STRUCTURE-FUNCTION STUDIES OF SCAFFOLDING PROTEINS INVOLVED IN THE FORMATION OF NEURONAL CONNECTIONS: AIDA-1 AND CASKIN2

EKATERINA SMIRNOVA

A DISSERTATION SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CHEMISTRY
YORK UNIVERSITY
TORONTO, ONTARIO

February 2017

Abstract

Modular proteins serve assembly platforms and often actively regulate cellular signaling events. An intrinsic diversity of interaction modules, typical for scaffolding proteins, facilitates the organization of numerous protein partners into signaling cascades, contributing to the spatial precision, efficiency and fidelity of signal transduction.

The role of complex molecular dynamics of postsynaptic density (PSD) proteins in synaptic plasticity is relatively new and yet to be fully understood. AIDA-1 is one of the most abundant members of the PSD protein family. Growing research evidence of multiple protein partnerships suggests that AIDA-1 functions as an essential PSD molecular scaffold, NMDA receptor functional mediator, and a synapse-to-nucleus messenger. The NMR structure of AIDA-1 carboxy-terminal phosphotyrosine binding domain (PTB), presented in this study, provided the structural basis for comparative analysis with the other PTB domain-containing proteins, Fe65 and X11/Mint1, that also participate in amyloid beta precursor protein (AβPP) processing and amyloid beta peptide (Aβ) secretion. A combination of peptide arrays, mutagenesis and fluorescence based assays was employed to characterize the affinity and specificity of the AIDA-1 PTB domain and AβPP intracellular domain (AICD) interaction.

Another modular protein of these studies is a pre-synaptic scaffolding protein, Caskin2. Presently, its function within the synapse is less clear compared to its more widely studied homolog, Caskin1. However, the structural differences between the two identified by our research suggest the possibility of distinct functional outcomes in the neuron. We demonstrated that Caskin2 Sterile Alpha Motif (SAM) assembles into an oligomeric architecture different from Caskin1, with the minimal repeating unit being a dimer, rather than a monomer. In invertebrates, Caskin has been functionally linked to LAR receptor tyrosine phosphatase functional pathways, implicated in axonogenesis and synaptogenesis. Using a combination of biophysical and biochemical methods, the partnership between Caskin2 and LAR *Homo sapiens* homologs was confirmed and characterized. These integrated structural and functional studies provide a platform for further elucidation of AIDA-1 and Caskin cellular functions.

I dedicate this dissertation to my loving family:
my husband, Mikhail Smirnov, my daughter, Anna Smirnova,
and my parents, Tamara and Petr Gusev.

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor Dr. Logan W. Donaldson for providing me with the opportunity to work in his laboratory and explore the whole new world of protein biochemistry. I am grateful for his guidance, support, and patience. Over the years under his wing, I believe I metamorphosed from a somewhat lost graduate student into a dedicated researcher. Thank you for believing in me and letting me take part in your research. You have encouraged me to listen and absorb, to search and question, to try fail and succeed.

I wish to thank my advisory committee members Dr. Philip E. Johnson and Dr. Yi Shen for providing valuable feedbacks, insightful comments and directions. My external committee member Steven Rafferty for joining in. I extend my gratitude to our collaborators and awesome neighbors Dr. Gorg Zoidl (who also readily joined my committee), Christiane Zoidl and their lab members for the great opportunity to take part in the collaboration study and for your friendship. Thank you for including me in your labfamily circle. I would like to extend my sincerest gratitude to Howard Hunter; I cannot thank you enough for your patience and professionalism. Thank you for guiding me through the basics of NMR experiments and data analysis. I extend my gratitude to teachers and mentors that had influenced and inspired me at different stages of my life and education journey: Pravdina Olga Vladimirovna, Zobkova Svetlana Vasilievna, Kasaikina Olga Tarasovna and Shmyreva Janetta Vladimirovna.

I cannot express an extent of my gratitude to my lab mate, my mentor and my soulmate Dr. Jamie J. Kwan. Thank you for always being there for me, believing in me and helping me to become a better scientist. Learning from you and being your friend have influenced me in so many ways. I extend my gratitude to my entire Lab Family and LSB friends that I have met over the years, worked side by side, and shared laughs and truly unforgettable moments: Sherry, Riya, Sarah, Albana, Shahin, Agnesa, Shaolong, Peter,

Anna, Rawaa, Sam, Jyotsna, Stefano, Cherie, Paige, Ira, Varvara, Sahar, and many others! I wish to include here the undergraduate students, Liora, Gabby, Ambica, Shaon and the others that allowed me to mentor them and, at the same time, to learn from them. Thanks for taking on your shoulders a fraction of work, especially endless protein purifications, the experience of French Pressure Cell weight unites us forever! Especial gratitude to the fellow graduate students Sladjana Slavko, Jiayin Bao, Oren Reinstein, Christina Lento and Rayan Siu for taking your time to share your knowledge with me about certain lab techniques and data analysis. Parts of this work would not be possible without your input.

I wish to thank my family for their unconditional love, all the patience, and constant support as we were taking this long journey together. Thank you, Mikhail, for your encouragement and support, for all the suffering through my technical talks. Thank you, Anna, for being so adorably proud of me and taking every opportunity to endorse your mom as a real scientist! Thank you both for keeping our home the warm and light Harbor and for being a source of endless happiness in my life. Thank you, Mom and Dad, for your love and support, for teaching me the values of education from early years and planting the seeds of my future ambitions. I love you all so much!

This dissertation is the combination of research resulting in two peer-reviewed publications as well as supplementary unpublished data. An additional paper produced by a collaborative effort is included as an appendix.

Chapter 1: The introduction aims to provide a general background and prepare the reader for the biochemical and biophysical content of the following chapters. It is composed of sections presenting modular domains and multidomain scaffolding proteins as an indispensable instrument in cellular signaling, including a detailed summary of phosphotyrosine binding domain (PTB) and sterile alpha motif (SAM) domains in terms of structural and functional highlights. The following sections then provide an overview of how these domains operate within the larger framework of AIDA-1 and Caskin2 neuronal proteins.

Chapter 2: The research presented in this chapter has been published in *PLoS One* journal (Smirnova *et al*., 2013). The solution structure of AIDA-1 PTB domain and its interaction with amyloid-beta precursor protein APP intracellular domain is presented. A consensus sequence around an NxxY motif was identified by a comprehensive peptide survey, accompanied by molecular modeling and binding affinity determinations by fluorescence anisotropy.

> Smirnova E., Shanbhag R., Kurabi A., Mobli M., Kwan J.J., Donaldson W.L. (2013) "Solution Structure and Peptide Binding of the PTB Domain from the AIDA-1 Postsynaptic Signaling Scaffolding Protein". *PLoS ONE* 8(6): e65605.

Authors' contributions:

The AIDA-1 PTB NMR structure reported in this publication was solved and refined by Dr. Logan Donaldson. I joined the laboratory as a MSc. student soon after the initial PTB structure calculations were accomplished. Additional NMR data acquisitions were done by Dr. Mehdi Mobli and Dr. Logan Donaldson at the Institute for Molecular Bioscience, University of Queensland, (Brisbane, Australia). I was responsible for preparing isotopically enriched protein samples for NMR studies. Preliminary interaction studies were accomplished by Riya Shanbhag and Arwa Kurabi. My contributions to the publication include CD spectroscopy, fluorescence anisotropy study and the peptide array and corresponding data analysis. The manuscript was written by Dr. L. Donaldson with subsections written by myself.

Chapter 3: The structure-based study contained in this chapter was published in *Cell Communication and Signaling* journal (Smirnova *et al.*, 2016). The crystal structure of the Caskin2 SAM tandem and oligomerization model distinct from its homolog Caskin1 are presented in this chapter. The structural architecture was further analyzed by a combination of biochemical and biophysical methods including mutagenesis, NMR, analytical ultracentrifugation (AUC) and fluorescence microscopy in live cells.

Smirnova E., Kwan J.J., Siu R., Gao X., Zoidl G., Demeler B., Saridakis V., and Donaldson L.W. (2016) "A New Mode of SAM Domain Mediated Oligomerization Observed in the Caskin2 Neuronal Scaffolding Protein". *Cell Communication and Signaling*, 1–14.

Authors' contributions:

The Caskin2 SAM SAM crystallization was accomplished by Dr. Jamie Kwan and the crystal structure was solved with an assistance of Dr. Vivian Saridakis. Cloning and site-directed mutagenesis were completed by myself and Dr. Jamie Kwan. I was responsible for preparation of isotopically labeled NMR samples as well as preparation of several protein sample sets for AUC analysis, which was performed by Dr. Borries Demeler at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center (San Antonio), and analyzed by Dr. Demeler, Dr. Donaldson and myself. In vivo experiments in Neuro2a cells and confocal microscopy data acquisition were accomplished by myself and Ryan Siu. The earlier draft of the publication was prepared by myself with incorporated AUC section prepared by Dr. B. Demeler. The final version of manuscript was prepared by Dr. Logan Donaldson after deliberations with myself, Dr Jamie Kwan and Dr. Borries Demeler and approved by all authors.

Chapter 4: The research contained in this chapter outlines the Caskin2 functional study that has not yet been published. A combination of *in vitro* and *in vivo* assays provide evidence that *Homo sapiens* Caskin2 interacts with the D2 domain of leukocyte common antigen-related (LAR) tyrosine phosphatase through its SAM1-SAM2 module. Spot blot experiments were performed with an assistance of Liora Naroditsky (the summer research student). Cloning of the C-term deletion ΔCSS construct was done by myself. All immunoprecipitation experiments were conducted by me. NMR titrations and analysis were performed by myself and Dr. Logan W. Donaldson. Cloning of EGFP and dsRED constructs were done by Jamie J. Kwan and myself. Neoro2A Cell culture, transient transfections and microscopy slide preparations were done by Dr. Zoidl Ph. D. student

Ryan C. F. Siu. Fluorescence anisotropy titrations, confocal microscopy imaging and data analysis were performed by myself consulting with Dr. Georg Zoidl and Ryan C. F. Siu.

Chapter 5: The final chapter summarizes the main outcomes of these studies and provides future directions.

Appendix A: The fundamental aspects of protein structure determination by nuclear magnetic resonance (NMR), a tool that is essential for the exploration of structure-function relationships.

Appendix B: A combined ITC and NMR spectroscopy study, to characterize a gap junction protein Connexin 36 (Cx36) and its interaction with calmodulin (CaM) has been included as an additional research accomplishment. This work was completed as part of a collaboration with the laboratory of Dr. G. Zoidl and published in *Frontiers in Molecular Neuroscience,* (2016).

