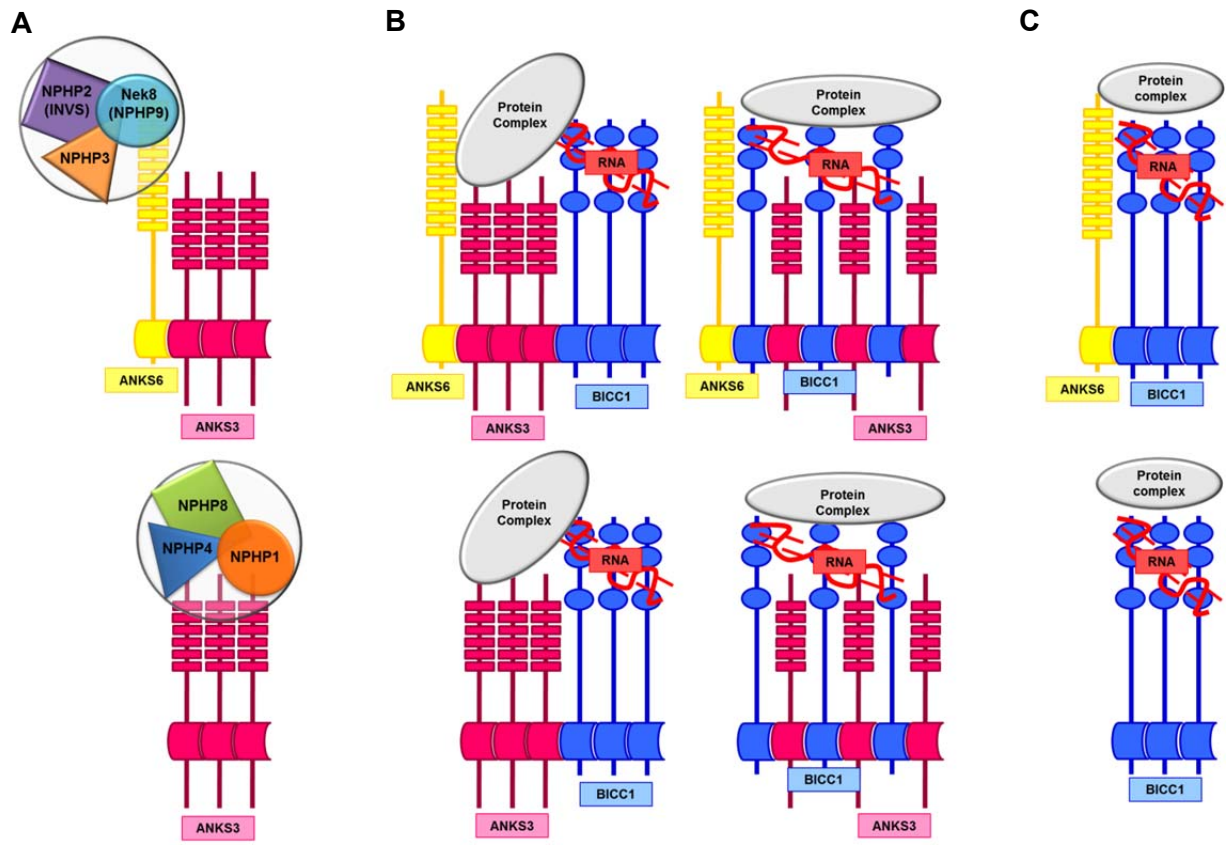


Figure 6-1. Possible BICC1, ANKS3, and ANKS6 SAM-domain-mediated scaffolds.

A) Possible scaffolds formed by the ANKS3 SAM domain and its hetero-interaction with the SAM domain of ANKS6. Co-IP experiments have revealed some components of the associated protein complexes (Yakulov et al., 2015). B) The SAM domains of BICC1, ANKS3, and ANKS6 may be present in the same scaffold. The SAM domains of ANKS3 and BICC1 may associate as a block co-polymer (left) or an alternating co-polymer (right). ANKS6 can bind the free EH surface of either the ANKS3 or BICC1 SAM domains. This scaffold may (top) or may not (bottom) involve binding of ANKS6. C) A BICC1 SAM domain polymer scaffold may (top) or may not (bottom) bind ANKS6.

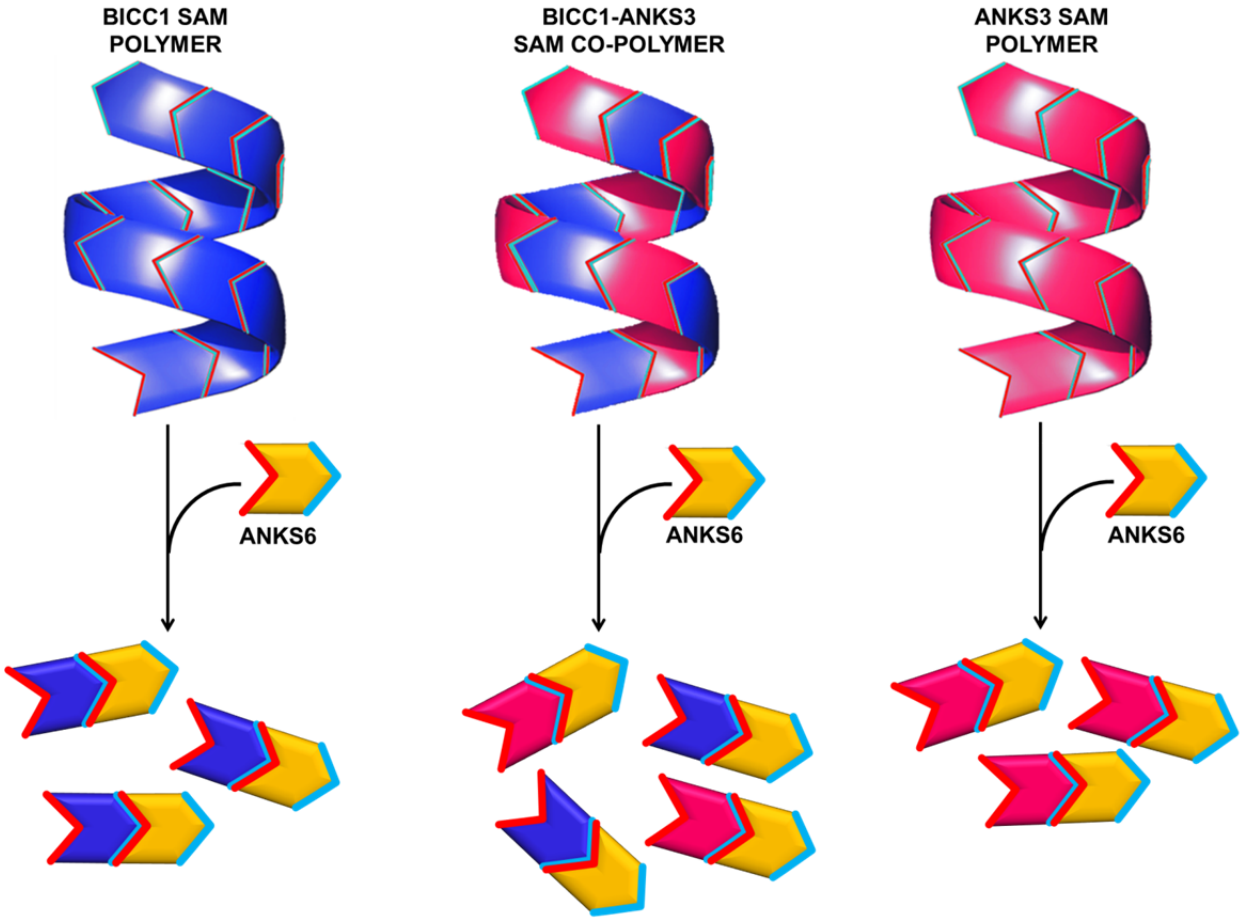
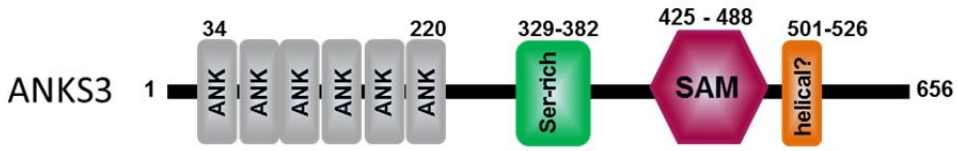


Figure 6-2. The ANKS6 SAM domain may depolymerize BICC1 and ANKS3 polymers.

Polymers of the BICC1 SAM domain, the ANKS3 SAM domain, or an alternating co-polymer may be depolymerized via a high affinity interaction with the SAM domain of ANKS6. Depolymerization may result in disassembled scaffolds or inhibition of polymer-mediated function.

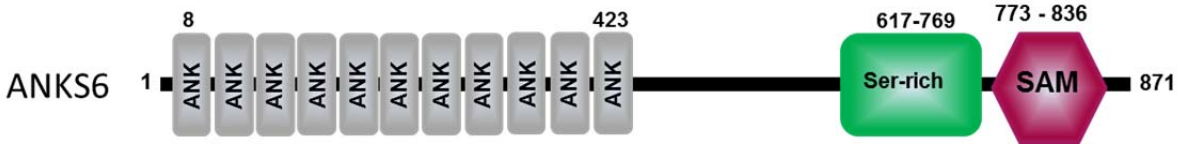
A



ANKS3 Constructs

Construct	Plasmid	Construct Number	Amino Acids	Expressed/Purified
ANKS3 FL_AR	pHis-SUMO	#8	1-230	No expression
ANKS3 FLΔSAM	pHis-SUMO	#23	1-420	Protein expressed with minor degradation
ANKS3 31_AR	pHis-SUMO	#14	31-230	Expressed protein was purified under denaturing conditions and refolded to a stable non-aggregated state
ANKS3 31_ΔSAM	pHis-SUMO	#26	31-420	No expression