> Siu R.C.F., Smirnova E., Brown C.A., Zoidl C., Spray D.C., Donaldson L.W., Zoidl G. (2016) "Structural and functional consequences of Connexin 36 (Cx36) interaction with Calmodulin". *Frontiers in Molecular Neuroscience. 9:120.*

Authors' contributions:

I performed and analyzed the ITC titrations of CaM with wild type and three mutant peptides of Cx36, deriving the dissociation constants for each. Two representative titrations were selected for the publication. I was responsible writing the sections related to ITC experiments, as well as took part in the manuscript editing at all the revision stages. I also assisted with protein sample preparations for NMR studies and protein expression and purifications of CAMKII protein fragments for the succeeding part of this study.

Table of Contents

List of Tables

Chapter 2:

Chapter 3:

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

List of Abbreviations

aa.	amino acid(s)
AICD	AβPP intracellular domain
AMPA	2-amino-3-(3- hydroxy-5-methyl-isoxazol-4-yl) propanoic acid
ASD	idiopathic autism spectrum disorders
Aβ	amyloid beta
AβPP	amyloid beta precursor protein
BLAST	basic local alignment search tool
CaMKII	Ca2+/calmodulin-dependent protein kinase II
CASK	Ca2+/calmodulin-associated Ser/Thr kinase
cKO	conditional knock-out
COSY	Correlation Spectroscopy
D2O	deuterium oxide
Dab	Disabled
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EDTA	ethylene diamide tetracetic acid
EGFR	epidermal growth factor receptor
EH	end-helix
EH/ML	end-helix/mid-loop

Eph	ephrin receptor
ER	endoplasmic reticulum
FGFR1	fibroblast growth factor receptor 1
FID	free induction decay
FRS2	fibroblast growth factor receptor substrate 2
GIT1	G-protein-coupled receptor kinase-interacting protein 1
GKAP	guanylate kinase-associated *protein*
HLH	helix loop helix
hr	hour
HSQC	Heteronuclear Single Quantum Coherence (HSQC) or Heteronuclear Single Quantum Correlation
IDE	insulin-degrading enzyme
imAPP	immature APP
IPTG	isopropyl beta-D-thiogalactoside
IRS-1	protein insulin receptor substrate-1
kb	kilobase
K_d	dissociation constant
kDa	kilo Daltons
LAR	Leukocyte common antigen-related
LB	Luria-Bertani
LDL	low-density lipoprotein
LDLR	low-density lipoprotein (LDL) receptor
LOF	loss-of-function phenotypes

MAGUK	membrane associated guanylate kinase
MAPK	mitogen-activated protein kinase
miRNA	Micro RNA
ML	mid-loop
mRNA	messenger RNA
MW	molecular weight
Nak	Numb-associated kinase
NCR	N-terminal conserved region
NIP	Numb-interacting protein
NMDA(R)	N-methyl-D-aspartate (receptor)
NOE	nuclear Overhauser effect
PBS	phosphate buffered saline
Pcdh 18	Protocadherin 18
PcG	polycomb group
PDZ	Post synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and Zonula occludens-1 protein (zo-1) (acronym derived from the names of the first three proteins in which the domain was observed); also, referred as PSD95-like domain
PH	pleckstrin homology
Ph	polyhomeotic
PID	phosphotyrosine interaction domains
PKC3	protein kinase C3
PLC	phospholipase C-gamma
PNT	Pointed

ppm	part per million
PSD	postsynaptic density
PTB	phosphotyrosine binding domain
PTP	protein tyrosine phosphatase
PTP-BL	protein tyrosine phosphatase-basophil like
PVDF	polyvinylidene fluoride
RF	radio frequency
RNA	ribonucleic acid
RPTP (PTP)	receptor-like protein tyrosine phosphatase
rRNA	ribosomal ribonucleic acid
RTK	receptor tyrosine kinase
RTK	receptor tyrosine kinase
SAM	sterile alpha motif
Scm	Sex-comb-on-midleg
SDM	Site Directed Mutagenesis
SDS	sodium dodecyl sulphate
SEP	(yeast sterility, Ets-related, PcG) protein family
SH2	Sarc homology-2
SH3	Sarc homology-3
Ship2	Src homology 2 domain-containing phosphoinositide-5-phosphatase 2
siRNA	small interfering RNA
SMART	simple modular architecture research tool
snRNPs	small nuclear ribonucleoproteins

SPM	(Scm, Ph, lethal-3 malignant brain tumor) protein family
SRE	Smaug recognition element
SRE	Smaug recognition element
TBS	tris buffered saline
TEL	translocation Ets leukemia
TOCSY	Total Correlated Spectroscopy
TrkA	Tropomyosin receptor kinase A
TROSY	transverse-relaxation optimized spectroscopy
YT	yeast tryptone

CHAPTER 1:

GENERAL INTRODUCTION

1.1. Cell Signaling and Signal Transduction

Cell signaling is a complex process regulating virtually every aspect of cellular function and essentially cell survival itself. It allows cells to detect and process sensory information from external stimuli and communicate with each other via distinct signaling pathways producing a coherent response. *Intercellular signaling* controls cell division and proliferation, differentiation and development as well as metabolic fluxes in different tissues. It works in synchrony with *intracellular signaling* which coordinates individual cell metabolism, protein expression, cell motility, and morphology. A human nervous system is incredibly complex with billions of neurons communicating every second via axons with their targets and their proteins, including cell surface receptors, downstream effectors, scaffolding and adaptor proteins. These proteins are the workhorses controlling the integrity of such complex processes as neurotransmission, synaptic plasticity, and synaptogenesis.

The complete sequencing of several eukaryotic genomes, including the human genome, provided a foundation for a whole new paradigm of proteomic studies with the ultimate goal of identifying and characterizing all protein networks and their connections with particular cellular mechanisms or functional pathways. Such knowledge ultimately drives a better understanding of malfunctions within organisms and opens up new possibilities for human disease therapeutics.

1.2. Modular Proteins in Cell Signaling and Signal Transduction

A comparison of the proteomes across different species and development of sequence alignment tools and protein databases led to the realization that eukaryotes share a number of recognizable protein families with the same type and number of protein domains (also called peptide recognition modules or PRMs). Evidently, the higher eukaryotes gained an evolutionary advantage by assembling these modules in numerous combinations; consequently, the human genome comprises almost twice of the multidomain combinations compared to worm or fly (Lander *et al.*, 2001).

The protein domain is traditionally defined as an evolutionarily conserved, independently folded globular structure, with an assigned or unknown biological function (Cesareni *et al.*, 2005). The primary functional outcomes of modular domains include colocalization of enzymes with their substrates resulting in signaling cascades and recognition of specific post-translational modifications (PTMs) and finally, cross-talk signaling modules that link different pathways. The modular signaling proteins such as kinases and phosphatases, although they harbor catalytic functions, are distinct from other

enzymes that are designed to process a large amount of substrate promptly (fast substrate turnover). Their catalytic center is often separated from their substrate binding motifs that specialize in recruiting specific protein targets (Cohen *et al.*, 1995). In addition, they possess other protein binding and extracellular communication modules. Kinases and phosphatases assemble into dynamic signal transduction cascades in conjunction with a non-catalytic adaptor and scaffolding proteins. This exceptional complexity allows them to link to downstream and upstream proteins and concentrate the signaling event at particular subcellular regions, therefore providing fine control of signal transduction.

In addition to polypeptide binding domains, there are a number of specialized lipid and carbohydrate binding modules. Overall, although employing diverse interaction models, modular domain binding is characterized by specific sequence recognition. For instance, PTB domains recognize NPxY in both a phosphorylation-dependent and -independent fashion; Sarc homology-2 (SH2) are phospho-tyrosine binding modules; a large group of polyproline sequence recognition domains includes the SH3, WW, GYE, UEV and EVH1 domains, PDZ domains which recognize the binding motifs at their C-terminal ends of their ligands; PH and PX domains bind phospholipids; and the SAM domain is known for its ability to assemble SAM-containing protein into homo- and/or hetero-oligomeric structures.

Although protein domains vary in size from as small as ~30 amino acids to as big as several hundred residues, the majority of modular domains are composed of ~50-150 residues (Petsko & Ringe, 2004). The modest interaction surface is energetically compensated by the extensive hydrogen bonding network and electrostatic interactions.

At the same time, the affinity of an enzyme for its binding epitope should not be too strong, in fact, it is typically found within the range from 10^{-3} M to 10^{-9} M (Petsko & Ringe, 2004), to assure fast dissociation rates required for dynamic signaling systems (Cesareni *et al.*, 2005). Another consequence of the small size is that the binding groove is often accompanied by various supporting sequences especially common in the unstructured loops. For example, the LAR receptor-like tyrosine phosphatase catalytic cleft formed by a motif named the PTP loop and two other substrate recognition regions: WPD and p-Tyr-binding loop (Nam *et al.*, 1999). Another reason for the small size is that the protein-interaction modules are often found in clusters, so called supramodules that ensure binding selectivity, such as phospholipase C-gamma (PLC) consisting of two PH, an SH2 and an SH3 domains which together engage in lipase recognition (Cesareni *et al.*, 2005). The small size of the minimal binding epitope is also often compensated by the adjacent residues that support binding specificity. For instance, a typical PTB ligand recognition spreads to the amino acids N-terminal from pY, while in the case of SH2 domains residues carboxy-terminal from phospho-tyrosine define their ligand specificity (Uhlik *et al.*, 2002).

Since the number of modular domains in any given organism is quite limited, most of them serve multiple functions. A certain degree of molecular cross-talk and redundancy is a natural evolutionary outcome and an essential instrument of cell survival. The observation that increasing complexity of the organism correlates with growing complexity of modular domain architecture supports the hypothesis that requirement for new signaling networks supporting more complex cellular functions prompted a new

domain evolution from preexisting domains as well as multi-modular assemblies (Ernst *et al.*, 2009; Jin *et al.*, 2009). For instance, certain non-catalytic domains of adaptor and scaffolding proteins could have originally diverged from enzymatic domains. An example of such domain evolution could be found in the MAGUK (membrane associated guanylate kinase) adaptor protein family that plays a critical role at neuronal synapses supporting a number of protein-protein scaffolds. The C-terminal domain of MAGUKs are strikingly similar to a guanylate kinase domain, but lacks its catalytic function. Studies indicate that its pSer/pThr-binding pocket most likely had evolved from the catalytic GMP domain of guanylate kinases (Jin & Pawson, 2012; Zhu et al., 2011).