B



ANKS6 Constructs

Construct	Plasmid	Construct Number	Amino Acids	Expressed/Purified
ANKS6 ANK	pHis-SUMO	#4	1-425	Expressed protein was purified under denaturing conditions and refolded to a stable non-aggregated state
ANKS6 + link	pHis-SUMO	#2	1-757	Protein expressed but extensively degraded

Figure 6-3. ANKS3 and ANKS6 ankyrin repeat constructs

A) The domain structure of human ANKS3 and the ankyrin repeat constructs which were generated. B) The domain structure of human ANKS6 and the ankyrin repeat constructs which were generated.

References

Bakey, Z., Bihoreau, M.-T., Piedagnel, R., Delestré, L., Arnould, C., de Villiers, A. d'Hotman, Devuyst, O., Hoffmann, S., Ronco, P., Gauguier, D., et al. (2015). The SAM domain of ANKS6 has different interacting partners and mutations can induce different cystic phenotypes. Kidney Int.

Bouvrette, D.J., Sittaramane, V., Heidel, J.R., Chandrasekhar, A., and Bryda, E.C. (2010). Knockdown of bicaudal C in zebrafish (Danio rerio) causes cystic kidneys: a nonmammalian model of polycystic kidney disease. Comp. Med. *60*, 96.

Brown, J.H., Bihoreau, M.-T., Hoffmann, S., Kranzlin, B., Tychinskaya, I., Obermuller, N., Podlich, D., Boehn, S.N., Kaisaki, P.J., Megel, N., et al. (2005). Missense Mutation in Sterile Motif of Novel Protein SamCystin is Associated with Polycystic Kidney Disease in (cy/+) Rat. J. Am. Soc. Nephrol. *16*, 3517–3526.

Bryda, E.C., Flaherty, L., Cogswell, C., Price, S.J., Hou, X., and Guay-Woodford, L.M. (2003). Positional cloning of jcpk/bpk locus of the mouse. Mamm. Genome *14*, 242–249.

Czarnecki, P.G., Gabriel, G.C., Manning, D.K., Sergeev, M., Lemke, K., Klena, N.T., Liu, X., Chen, Y., Li, Y., San Agustin, J.T., et al. (2015). ANKS6 is the critical activator of NEK8 kinase in embryonic situs determination and organ patterning. Nat. Commun. *6*, 6023.

Hoff, S., Halbritter, J., Epting, D., Frank, V., Nguyen, T.-M.T., van Reeuwijk, J., Boehlke, C., Schell, C., Yasunaga, T., Helmstädter, M., et al. (2013). ANKS6 is a central component of a nephronophthisis module linking NEK8 to INVS and NPHP3. Nat. Genet. *45*, 951–956.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Kraus, M.R.-C., Clauin, S., Pfister, Y., Di Maïo, M., Ulinski, T., Constam, D., Bellanné-Chantelot, C., and Grapin-Botton, A. (2012). Two mutations in human BICC1 resulting in Wnt pathway hyperactivity associated with cystic renal dysplasia. Hum. Mutat. *33*, 86–90.

Leettola,, C.N., Knight, M.J., Cascio, D., Hoffmann, S., and Bowie, J.U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. BMC Struct. Biol. *14*.

Maisonneuve, C., Guilleret, I., Vick, P., Weber, T., Andre, P., Beyer, T., Blum, M., and Constam, D.B. (2009). Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. Development *136*, 3019–3030.

Mohieldin, A.M., Haymour, H.S., Lo, S.T., AbouAlaiwi, W.A., Atkinson, K.F., Ward, C.J., Gao, M., Wessely, O., and Nauli, S.M. (2015). Protein composition and movements of membrane swellings associated with primary cilia. Cell. Mol. Life Sci. *72*, 2415–2429.

Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., and Peng, Z. (2004). The ankyrin repeat as molecular architecture for protein recognition. Protein Sci. *13*, 1435–1448.

Neudecker, S., Walz, R., Menon, K., Maier, E., Bihoreau, M.-T., Obermüller, N., Kränzlin, B., Gretz, N., and Hoffmann, S.C. (2010). Transgenic Overexpression of Anks6(p.R823W) Causes Polycystic Kidney Disease in Rats. Am. J. Pathol. *177*, 3000–3009.

Sedgwick, S.G., and Smerdon, S.J. (1999). The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem. Sci. *24*, 311–316.

Stagner, E.E., Bouvrette, D.J., Cheng, J., and Bryda, E.C. (2009). The polycystic kidney disease-related proteins Bicc1 and SamCystin interact. Biochem. Biophys. Res. Commun. *383*, 16–21.

Szymanska, K., and Johnson, C.A. (2012). The transition zone: an essential functional compartment of cilia. Cilia *1*.