In the following sections, I will discuss in greater detail the structural and functional aspects of SAM and PTB domains as the key elements of AIDA-1 and Caskin2 structural and functional studies.

1.3. The Sterile Alpha Motif (SAM) Domain

Over two decades ago, Ponting first identified SAM domains based on the predicted high-level alpha-helical content and significant sequence similarity among 14 eukaryotic proteins (Ponting, 1995). The conservation of helical structure and the fact that four of these proteins were linked to yeast sexual differentiation prompted the name sterile alpha motif (SAM) domain (Kim, 2003). Later, in 1997, Shultz *et al.* united a number of previously recognized modules named SPM (Scm, Ph, lethal-3 malignant brain tumor), SEP (yeast sterility, Ets-related, PcG proteins), PNT (pointed), NCR (N-terminal conserved region) and HLH (helix – loop – helix) under this name (Schultz *et al.*, 1997).

Currently, the SMART (Simple Modular Architecture Research Tool) database (Schultz *et al.*, 1998) identifies over 6000 SAM domain-containing proteins in mammalian genomes alone, and over 14000 in all genomes. SAM domain representation in the human genome (>200 by homology) comparable to the most common protein-protein interaction domains such as SH2 and SH3. SAM domains play diverse roles facilitating an array of protein-ligand interactions, and they generally exist as single units. However, tandem and even triplet (Taru & Jin, 2011) modules have been found in higher eukaryotes which suggests that they offer more complexity to the signaling interaction platform. SAM domains are often present in multidomain proteins of a remarkably versatile range of functions, including scaffolding and adaptor proteins, transcription and translational regulators, tyrosine kinases and serine/threonine kinases, and even nucleic acid and lipid binding proteins (Kim, 2003; Qiao, 2005).

1.3.1. Oligomerization States of SAM Domains Associate with Their Physiological Functions

The most common and well-characterized protein partnerships of SAM domains are other SAM domains. Although the SAM domain structures determined to date typically form conserved alpha helical bundles, the surface complementarity leads to a great variety of homotypic and heterotypic SAM-SAM interactions, especially in transcription factors (C. A. Kim, Gingery, Pilpa, & Bowie, 2002; C. A. Kim *et al.*, 2001) and neuronal signaling protein assemblies (Baron *et al.*, 2006; Bourgeron, 2009; Harada *et al.*, 2008). The ability to polymerize appears to be a fundamental property of tandem

and multiple SAM domain proteins and has been connected to their regulatory and structural roles in building cellular complexes. Knight *et al.* (2011) first used the term 'polymerizome' for SAM domain-mediated protein assemblies. This section aims to review a number of reported cases where the SAM domain oligomerization state was central to the activity of the protein within the biological system, whether it was transcriptional regulation or pre/post-synaptic scaffolding.

A transcriptional tumor suppressor, TEL (translocation Ets leukemia), was shown to self-assemble into a polymeric helix in a head-to-tail orientation with six SAM domains per turn (C. A. Kim *et al.*, 2001). Mutations specifically targeting polymerization contacts, specifically Lys99, increased nuclear retention and hindered TEL repression *in vitro* (Tran *et al.*, 2002; Wood *et al.*, 2003). Remarkably, two other transcriptional repression proteins from the polycomb group (PcG), polyhomeotic (Ph) and Sex-comb-on-midleg (Scm) form oligomers analogous to TEL although the intermolecular interface is supported by different residues (Kim *et al.*, 2002; Kim & Kim, 2005). Despite very modest sequence similarity, these all exhibit a left-handed head to tail helical type of polymeric architecture that is linked to their biological function. The loss of PcG repression *Drosophila* phenotypes were observed *in vivo* for mutations disrupting Scm SAM domain self-association. Moreover, negative phenotypes resulted from an overexpression of isolated Scm-SAM domain as a consequence of its competition with functional full-length Scm protein for its binding partners (Peterson *et al.*, 2004). The cooperativity of functions between Ph and Scm was suggested since both proteins colocalize with polytene chromosomes in fly and, in addition to self-oligomerization, can interact with each other

in a heterotypic manner via C-terminal SAM domains (Kyba & Brock, 1998; Peterson *et al.*, 1997). Altogether, accumulated experimental evidence suggests the existence of a mechanism in which PcG family of proteins provides a platform for chromatin modulation and long-term repression mediated by their polymerization states (Kim *et al.*, 2002; Peterson *et al.*, 2004).

Perhaps the most studied heterodimeric SAM domain interactions are scaffolding proteins Ste4 and Byr2 (*Schizosaccharomyces pombe*) that are involved in MAP kinase controlled yeast sexual differentiation and their orthologs in *Saccharomyces cerevisiae,* Ste11 and Ste50. Similar to the common polymerization mode described above, Byr2-SAM / Ste4-SAM (Ramachander & Bowie, 2004; Ramachander *et al.*, 2002) as well as Ste11 / Ste50 (Grimshaw *et al.*, 2004; Kwan *et al.*, 2004) utilize their EH/ML (end-helix/mid-loop) interface to form a heterodimer with nanomolar affinities. Amongst them only Ste11 demonstrated weak (~0.5mM) homodimerization potential (Grimshaw *et al.*, 2004). Three SAM domain containing proteins Mae, Yan and Pnt-P2 regulate transcription in *Drosophila* eye development controlled by receptor tyrosine kinase (RTK) pathways (Qiao *et al.*, 2006). Mae-SAM and Yan-SAM can form closed hetero-oligomers (Qiao *et al.*, 2004) resulting in Yan depolymerization and exposing MAPK phosphorylation site. Following the phosphorylation event, Mae is displaced by CRM1 permitting Yan nuclear export (Song *et al.*, 2005; Tootle *et al.*, 2003). The interaction between Mae-SAM ML surface and EH surface of Pnt-P2-SAM prevents MAPK phosphorylation by blocking its target surface and prevents transcriptional activation activity of Pnt-P2 as part of the MAPK signaling pathway regulation (Qiao *et al.*, 2006).

The scaffolding protein Shank organizes multiple proteins assemblies at the postsynaptic density (PSD). Shank family proteins (Shank1, Shank2, Shank3) have been associated with a number of neurodegenerative conditions termed idiopathic autism spectrum disorders (ASD) (Kim & Sheng, 2004; Sheng & Kim, 2000; reviewed in Jiang & Ehlers, 2013). Shank 3, in particular, defines the size and shape of dendritic protrusions, and deletion of various domains of Shank3 led to abnormal dendritic spine development in mice (Roussignol *et al.*, 2005). All Shank family proteins have a conserved single SAM domain at the C-terminus. Shank3 self-associates (Naisbitt *et al.*, 1999) into a helical polymer that is 70 Å in diameter, which in turn assembles into large sheets over 100nm wide (Baron *et al.*, 2006). The PDZ domain-mediated interactions with NMDA, AMPA and mGluRs receptors define the specific glutamatergic synaptic localization of Shanks (Jiang & Ehlers, 2013) and the SAM domain amplifies their presence by polymerization at the C-terminus (Hayashi *et al.*, 2009; Naisbitt *et al.*, 1999). It was established that in Shank3-SAM higher order assembly, Zn^{2+} plays an important role in stabilizing the formation of two-dimensional sheets of helical fibers (Gundelfinger *et al.*, 2006). Zn^{2+} stabilized salt bridges exist at the site where intra- and inter- Shank3-SAM polymer interfaces meet. Therefore, it was proposed that Zn^{2+}-mediated assembly and packing density of Shank oligomers contributes to mechanisms regulating synaptic formation, maturation and structural plasticity (Baron *et al.*, 2006). Examples of SAM domains employing common structural interfaces for homo- and heterotypic interactions are presented in **Figure 1.1.**

Overall, these examples demonstrate that many SAM domain proteins use their surface complementarity to produce an array of oligomeric states (Ramachander & Bowie, 2004) as well as a strong basis for heterotypic SAM-SAM domain associations. The oligomeric state is often a driving force for a specific cellular function such as transcriptional repression or pre/post-synaptic scaffolding or playing part of a greater regulatory mechanism such as synaptic plasticity.

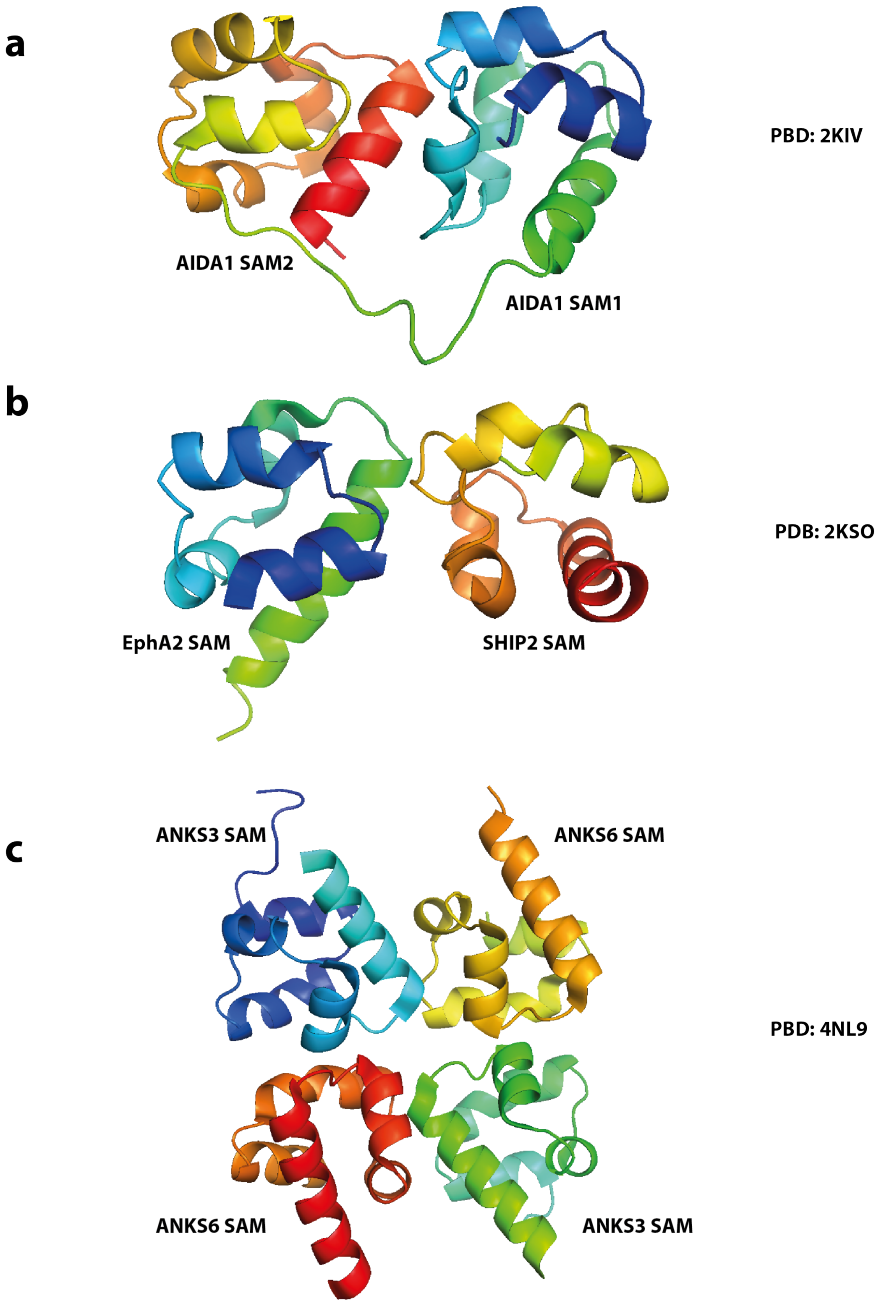


Figure 1. 1 – Structural diversity of SAM domains. Representative examples of SAM-SAM homo- and heterotypic interactions. **(a)** AIDA-1 SAM1-SAM2 tandem (PDB: 2KIV) intramolecular interactions; **(b)** heterotypic EphA2:SHIP2 SAM:SAM complex (PBD: 2KSO); **(c)** human Anks3-SAM/Anks6-SAM heterooligomeric packing (PBD: 4NL9). The cartoon representations were generated using MacPyMOL program with color scheme from blue (N-terminus) to red (C-terminus).

1.3.2. SAM Domain Interactions Beyond SAM-SAM Type

C-terminal SAM domains are also present in all Eph family receptor tyrosine kinases (RTKs) (Pawson & Nash, 2000). Eph family receptor tyrosine kinases mediate axonal pathfinding, neuronal cell migration, angiogenesis and capillary morphogenesis (Smalla *et al.*, 1999; Stapleton *et al.*, 1999). Although the biological significance of homo-dimerization and oligomerization of Ephrin SAM domains remain unclear (Stapleton *et al.*, 1999; Thanos, 1999) its potential impact on cell signaling via receptor clustering is anticipated (Qiao, 2005). EphA RTKs attenuate cell migration (Borthakur *et al.*, 2014). Structural studies, supported by *in vitro* experiments, have reported the requirement for phosphorylation of two tyrosine residues in EphA2 SAM domain for it to be recruited by the adaptor protein Grb7 [also by both Grb7 and Grb10 in the case of EphB1 RTK (Han *et al.*, 2002)]. A reverse signaling event, dephosphorylation of EphA2 at Tyr930 by receptor protein tyrosine phosphatase LAR, decouples it from another adaptor protein, Nck1, likewise involved in EphA-mediated cell migration (Hu *et al.*, 2009; Lee & Bennett, 2013). A number of recent studies elucidated the details of the heterotypic association of EphA2 and the lipid phosphatase Ship2 that leads to inhibition of receptor endocytosis as well as enhancement of Eph kinase activation (Lee *et al.*, 2012; Leone, Cellitti, & Pellecchia, 2008; Zhuang *et al.*, 2007).

Traditionally, SAM domains were viewed solely as protein-protein interaction modules until RNA binding SAM domains of Smaug and its ortholog Vts1 emerged in the literature (Aviv *et al.*, 2006; 2003; Green *et al.*, 2003). The primary function of Smaug

is morphogen gradient control by repression of Nanos mRNA translation in developing *Drosophila* embryos. The surface exposed positively charged cluster (Green *et al.*, 2003) of the Smaug SAM domain directly targets non-stem-loop RNAs at hairpins termed the Smaug recognition element (SRE) (Johnson & Donaldson, 2006; Oberstrass *et al.*, 2006). Essentially, Smaug fulfills the function of a translational switch that controls Nanos distribution and proper abdominal segmentation in the early embryogenesis stages in *Drosophila* (Dahanukar, Walker, & Wharton, 1999).

1.3.3. Phosphotyrosine Binding Domain (PTB)

The PTB domain along with the SH2 domain were originally classified as a phosphotyrosine interaction domains (PID) that play an active part in phosphotyrosine related cell signaling. In the human genome, there are approximately 60 proteins that have PTB domains (Uhlik *et al.*, 2005), and while they are also present in *Drosophila* and *C. elegans* genomes, and none found in *Arabidopsis thaliana* or *S. cerevisiae* (Yaffe, 2002).

The SH2-domain-containing adaptor molecule (Shc) and the docking protein insulin receptor substrate-1 (IRS-1) were the first two proteins where PTB domains were independently identified (Yaffe, 2002) as NPXpY-motif binding modules. Successively two PTB-domain neuronal proteins, Fe65 and X11/Mint (Borg *et al.*, 1996; Zambrano *et al.*, 1997), well-known nowadays as amyloid precursor protein (APP) partners, have been shown to specifically target the non-phosphorylated NPTY sequence of the APP intracellular domain, suggesting that the PTB module has a separate set of functions aside from kinase signaling. Numerous examples of PTB domain substrate recognition through

non-canonical NPXY sequences have been identified and described (reviewed in Uhlik *et al.* 2005). Furthermore, examples of PTB-binding motifs that lack tyrosine residues entirely include: the PTB domain of cell-fate-determinant protein Numb interaction with Numb-associated kinase Nak; Protein Kinase C3 (PKC3) and Numb-interacting protein (NIP); hSNT-1 and -2 PTB binding to fibroblast Growth Factor Receptor 1 (FGFR1); mammalian protein Disabled (Dab) and Protocadherin 18 (Pcdh 18) (Uhlik *et al.*, 2005; Yaffe, 2002). The PTB domain of Shc (Ravichandran *et al.*, 1997) and mDab (Howell *et al.*, 1999a) have been reported to interact with phospholipids (Yan, Kuti, & Zhou, 2002a) which further contributes to the complexity and versatility of PTB supported functions.

1.3.4. Classification and Structural Features of PTB Domains

Traditionally, PTB domains were segmented into two groups based on structural organization and ligand binding specificity: two phosphotyrosine-dependent, Shc-like and IRS-like, more recently Dab-like phosphotyrosine-independent PTB was defined as a separate subgroup (Forman-Kay & Pawson, 1999; Uhlik *et al.*, 2005; Yaffe, 2002). Unlike the SH2 domain family that is characterized by high sequence similarity, PTB domains exhibit a surprisingly low level of sequence conservation. Nonetheless, they adopt a similar structural fold (also referred as pleckstrin homology (PH) domain superfold (Uhlik *et al.*, 2005) that consists of an orthogonal β-sandwich capped with a C-terminal α-helix. The conserved glycine-based loop differentiates the Shc group from IRS and Dab. In addition, Shc and Dab1 PTB domains lack a β1' strand, but contain two additional α-helices, one N-terminal from the β-sandwich and another between β1 and β2

(Uhlik *et al.*, 2002). Despite the structural differences all PTB domains have a distinct peptide binding pocket lined up by residues from the β5 strand and C-terminal α helix. The other common structural feature is a highly basic phospholipid-binding surface formed by the N-terminal loops (Uhlik *et al.*, 2005). A representative example from each group is shown in **Figure 1.2.** The structural foundations of PTB domains distinct ligand binding modes will be discussed in greater details in the following section.

1.3.5. Structural Foundation for Canonical NPX(Y/Py) Versus Non-Canonical Peptide Binding and Interactions with Phospholipids

As noted earlier, PTB domains engage multiple ligands highlighting their functional significance and evolution as a modular binding domain. A typical PTB ligand recognition site extends to the amino acids N-terminal from pY; while in the case of SH2 domains residues carboxy-terminal to the phospho-tyrosine define ligand specificity (Uhlik *et al.*, 2005; Yaffe, 2002). With the accumulation of substantial structural data on ligand-bound PTB domains, the general basis for the "classic" PTB NPX(p)Y motif recognition has emerged (Eck *et al.*, 1996; Shi *et al.*, 2002; Stolt *et al.*, 2003; Z. Zhang, 1997; Zhou *et al.*, 1996; 1995; Zwahlen, Li, Kay, Pawson, & Forman-Kay, 2000) *et cetera*. A consensus Asn-Pro-X-Tyr motif adopts a type-I β-turn conformation which ensures precise positioning of the Y residue in the L-shaped hydrophobic binding groove of PTB also called the "anchoring pocket". The stretch of N-terminal residues forms a pseudo anti-parallel β-sheet through hydrogen bonding with the β-5 strand and the C-terminal α-helix (Uhlik *et al.*, 2005; Yaffe, 2002). One of the earliest solved pY-ligand bound structures, IRS-1 (Eck *et al.*, 1996; Zhou *et al.*, 1996) and Shc (Zhou *et al.*, 1995),

demonstrates the phosphorylated tyrosine coordination by basic residues which explains the high-affinity binding observed for phospho-tyrosine ligands (Farooq *et al.*, 1999; Wolf *et al.*, 1995). In the case of non-phosphorylated ligands similarly tight binding is achieved through compensation by extended recognition of carboxy-terminal residues adjacent to tyrosine (or non-canonical phenylalanine) in addition to amino-terminal residues (Forman-Kay & Pawson, 1999; Yaffe, 2002).