Taskiran, E.Z., Korkmaz, E., Gucer, S., Kosukcu, C., Kaymaz, F., Koyunlar, C., Bryda, E.C., Chaki, M., Lu, D., Vadnagara, K., et al. (2014). Mutations in ANKS6 Cause a Nephronophthisis-Like Phenotype with ESRD. J. Am. Soc. Nephrol. *25*, 1653–1661.

Yakulov, T.A., Yasunaga, T., Ramachandran, H., Engel, C., Müller, B., Hoff, S., Dengjel, J., Lienkamp, S.S., and Walz, G. (2015). Anks3 interacts with nephronophthisis proteins and is required for normal renal development. Kidney Int.

Zhang, Y., Cooke, A., Park, S., Dewey, C.N., Wickens, M., and Sheets, M.D. (2013). Bicaudal-C spatially controls translation of vertebrate maternal mRNAs. RNA *19*, 1575–1582.

Appendix I

The F368C mutation causes unfolding of the ZAK SAM domain and is responsible for developmental limb defects

Abstract

Mutations affecting the SAM domain of the ZAK protein have recently been identified as causative of split foot malformations in humans and of hindlimb defects in mice. In humans, the F368C point mutation is causative of disease. We show here that this point mutation destabilizes the SAM domain and makes the resulting protein aggregation prone. Prior research has shown that ZAK self-associates. We find that the SAM domain is monomeric and dimerization of the protein is driven by a leucine zipper immediately N-terminal to the SAM domain. In an attempt to identify an interaction partner of the ZAK SAM domain, we used a negGFP native gel assay to screen against 54 different human SAM domains but did not detect any potential hetero-interactions. As destabilization of the SAM domain results in disease, future research should focus on identifying an interaction partner of the ZAK SAM domain.

Introduction

Split hand/foot malformation (SHFM) is a heterogeneous congenital defect in which malformations affect the distal limbs, often giving rise to a “lobster-claw” like cleft hand or cleft foot (Elliott et al., 2005). Causative mutations have only been identified in approximately 50% of patients (Sowińska-Seidler et al., 2014). Using homozygosity mapping and exome sequencing a research group led by Dr. Guntram Borck at the University of Ulm, Germany has identified the Phe368Cys mutation in the SAM domain of the ZAK protein as causative of a unique autosomal recessive split foot-hearing loss syndrome (Kakar et al., 2015). The importance of this SAM domain is further demonstrated by targeted deletions of the ZAK SAM domain in mice which result in complex hindlimb defects (Kakar et al., 2015).

ZAK is a member of the MAPKKK family of signal transducers and is so named because it is a leucine zipper (LZ) and sterile alpha motif (SAM) domain containing serine/threonine kinase. The ZAK protein is known to self-associate but it is currently unclear whether this capacity is due to the leucine zipper or the SAM domain (Liu et al., 2000). We were initially contacted by Dr. Borck to provide insight into how the F368C mutation affected the structure of the ZAK SAM domain and whether the ZAK SAM domain formed homotypic or heterotypic interactions with other SAM domains that were altered in the presence of the F368C mutation.

Using circular dichroism we determined that the F368C mutation causes a dramatic destabilization and unfolding of the ZAK SAM domain. We also show that the ZAK SAM domain is monomeric and that the leucine zipper is responsible for dimerization of the protein. Though we were unable to detect any heterotypic interactions between the ZAK SAM domain and the 54 different human SAM domains we screened against, our understanding of the structural effects of the F368C mutation provides a molecular basis for future studies investigating the function and binding partners of the ZAK SAM domain.

Results

The F368C mutation causes ZAK SAM unfolding

To understand how the F368C mutation affects the structure and thereby the function of the ZAK SAM domain we generated a PHYRE model of ZAK SAM (Fig. A-1A) (Kelley and Sternberg, 2009), based on the structure of DGKδ SAM (30% identity, PDB ID: 3BQ7) (Harada et al., 2008). The structural model predicts that Phe 368 is located in the hydrophobic core of the protein. Mutation to the Cys might therefore leave a hole in the hydrophobic core and lead to SAM domain instability. To test this hypothesis, we purified both the wild-type and the F368C mutant of the ZAK SAM domain and assessed the secondary structure content of each using circular dichroism (CD). Indeed the far-UV CD spectrum of the F368C mutant shows an approximately 30% loss of alpha-helicity, accompanied by an increase in random coil, compared to wild-type (Fig. A1-1B). Moreover, thermal unfolding experiments showed a cooperative, albeit irreversible, unfolding transition with a midpoint of 54^{o}C for the wild-type ZAK SAM domain, while the F368C mutant showed a complete loss of a cooperative unfolding transition (Fig. A1-1C). These results indicate that the F368C mutation strongly destabilizes the ZAK-SAM domain.