A closer look at Dab-like non-phosphorylated ligand binding specificity is essential as it is most relevant to my structural study of the APP interaction with the AIDA-1 PTB domain (described in **Chapter 2**). With the Dab-like PTB mode of binding, phosphorylation of tyrosine is not only unnecessary but in most cases specifically disfavored since it perturbs the binding (Howell *et al.*, 1999; Yun *et al.*, 2003; Z. Zhang, 1997). The role of a hydrophobic anchoring pocket is less prominent and Tyr at position 0 in X11-APP complex (Z. Zhang, 1997) or Phe in Numb-NAK complex (Yun *et al.*, 2003) does not play a central role in sharp contrast to Shc-like and IRS-like type of ligand recognition. Furthermore, the X11 PTB-APP interaction remains preserved even when Tyr0 was substituted by Ala (Borg *et al.*, 1996). On the other hand, a large number of hydrophobic interactions and hydrogen bonds occur between β-5 strand and the peptide binding sequences in both N-terminal and C-terminal directions from position 0. In the case of Fe65 interaction with the APP intracellular domain (AICD), the minimal AICD peptide binding motif extends to as much as 32 amino acids and aligns the entire PTB binding groove. In fact, the Fe65 PTB2 domain demonstrated ~100-fold difference in affinity between an 11 aa. minimal sequence (K_d = 100 mM) and an amino-terminally

extended 32 aa. (K_d = 0.2 mM) (Mulvihill & Komives, 2011; Radzimanowski *et al.*, 2008). Another unique aspect of the Fe65-APP interaction is that the phosphorylation on threonine preceding tyrosine acts as a conformational switch that forces proline (P669) to transition from *trans* to *cis* conformation thereby precluding APP ligand binding (Radzimanowski *et al.*, 2008). To add another layer of complexity, the PTB domain of SNT1 (also referred as FRS2 -fibroblast growth factor receptor substrate 2) was shown to engage two completely different ligands: first, the TrkA (Tropomyosin receptor kinase A) receptor which is classified as IRS-type of binding and second, FGFR1 which not only lacks the NPTY motif but does not have any Asp, Tyr or Phe within the minimal binding sequence (Dhalluin *et al.*, 2000; Yan *et al.*, 2002). Another exception is Talin which interacts with the cytoplasmic tail of β3 integrin. Although the binding mode structurally resembles the IRS-type, it occurs in a phosphorylation-independent manner. Furthermore, the Tailin PTB binding pocket lacks both arginine residues typical for IRS mode of binding (García-Alvarez *et al.*, 2003).

Another characteristic feature of PTB as an interaction hub is its phospholipid binding capability. It is achieved through a surface-exposed highly basic cluster distinct from the peptide binding pocket described earlier (**Figure 1.2**). The numerous examples of direct binding to liposome-associated or free phospholipid head groups (Uhlik *et al.*, 2005) reported to date, including Dab1, Dab2, X11, Numb, ARH, IRS-1, Shc, and Talin PTB domains (Dho *et al.*, 1999; Martel *et al.*, 2001; Mishra *et al.*, 2002; Okamoto & Sudhof, 1997; Ravichandran *et al.*, 1997; Takeuchi *et al.*, 1998; Yun *et al.*, 2003) suggest the evolutionary conservation of lipid-binding function in PTB domains.

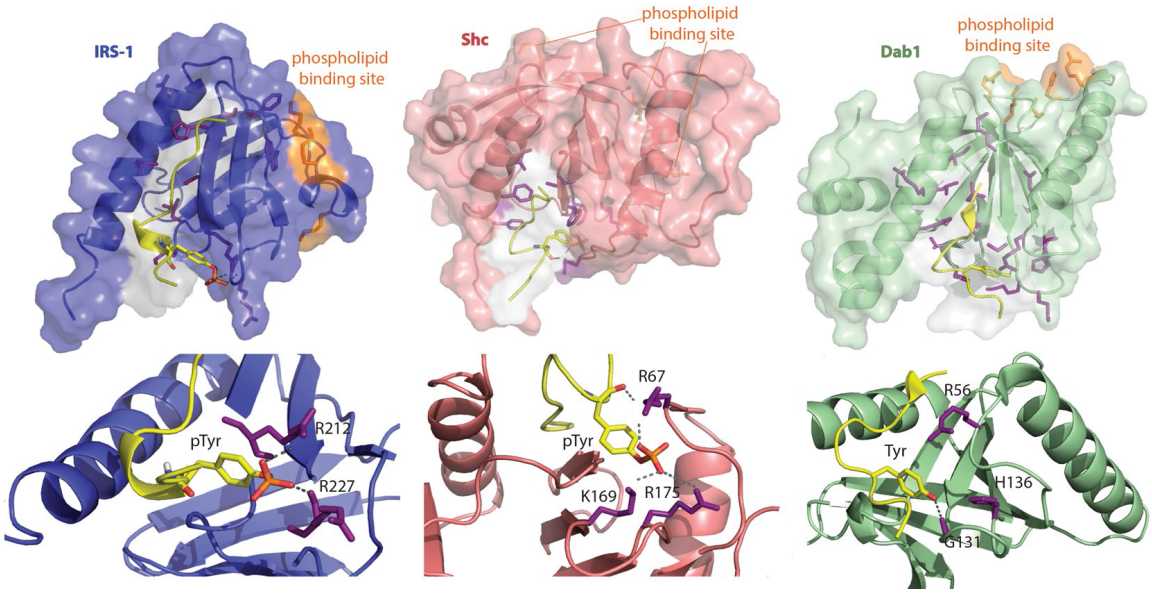


Figure 1.2 – The representative PTB structures of IRS-like, Shc-like and Dab-like families. *Top panel:* Structures of IRS-1 PTB in complex with the IL-4R peptide (PDB: 1IRS), Shc PTB in complex with the TrkA receptor peptide (PDB:1SHC), and Dab1 PTB in complex with ApoER2 receptor and PI-4,5P_2 (PDB: 1NU2). The residues contributing to PTB binding cleft are shown in purple; the basic surface exposed side chains forming phospholipid-binding clusters highlighted in orange (in both stick and molecular surface formats); and the ligand peptides (in yellow, cartoon representation). *Bottom panel:* close-up view of the tyrosine- or phospho-tyrosine-coordinating residues of each PTB domain. Dashed lines represent electrostatic interactions anchoring pTyr/Tyr in the binding pockets of IRS-1, Shc and amino acids at corresponding positions in Dab-1. Figure revised from (Uhlik *et al.*, 2005).

1.3.6. PTB Domains in Signal Transduction

Remarkably, while the majority of the PTB domains exist along with multiple interaction domains within a protein, close to a third of PTB domain containing proteins do not have any other defined modules (**Figure 1.3**). This, combined with the absence of catalytic activity, solidifies them as ultimate adaptor and scaffolding proteins (Uhlik *et al.*, 2005). In line with the classification outlined in the previous section, distinct groups of PTB domain proteins fulfill different cellular functions. They are involved in tyrosine kinase, cytokine receptor signaling; APP regulation; integrin related cell adhesion; asymmetric cell division and low-density lipoprotein (LDL) controlled endocytosis (Cesareni *et al.*, 2005; Uhlik *et al.*, 2005; K. S. Yan, Kuti, & Zhou, 2002a). One of the most well-studied adaptor proteins contains an SH2 domain in addition to a PTB domain. The Shc PTB connected to activated growth factor receptors undergoes phosphorylation which in turn permits Gb2 adapter recruitment followed by Ras (nucleotide exchange factor) mediated MAPK pathway activation (Ravichandran, 2001). Shc has been reported to bind as many as 15 different phospho-tyrosine activated growth factor and cytokine receptors (Uhlik *et al.*, 2005). IRS and Dock family proteins are notably involved in various signal transmission assemblies, including insulin receptor, B-cell receptor, CD2 and Eph family receptors mediated signaling (Cesareni *et al.*, 2005; Yaffe, 2002).

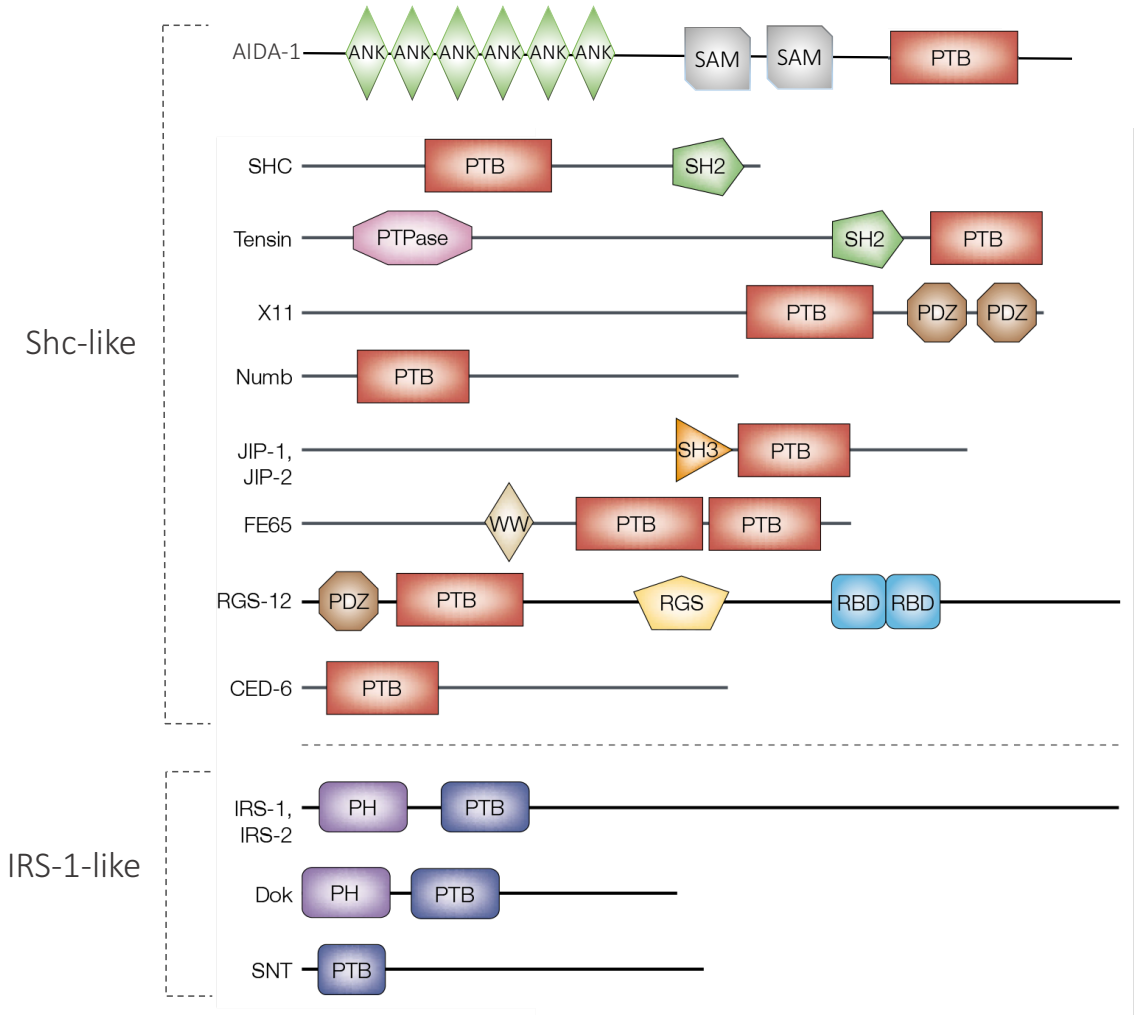


Figure 1.3 – Modular architecture of PTB-domain-containing proteins. *Top panel* depicts Shc-like PTB domain family representatives (colored terracotta), including (p)Tyr-independent Dab-like subfamily, whereas *bottom panel* illustrates IRS-1-like PTB domain family representatives (colored violet).

Domain annotations: ANK, ankyrin repeats, SAM, sterile alpha motif; IRS-1/2, insulin-receptor substrate-1/2; JIP, c-Jun amino-terminal kinase-interacting protein; PDZ, domain present in PSD-95, discs large and ZO-1; PH, Pleckstrin homology domain; PTB, phosphotyrosine binding; PTPase, protein tyrosine phosphatase; RBD, Raf-like Ras-binding domain; RGS, regulator of G-protein-signalling domain; SHC, Src-homology-2-containing transforming protein; SNT, suc1-associated neurotrophic-factor target; SH2, Src-homology-2; SH3, Src-homology-3; WW, domain with two conserved tryptophan (W) residues. Figure adapted from (Yaffe, 2002) with the addition of AIDA-1 schematic diagram.

The numerous PTB adapters, including Shc, JIP, CED-6, ARH, Fe65, X11, CAPON, ICAP [reviewed in (Uhlik *et al.*, 2005)] have been found to bind low-density lipoprotein receptor family proteins and are believed to directly affect the regulation of LDL receptor (LDLR) endocytosis. Furthermore, two of these proteins, Fe65 and X11, are also involved in amyloid precursor protein (APP) processing (Cesareni *et al.*, 2005; King & Scott Turner, 2004). Fe65 employs two of its PTB domains for simultaneous recruitment of both APP and lipoprotein receptor LRP intracellular tails and is linked to upregulation of APP processing and generation of amylogenic fragments that are known determinants of Alzheimer's disease pathogenesis (McLoughlin & Miller, 2008). Conversely, X11 is believed to have a positive regulatory function in APP exocytosis and an inhibitory role in its endocytosis (Cesareni *et al.*, 2005; King & Turner, 2004; Sakuma *et al.*, 2009). It is not surprising that many PTB domain proteins are involved in integrin-cytoskeleton signaling pathways that regulate focal adhesion development (Calderwood *et al.*, 2003) since β-integrin tails offer double NPxY motifs for their recognition. Finally, Numb is a cell-fate determinant proposed to function downstream of transmembrane receptor Notch, which activates transcription of many cell-fate mediating genes. The Numb "recycling inhibition" model of Notch signaling through regulation of Notch-Sanpodo oligomers trafficking has been most recently proposed (Couturier, Mazouni, & Schweisguth, 2013).

Thus, the versatility of PTB domain functions has been proven repeatedly with numerous example that highlight evolutionary complexity of cell signaling and the critical role of protein-protein interaction modules in maintaining specificity and selectivity in

signaling cascades.

1.4. Multidomain Scaffolding Proteins

Scaffold, adaptor, anchoring proteins support complexity and specificity required for signal transduction pathways. Scaffolding proteins do not typically possess an enzymatic function, yet their role as molecular hubs is particularly important in organizing numerous protein complexes of signaling machinery. The presence of multiple modular domains, such as in Fe65, Mint, AIDA-1, Numb, MAGUK family, Caskins, Shank, Mint, SARM, Liprin-α scaffolding proteins, permits a panoply of protein partner recognition. Therefore, modular proteins serve as an assembly platform for protein signaling networks (Pawson & Nash, 2000; Pawson & Scott, 1997). In addition to the distinct protein interaction domains, intrinsically unstructured segments of scaffolding proteins are known to contribute to their functional specificity by carrying various post translational modifications, recruiting ligands and contributing to structural versatility (Van der Lee et al., 2014). Scaffolding proteins do not only passively serve as molecular docking sites, organizing multiple protein partners; they often carry regulatory modifications and actively mediating signaling cascades (Burack & Shaw, 2000; Smith & Scott, 2013).

Mitogen-activated protein kinase (MAPK) cascades are perhaps the most extensively studied signal transduction scaffolds. MAPK signaling pathways regulate many cellular processes, including cell differentiation, trafficking, division and apoptosis (Garrington & Johnson, 1999; Schaeffer & Weber, 1999). Growth factor receptor-bound protein 2 (Grb2), Ste5 (in yeast), mammalian kinase suppressor of Ras (KSR), guanine

exchange factor (GEF) and JNK-interacting protein (JIP) are scaffolding organizers of multi-MAP kinase assemblies crucial for initiation of MAPK phosphorylation cascades (Meister *et al.*, 2013; Smith & Scott, 2013).

Adaptor and scaffolding proteins maintain signal specificity, subcellular co-localization of signaling scaffold components and amplify signal transduction. More complex regulatory roles of scaffolds such as Ste5p. Ferrell *et al.* (2000) extend their function to the regulation of the strength of signaling response in either a graded or switch-like manner, in which case overexpression of the scaffolding protein could lead to a lower signaling output. For instance, the KSR-1 scaffold mediates Ras signaling response favorably at low and negatively at high expression levels (Burack & Shaw, 2000; Ferrell, 2000). Smith and Scott, (2013) reviewed the number of studies that demonstrated the phosphorylation-dependent recruitment of signaling components by scaffolding proteins, where modification of scaffolding proteins themselves defined the outcomes of the signaling process they regulate. For example, the Wnt-β-catenin signaling pathway regulated gene expression relies on the Axin scaffold. Phosphorylation of Axin promotes destruction complex formation in β-catenin degradation, as part of the "destruction" pathway; and dephosphorylated Axin complex facilitates the stabilization pathway and destabilization of the destruction complex, leading to β-catenin nuclear accumulation where it serves as a transcription co-activator of Wnt responsive genes. This phospho-dependent switch is regulated by glycogen synthase kinase 3β (GSK3β), PP1cγ protein phosphatase and its inhibitor (Kim *et al.*, 2013; Smith & Scott, 2013).

It is not surprising that oligomerization is an intrinsic property and one of the

regulatory mechanisms of scaffolding proteins. In the context of cellular signaling, repeated rounds of a quick scaffold assembly, dissociation and reinitiating is required. Therefore, scaffolding proteins often present themselves as multimeric complexes (Pan *et al.*, 2012). Sterile alpha motif, as discussed earlier, is one of the most common homo- and hetero- oligomerization modules of scaffolding proteins.

A family of neuronal scaffolding proteins, including Homer/Velis, PSD95, Shank, CASK and Caskin are indispensable for proper synaptic function. Homer/Velis link neurotransmitter receptors to intracellular effectors and postsynaptic density (PSD) scaffolds; Shank and PSD95 are implicated in receptor trafficking, receptor organization at the surface plasma membrane and activity-dependent PSD remodeling (Iasevoli, Tomasetti & de Bartolomeis, 2013). Signal transduction utilizing multiple protein domains could be exemplified by the MAGUK family of neuronal scaffolding protein CASK ($Ca^{2+/}$calmodulin-associated Ser/Thr kinase). SH3 and guanylate kinase domain (GK) compose a structural supramodule (Zhu *et al.*, 2011) with GK providing the binding pocket specific for either pSer- or pThr-containing targets, while PDZ domain is responsible for CASK multimeric assembly via complementary surfaces (Rademacher *et al.*, 2013).

Evidence from genetic, biochemical and clinical studies promote PSD-associated scaffolding proteins as important determinants of synaptic plasticity associated with learning and memory formation (Iasevoli *et al.*, 2013). In this respect, the mechanistic understanding of neuronal scaffolding protein functions is necessary for development of

therapeutic strategies treating/controlling neurodegenerative disorders such as schizophrenia, dementia, and autism. Structure-directed studies of neuronal scaffolding proteins AIDA-1 and Caskin2 are therefore the foci of this dissertation. The following sections aim to introduce the structural attributes of the two proteins and the context of their plausible neuronal functions.

1.5. Amyloid Beta Precursor Intracellular Domain-Associated Protein-1 (AIDA-1) a Major Synaptic Scaffolding Protein with Multiple Functions

The AIDA-1 protein was originally discovered in a yeast-two hybrid screen as a new interacting partner of the C-terminal cytoplasmic fragment of a membrane-associated amyloid precursor protein (AβPP) (Ghersi *et al.*, 2004). This fragment, termed AβPP intracellular domain (AID, or AICD), is produced by the subsequent proteolytic processing of AβPP that also generates a short polypeptide fragment β-amyloid (or Aβ). Aβ aggregation in the brain leads to plaque formation and is one of the determinants of Alzheimer's disease (AD), a neurodegenerative disorder characterized by loss of synapses and neuronal cell death (Awasthi *et al.*, 2009). AIDA-1 is encoded by a single gene, *ANKS1*, on human chromosome 12. Five different isoforms of AIDA-1 depicted in **Figure 1.4** as a result of alternative splicing have been identified (Ghersi *et al.*, 2004).

In addition to the protein-protein interaction domain module composed of two SAM domains and a single PTB domain that is present in all isoforms, the longest isoform AIDA-1b also includes six N-terminal ankyrin repeats. A similar supramodular layout of multiple protein-protein interacting domains was also found in two neuronal proteins,

Mint1/X11 and Fe65, that regulate AβPP processing and intracellular trafficking (Dumanis *et al.*, 2012; Han *et al.*, 2016b etc.; Saito *et al.*, 2011; Sakuma *et al.*, 2009; X. Xie *et al.*, 2012). In addition to a regulatory role in APP processing originally proposed by Ghersi *et al.* (2004), AIDA-1 has since been implicated in Cajal body regulation, nucleolar formation and stability, nuclei-to-synapse signaling and PSD structural remodeling long term memory formation and synaptic plasticity (Xu & Hebert, 2005; Dosemeci *et al.* 2015, Jordan *et al.* 2009, Tindi *et al.*, 2015). Recognized AIDA-1 protein partnerships in conjunction with associated pathways and proposed functions will be discussed in the following subsections.