The ZAK SAM domain is monomeric

Previous work has indicated that the ZAK protein self-associates *in vivo* (Liu et al., 2000). Because the protein contains a SAM domain and a leucine zipper domain, both of which are protein-protein interaction domains, it was unclear which domain is responsible for oligomerization. We prepared negative GFP fusions (net charge -30) of the ZAK SAM domain and used a native gel assay developed previously to assess the oligomeric state of the protein (Knight et al., 2011). The negatively charged GFP causes the fusion proteins to migrate

towards the cathode on a native gel with migration distance based on the extent of oligomerization. The ZAK SAM domain appears monomeric since it migrates as a distinct band and with the same mobility as a known monomeric SAM domain, ANKS6 (Leettola, et al., 2014) (Fig. A1-2A). In contrast, a negative GFP fusion encompassing the upstream leucine-zipper region and the SAM domain (ZAK-LZ-SAM) migrates significantly slower. The F368C mutation in the ZAK SAM domain causes a retarded migration and this mutation in the context of the ZAK-LZ-SAM construct is aggregation prone, consistent with the mutation causing protein unfolding.

To further characterize the oligomeric state of the ZAK-SAM and ZAK-LZ-SAM proteins, we purified the proteins and assessed their size using SEC-MALS (Fig. A1-2B). Consistent with our native gel data, ZAK-SAM is monomeric with an observed molecular weight (MW) of 10.2 kDa (calculated MW = 8.9 kDa). The leucine zipper fused to the SAM domain had an observed MW of 32.3 kDa (predicted MW = 14.8 kDa), consistent with a dimeric state of the protein. Hence, the leucine zipper and not the SAM domain is responsible for generating dimeric ZAK.

We screened a large library of different negGFP-SAM fusions for binding to the negGFP ZAK-SAM using a native gel binding assay as described previously (Leettola, et al., 2014). To this end, we mixed negGFP ZAK-SAM in 1:1 ratios with 54 other human SAM domains (Table A1-1) and looked for interactions by native gel shifts. We did not find any such interactions. The Ste4 protein of *S. pombe* contains a SAM domain followed by a leucine zipper which forms a trimer and only in this oligomeric state does Ste4 SAM engage in a hetero-interaction with the Byr2 SAM domain (Ramachander, 2002). We reasoned the leucine zipper of ZAK may act similarly and that a hetero-SAM binding surface may only be presented when the protein dimerizes. We therefore repeated our negGFP native gel binding assay by mixing negGFP-ZAK-LZ-SAM in 3:1 ratios with the same 54 other human SAM domains, yet still saw no

interactions. It remains possible that the ZAK SAM domain engages in a hetero-SAM interaction we did not screen for or binds with affinity that is too weak to be detected by this assay. Alternatively, the ZAK SAM domain may bind a different protein target altogether and was suggested previously to bind the zinc finger protein ZZaPK (Yang, 2003).

Discussion

The F368C mutation in the SAM domain of ZAK is a newly identified mutation causative of a unique autosomal recessive split foot-hearing loss syndrome in humans (Kakar et al., 2015). We have determined that this point mutation destabilizes the SAM domain and results in protein aggregation. However, the mechanistic basis of how this point mutation leads to disease remains unknown. We did show that the ZAK SAM domain is monomeric and that the upstream leucine zipper drives protein dimerization. We attempted to identify an interaction partner of the ZAK SAM domain by using a negGFP native gel binding assay and screening against 54 different human SAM domains. However, no interaction was detected. This could be an artifact of the screen, which is known to produce false negatives (see Chapter 2). Alternatively, the ZAK SAM domain may engage in a very weak hetero-interaction not detectable in our screen or bind a SAM domain which we did not test.

SAM domains have also been shown to bind RNA (Aviv et al., 2003) and lipids (Barrera, 2003; Bhunia et al., 2009; Rufini et al., 2011), suggesting that these biological molecules may be binding partners of the ZAK SAM domain. Interestingly, the SAM domain-containing protein P63 is mutated in a subset of SHFM cases and has been shown to bind gangliosides (Rufini et al., 2011; Sowińska-Seidler et al., 2014). Furthermore, our collaborators have detected a decrease in *Tp63* expression in mice with a targeted deletion of the ZAK SAM domain, suggesting these two SAM domain-containing proteins are functionally linked (Kakar et al., 2015). In the P63 SAM domain, F565 is homologous by sequence alignment to F368 of the

ZAK SAM domain and the mutation F565L causes a related disorder known as ankyloblepharon-ectodermal dysplasia-clefting (AEC) syndrome. The F565L mutation is expected to create a packing defect in the hydrophobic core which destabilizes the SAM domain, much like the effect we observed for the F368C mutation of ZAK (Sathyamurthy et al., 2011). Given these similarities, a possible role of the ZAK SAM domain in lipid binding should be explored in future experiments.