1.5.1. AIDA-1 Associates with APP Intracellular Domain

Similar to Mint/11 and Fe65, AIDA-1 interacts with APP through its PTB domain in a phosphorylation-independent manner characteristic of the Dab-like PTB domain family. In our study, we employed a combination of biochemical methods to characterize the interaction between AIDA-1 PTB and APP and reported the AIDA-1 PTB NMR structure that permitted AIDA-1-APP ligand-binding by molecular modeling. This study resulted in the following publication: Smirnova E., Shanbhag R., Kurabi A., Mobli M., Kwan J.J., Donaldson L.W. (2013) *Solution Structure and Peptide Binding of the PTB Domain from the AIDA-1 Postsynaptic Signaling Scaffolding Protein.* PLoS ONE 8(6) which will be described in **Chapter 2.**

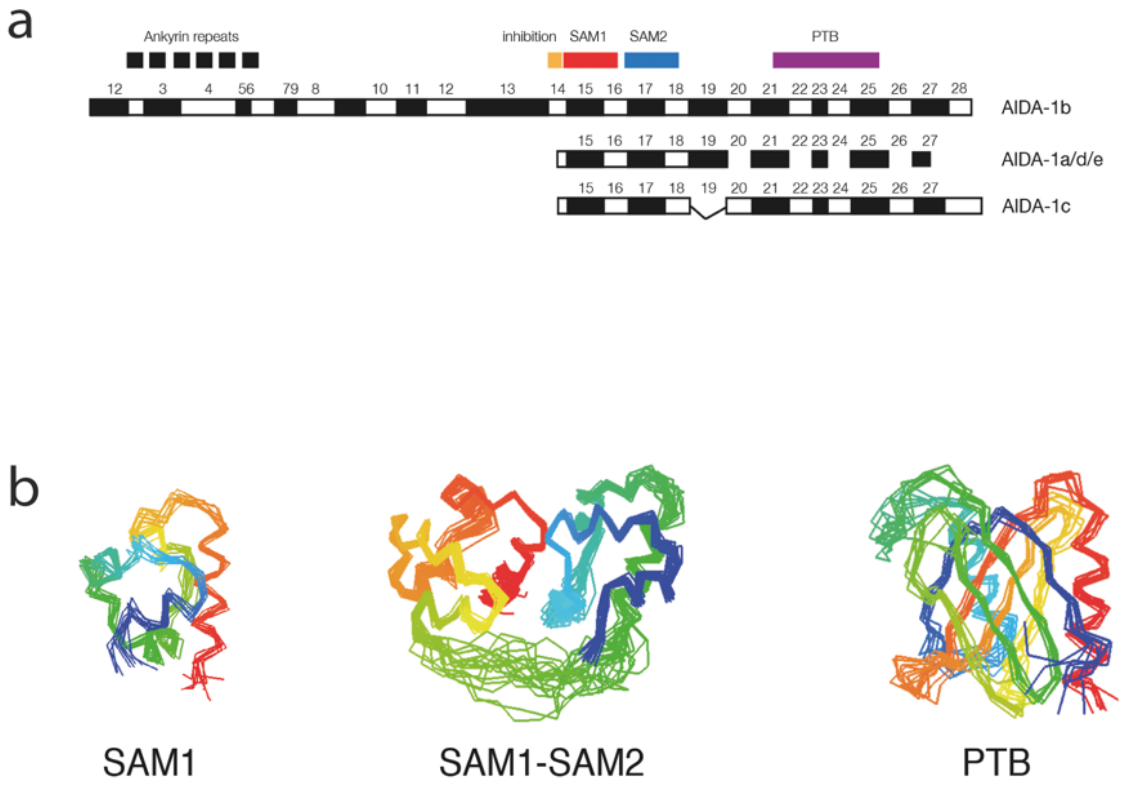


Figure 1.4 – AIDA-1 domain organization. **(a)** Five splice variants of AIDA-1. All isoforms contain double sterile alpha motif and phosphotyrosine binding domain. The longest isoform differs by presence of stretch of N-terminal ankyrin repeats and putative self-inhibition region encoded by exon 14. (**b)** Ribbon presentations of NMR solution structures of SAM1 (PDB: 2KE7), SAM1-SAM2 (PDB: 2KIV) and PTB (2M38) domains.

AICD is a 6kDa intracellular fragment that has held a research spotlight for over two decades, as part of effort to elucidate APP pathogenic and nonpathogenic pathways and to develop AD therapeutic strategies. Since the foundation of the amyloid cascade hypothesis [the chain of events neurotoxic effects leading to the cell death and neuronal degeneration in AD (Selkoe, 1991; Selkoe & Hardy, 2016)] more than twenty protein partners, including AIDA-1, have been observed to associate with AICD, reviewed in (Raychaudhuri & Mukhopadhyay, 2007) thereby (Multhaup *et al.*, 2015), linking it to diverse cellular processes including APP processing and trafficking (Lorenzo *et al.*, 2000; Marks & Berg, 2010), transcriptional regulation, (Cao & Sudhof, 2004), apoptosis (Passer *et al.*, 2000), calcium homeostasis (Hamid *et al.*, 2007) and AD pathogenesis, physiological and behavioral hallmarks (Beckett *et al.*, 2012; Galvan *et al.*, 2006).

Many scaffolding and protein adaptors employ PTB or SH2 domains to recognize a YENPTY motif of AICD. In contrast to AIDA-1, the roles of X11/Mint1 and Fe65 in APP metabolism are more characterized. Fe65 facilitates non-phosphorylated AICD transport to the nucleus (Bórquez & González-Billault, 2012; Cao & Sudhof, 2001) where in conjunction with Fe65, MED12 and transcription factors Tip60 or CP2/LSF/LBP1 AICD regulates gene expression acting as transcription co-factor (Beckett *et al.*, 2012; Minopoli *et al.*, 2001; Cao & Sudhof, 2001) [refer to **Figure 1.5**].

In contrast, the X11 protein family (X11s) is believed to have a positive regulatory function in APP exocytosis and an inhibitory role in its endocytosis (Cesareni *et al.*, 2005; King & Turner, 2004; Sakuma *et al.*, 2009). X11s regulate APP by at least two distinct events. Firstly, by suppression of APP maturation and direction of immature APP

(imAPP) to the early secretory pathway (Han *et al.*, 2016b; Saito *et al.*, 2011). Second, as observed in the late protein secretory pathway, by direct binding to AICD via its PTB domain X11 prevents γ-secretase cleavage (Saito *et al.*, 2008; Sakurai *et al.*, 2008), illustrated in **Figure 1.5**. Therefore, expression of X11 favors an overall metabolic stabilization of APP, promotes imAPP intracellular accumulation, and consequently reduces Aβ secretion (Saito *et al.*, 2011). The phosphorylation-dependent conformational switch mechanism has been proposed (illustrated in **Figure 1.6**) in which the phosphorylation on a threonine residue (Thr668), located 14 aa. N-terminally from the NPTY-motif anchoring tyrosine, forces proline (Pro669) to transition from *trans* to *cis* conformation thereby preventing Fe-65-APP binding (Ando *et al.*, 2001; Radzimanowski *et al.*, 2008; Suzuki & Nakaya, 2008). At the same time this conformational change does not prevent X11s from binding to the GYENPTY motif (Suzuki & Nakaya, 2008).

Noticeably, AIDA-1 isoforms have different subcellular localizations: while the longest AIDA-1b is normally found in the cytoplasm, AIDA-1 a/c/d/e have a predominantly nucleolar distribution. The short AIDA-1a isoform demonstrates preferential binding to AICD AβPP (cytoplasmic domain of APP) and reduces amylogenic fragment secretion by preventing γ-secretase cleavage, while isoform b was found to be AβPP-inactive (Ghersi *et al.*, 2004a). Moreover, later studies by our group supported observations by Ghersi *et al.* that the short region (~24 aa.) encoded by exon 14 prevented AIDA-1 from associating with AβPP which suggests the possibility of an AIDA-1 intrinsic self-inhibitory mechanism that directly impacts any AβPP related functions.