Methods

Cloning and mutagenesis

negGFP-fusions of hZAK [UniProt:Q9NYL2] encompassing the SAM domain (residues 333-410, “ZAK-SAM”) and including the upstream leucine zipper region (residues 286-410, “ZAK-LZ-SAM”) were generated by cloning into a negGFP vector described previously (Knight et al., 2011). Hexahistidine small ubiquitin-like modifier (SUMO) tagged constructs were generated by cloning the above sequences into a pHis-SUMO vector (Senturia et al., 2010). Site-directed mutagenesis was performed using the Quickchange method (Agilent). All plasmid sequences were verified by DNA sequencing (Genewiz).

negGFP native gel assay

negGFP-human-ZAK fusions transformed into ARI814 cells were expressed, harvested, and lysed as described previously (Knight et al., 2011; Leettola, et al., 2014). Expression levels were determined by fluorescence intensity(Knight et al., 2011) and based on fluorescence, equal amounts of protein were loaded on a 20% RunBlue 12-well Native gel (Expedeon). Gels were run at 90 V for 16 hours at 4°C and visualized on a Bio-Rad Molecular Imager FX Pro-Plus using an excitation wavelength of 488 nm and an emission wavelength of 510 nm. In an attempt to identify ZAK hetero-SAM interactions, negGFP-ZAK-SAM was mixed in a 1:1 ratio based on fluorescence and negGFP-ZAK-LZ-SAM was mixed in a 3:1 ratio based on fluorescence with 54 other negGFP fusions of hSAM domains. Mixes were allowed to equilibrate at 4°C for 3 hours before running on native gels. Neither screen detected any hetero-interactions.

Protein expression and purification

pHis-SUMO constructs (ZAK-SAM, ZAK-SAM F368C, and ZAK-LZ-SAM) were transformed into Rosetta(DE3) pLysS cells (Novagen) and 4 L of culture was grown to an OD_{600} of 0.6 at which point cells were induced with 1mM isopropyl β-D-galactopyranoside and grown for an additional 16 hours at 18^{o}C. Harvested cells were lysed and pHis-SUMO tagged proteins were purified using Ni-NTA (Qiagen) as described previously(Leettola, et al., 2014). Proteins were dialyzed into 20mM $NaHPO_4$ pH 8, 0.5M NaCl, 2mM βME and digested with the His_6-tagged catalytic domain of SUMO protease 1 (ULP1) at a 50:1 protein:protease molar ratio for 16 hours at 4°C(Malakhov et al., 2004). The cleaved His_6-SUMO tag and ULP1 protease were removed by subtractive Ni-NTA. Proteins were further purified by dialysis into either 20mM Tris pH 7.5 (ZAK-SAM and ZAK-SAM F368C) or 20mM Tris pH 7.0 (ZAK-LZ-SAM), 50mM NaCl, 2mM βME and bound to a HiTrap Q HP column (GE). Elution was achieved using a shallow gradient of NaCl (0.05-1 M across 20 column volumes) dissolved in either 20mM Tris pH 7.0 or pH 7.5, 2mM βME. The proteins eluted between 0.3-0.4 M NaCl.

Circular dichroism

Spectra were collected for protein samples at 0.2 mg/ml in 10 mM Tris pH 7.5, 75 mM NaCl, 1mM DTT at 25°C using a 1mm path length cuvette on a JASCO J-715 circular dichroism spectrophotometer equipped with a Peltier temperature control. Spectra were analysed for secondary structure content using the Neural Network algorithm available in SoftSec (Softwood Inc.). Thermal melts were performed by monitoring the change in CD signal at 222 nm across a temperature range of 25-80°C, with ramping of 1°C per minute. Wild-type ZAK-SAM and ZAK-SAM F368C thermal denaturation was not reversible.

SEC-MALS

Purified ZAK-SAM and ZAK-LZ-SAM were dialyzed into 20mM Tris pH 7.5, 0.15M NaCl, 2mM TCEP. 120µL of protein at 10mg/mL was loaded onto a WTC-030S5 analytical size-exclusion column (Wyatt Technology Co.) equilibrated in 20mM Tris pH 7.5, 0.15M NaCl, 2mM TCEP using an AKTA purifier (GE) with a flow rate of 0.7mL/min and analyzed using a miniDAWN TREOS (Wyatt Technology Co.). Eluted protein peaks were analysed for calculated molecular weight and monodispersity using ASTRA software (Wyatt Technology Co.)

Figures and Tables

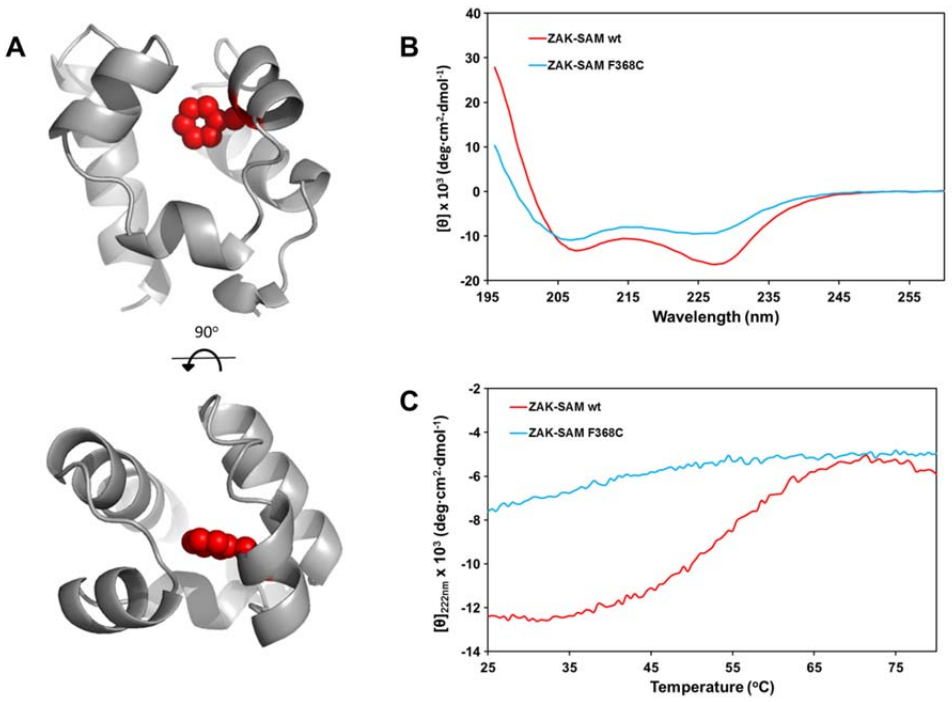


Figure A1-1. The F368C mutation causes ZAK SAM domain unfolding

A) Phyre model of the ZAK SAM domain. Phe 368 resides in the hydrophobic core of the SAM domain and is shown as spheres highlighted in red. B) CD spectra of wild-type ZAK-SAM and the F368C point mutant. The altered spectra of ZAK-SAM F368C correlates with a loss of alpha-helicity. C) Thermal denaturation curves of wild-type ZAK-SAM and ZAK-SAM F368C monitored by CD signal at 222nm. Wild-type ZAK-SAM exhibits cooperative unfolding while the F368C point mutant exhibits non-cooperative unfolding and is highly unstructured to begin with, as indicated by the reduced CD signal at 222nm.

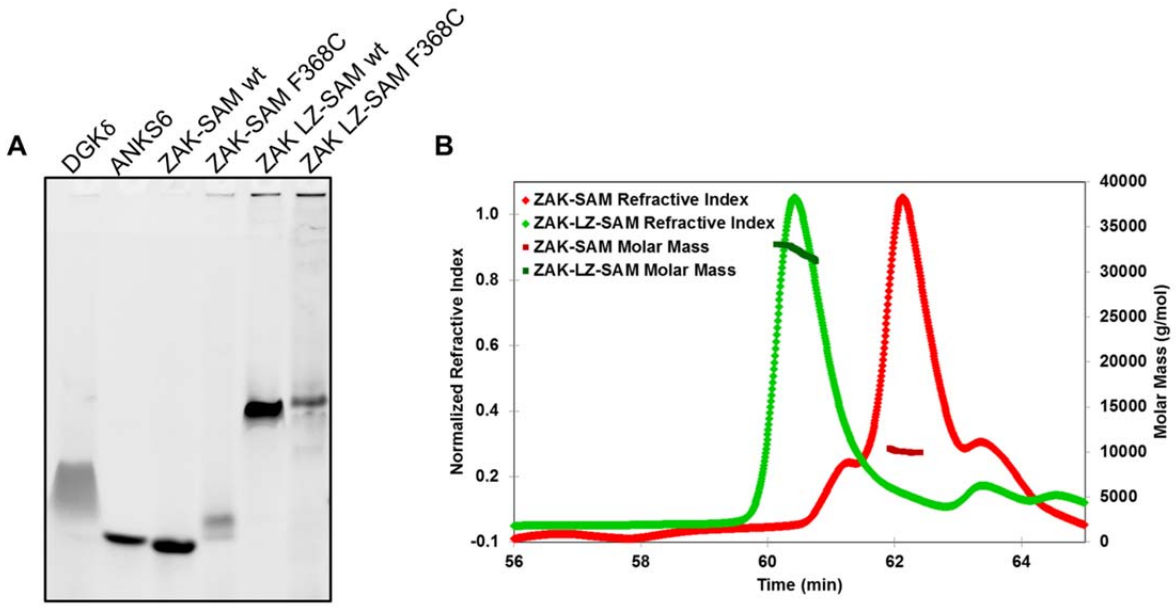


Figure A1-2. Dimerization of ZAK is mediated by the leucine zipper

A) negGFP fusions of wild-type and F368C mutants of ZAK-SAM and ZAK-LZ-SAM assessed by native gel electrophoresis. negGFP fusions of a monomeric SAM (ANKS6) and a polymeric SAM (DGKδ) are shown as controls. ZAK-SAM appears monomeric and addition of the leucine zipper causes a significant gel shift. B) SEC-MALS analysis of the ZAK SAM domain and the leucine zipper-SAM domain. The SAM domain alone is monomeric with an observed molecular weight of 10.2 kDa. The leucine zipper-SAM domain construct has an observed molecular weight of 32.3 kDa, corresponding to a homogenous population of dimer.

ANKS1A	EphA2	SAMD12
ANKS1B	EphA5	SAMD4B
ANKS3	EphA6	SAMD9
ANKS6	EphA7	SAMHD1
Arap1b	EphB1	SAMSN1
Atherin	EphB2	Sash1 - SAM1
BAR	EphB6	Sash1- SAM2
BICC1	Ese3b	Sash3
C140174	Ets1	SCMH1
Caskin1	GA-binding	Sec23
Caskin2	INPPL1	Shank2
Centaurin D1	LMBT-like	SLP76
CNKSR2	LRSAM	Smaug
DDHD-containing	MOB	StarD13
DGKh	Neurabin	Tankyrase1
DGKδ	Ph1	TP63
ELF3	Ph2	Usher1G
ELF5	Ph3	WDSUB1

Table A1-1. negGFP human SAM domains screened for interaction with ZAK

negGFP human SAM fusions tested for interaction with negGFP ZAK-SAM and negGFP ZAK-LZ-SAM using the native gel binding assay.

References

Aviv, T., Lin, Z., Rendl, L., M., Sicheri, F., and Smibert, C.A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. Nat. Struct. Biol. *10*, 614–621.

Barrera, F.N. (2003). Binding of the C-terminal Sterile Motif (SAM) Domain of Human p73 to Lipid Membranes. J. Biol. Chem. *278*, 46878–46885.

Bhunia, A., Domadia, P.N., Mohanram, H., and Bhattacharjya, S. (2009). NMR structural studies of the Ste11 SAM domain in the dodecyl phosphocholine micelle. Proteins Struct. Funct. Bioinforma. *74*, 328–343.

Elliott, A., Evans, J., and Chudley, A. (2005). Split Hand Foot Malformation (SHFM): Split hand/foot malformation. Clin. Genet. *68*, 501–505.

Harada, B.T., Knight, M.J., Imai, S., Qiao, F., Ramachander, R., Sawaya, M.R., Gingery, M., Sakane, F., and Bowie, J.U. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. Structure *16*, 380–387.

Kakar, N., Spielmann, M., Leettola, C., Kühl, S., Tayebi, N., Nürnberg, G., Sowada, N., Altmüller, J., Lupianez, D., Flöttmann, R., et al. (2015). Mutations in ZAK cause autosomal recessive split foot malformation in humans and complex hindlimb defects in mice. 26th Annual Meeting of the German Society of Humangenetics. Graz, Austria.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. Nat. Protoc. *4*, 363–371.

Knight, M.J., Leettola, C., Gingery, M., Li, H., and Bowie, J.U. (2011). A human sterile alpha motif domain polymerizome. Protein Sci. *20*, 1697–1706.

Leettola,, C.N., Knight, M.J., Cascio, D., Hoffmann, S., and Bowie, J.U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. BMC Struct. Biol. *14*.

Liu, T.-C., Huang, C.-J., Chu, Y.-C., Wei, C.-C., Chou, C.-C., Chou, M.-Y., Chou, C.-K., and Yang, J.-J. (2000). Cloning and Expression of ZAK, a Mixed Lineage Kinase-like Protein Containing a Leucine-Zipper and a Sterile-Alpha Motif. Biochem. Biophys. Res. Commun. *274*, 811–816.

Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D., and Butt, T.R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics *5*, 75–86.

Ramachander, R. (2002). Oligomerization-dependent Association of the SAM Domains from Schizosaccharomyces pombe Byr2 and Ste4. J. Biol. Chem. *277*, 39585–39593.

Rufini, S., Lena, A.M., Cadot, B., Mele, S., Amelio, I., Terrinoni, A., Desideri, A., Melino, G., and Candi, E. (2011). The sterile alpha-motif (SAM) domain of p63 binds in vitro monoasialoganglioside (GM1) micelles. Biochem. Pharmacol. *82*, 1262–1268.

Sathyamurthy, A., Freund, S.M.V., Johnson, C.M., Allen, M.D., and Bycroft, M. (2011). Structural basis of p63α SAM domain mutants involved in AEC syndrome: Mutants involved in AEC syndrome. FEBS J. *278*, 2680–2688.

Senturia, R., Faller, M., Yin, S., Loo, J.A., Cascio, D., Sawaya, M.R., Hwang, D., Clubb, R.T., and Guo, F. (2010). Structure of the dimerization domain of DiGeorge Critical Region 8. Protein Sci. *19*, 1354–1365.

Sowińska-Seidler, A., Socha, M., and Jamsheer, A. (2014). Split-hand/foot malformation - molecular cause and implications in genetic counseling. J. Appl. Genet. *55*, 105–115.

Yang, J.-J. (2003). A novel zinc finger protein, ZZaPK, interacts with ZAK and stimulates the ZAK-expressing cell re-entering the cell cycle. Biochem. Biophys. Res. Commun. *301*, 71–77.