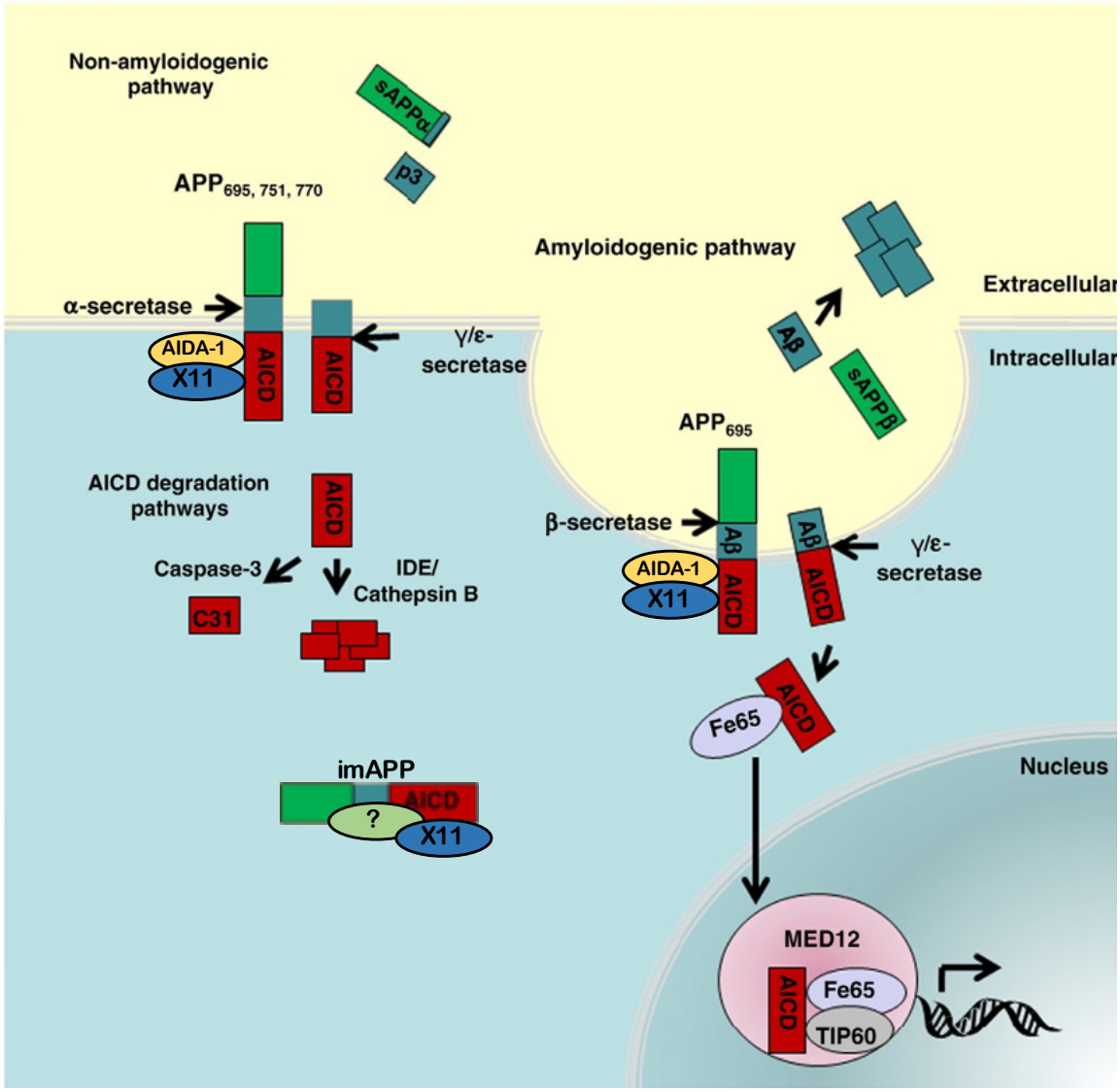


Figure 1.5 – The schematic diagram of APP alternative processing pathways. APP is subjected to α- or β-secretase cleavage resulting in soluble N-terminal fragments sAPPα or sAPPβ. The successive cleavage of remaining membrane-bound C-terminal fragments by γ-secretases to release a non-toxic p3 peptide and AICD (non-amyloidogenic) or Aβ and AICD (amyloidogenic pathway involves endocytosis). An alternative non-amylogenic pathway involves AICD degradation by the insulin-degrading enzyme (IDE) and cathepsin B or caspase-3. Fe65 facilitates AICD transport to the nucleus where in conjunction with Fe65, MED12 and Tip60 AICD regulates gene expression. X11 and AIDA-1 competitive binding to AICD suppress the γ-secretase cleavage. X11 also facilitates retention of imAPP in the cytoplasm, although the exact mechanism is unknown and may involve additional protein effectors. imAPP = immature APP. Revised from (Beckett *et al.*, 2012) with the addition of AIDA-1 and X11 APP-interacting proteins in line with the context of this thesis.

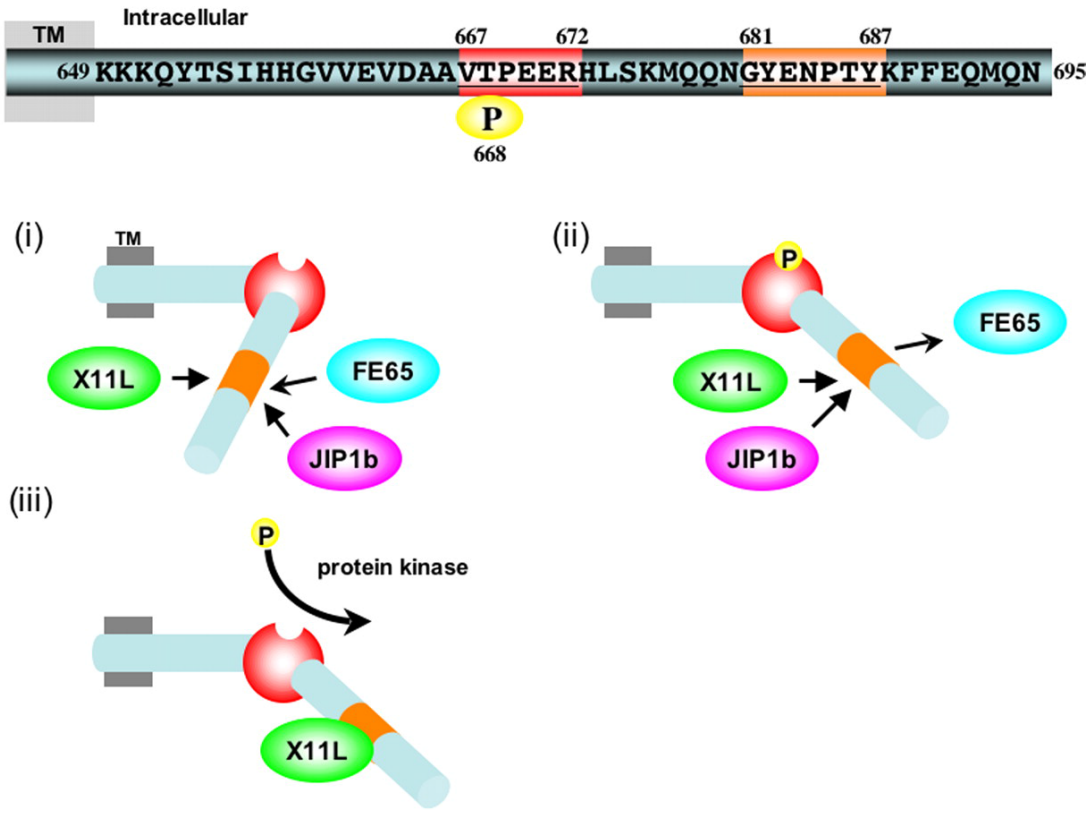


Figure 1.6 – A schematic representation of the APP cytoplasmic region conformational switch mechanism. Amino acid sequence of the APP cytoplasmic region presented on top of the panel (residue numbering based on APP695 isoform). Two motifs involved in the conformational switch mechanism, 667-VTPEER-672 and 681-GYENPTY-687 are highlighted and used in the panels below to indicate these regions in the schematic molecule models. Panel (*i*) The 681-GYENPTY-687 is a recognition motif for APP binding partners X11L, FE65, and JIP1b. Panel (*ii*), phosphorylation at Thr668 acts as a conformational switch and results in FE65 release from the 681-GYENPTY-687 motif (Ando *et al.*, 2001; Ramelot & Nicholson, 2001). Panel (*iii*), binding of X11L to the 681GYENPTY687 motif may facilitate the conformational change that leads to the upstream Thr668 exposure and susceptibility for phosphorylation by protein kinases such as JNK (Taru & Suzuki, 2004). TM, transmembrane domain. Figure adapted from (Suzuki & Nakaya, 2008).

Autoinhibition and ligand-induced competitive modes are common regulatory features of scaffolding proteins. For example, a Mint1(X11α) structural study revealed Mint1 phosphorylation regulated self-inhibition where the C-terminal linker adopts an α-helical structure and blocks the PTB binding cleft, hence, preventing APP from binding (Matos *et al.*, 2012). Interestingly, Mint2(X11β) was reported to have a different self-regulatory mechanism through an open to close conformational switch (Xie *et al.*, 2012). Other examples include integrin regulating Talin autoinhibition (Goksoy *et al.*, 2008), scaffolding protein GRIP1 (Long *et al.*, 2008) and PTP-BL tyrosine phosphatase (Van den Berk *et al.*, 2007) allosteric inhibition.

1.5.2. AIDA-1 Structural Role at PSD

AIDA-1 is recognized as a prominent protein of postsynaptic densities (PSDs) (Jacob *et al.*, 2010; Jordan, 2004) which are cellular protein-rich substructures that coordinate postsynaptic signal transduction that have been associated with mechanisms sustaining synaptic plasticity (Dosemeci *et al.*, 2015; Jacob *et al.*, 2010; Sheng & Hoogenraad, 2007). Recently, a quantitative mass spectrometry study reported the stoichiometric distribution of AIDA-1 at the PSD as 1:1:2 relative to total GKAP guanylate kinase-associated proteins and PSD-95 respectively (Lowenthal, Markey & Dosemeci, 2015). The fact that both GKAP (guanylate kinase-associated protein) and AIDA-1 directly interact with PSD-95 (Jordan *et al.*, 2007) and all three are found in the same electron-dense PSD layer (Jacob *et al.*, 2010) suggests that AIDA-1 serves an important structural function at the PSD core. Activity-induced PSD reorganization is

believed to be part of a complex mechanism regulating synaptic plasticity via fluctuations in synaptic strength. It also involves NMDA receptor reorganization, active protein trafficking, and activity-induced synaptic protein degradation. A recent research reports the reversible translocation of AIDA-1 outside of the deep PSD core under excitatory conditions, similar to another PSD member, SynGAP (Dosemeci *et al.*, 2015). While PSD-95 and GKAP proteins maintain the same positions at PSD layers, other proteins, such as CaMKII, Shank and CYLD, tend to cluster in the denser PSD zone. Since AIDA-1 directly associates with PSD-95, it has been suggested that its temporary shuffling out of the PSD core opens up a large number of PSD-95 binding sites and, consequently, allows for NMDAR reorganization during synaptic excitatory state (Dosemeci *et al.*, 2015).

1.5.3. AIDA-1 as Novel Synapse-to-Nucleolus Messenger and its Role in NMDAR Regulated LTP and Other Emerging Cellular Functions

In 2007, Jordan at al. reported N-methyl-D-Aspartate (NMDAR) dependent nuclear transport of the AIDA-1d isoform upon neuronal stimulation. NMDAR activation triggers the proteolytic cleavage and subsequent translocation of the AIDA-1d fragment to the nucleus where it associates with Cajal bodies and stabilizes its interaction with nucleoli (Jordan *et al.*, 2007). Another isoform, AIDA-1c, lacks exon 19 and was shown to interact with coilin, a major Cajal body protein marker (Xu & Hebert, 2005). Cajal bodies are nuclear suborganelles enriched with small nuclear ribonucleoproteins (snRNPs) and basal transcription factors specializing in pre-mRNA, pre-rRNA, siRNA

and miRNA processing (Pontes *et al.*, 2008) and telomerase formation (Jády *et al.*, 2006; Xu & Hebert, 2005). Silencing AIDA-1 through siRNA knockdown resulted in disruption of Cajal bodies and increased cell death rate (Xu & Hebert, 2005).

Among other "shuffling" proteins, identified by proteomics studies, AIDA-1 was suggested to serve in NMDAR-regulated neuronal signaling and protein trafficking (Dudek, 2007; Jordan & Kreutz, 2009). NMDA receptor stimulation leads to a chain of nuclear signaling events, characterized by increased nuclear transfer of several synaptic proteins, including AIDA-1, that were originally suggested by Jordan *et al.* (2009) as novel synapse-to-nucleus messengers. A growing body of evidence strongly suggests that NMDAR signaling controls complex processes regulating long term memory formation, thus, neuronal plasticity. The facilitated nuclear transport of AIDA-1 was supported by the presence of a nuclear localization sequence, 124-HRKR-127, within the SAM1-SAM2 domain interface (Kurabi *et al.*, 2009). Most recently, the breakthrough study by Bryen A. Jordan's group (2015) reported forebrain-specific conditional AIDA-1 knock-out (cKO) results in mouse models that redefine AIDA-1 function as a subunit-specific NMDAR transport facilitator. NMDAR channels are composed of 4 subunits, two GluN1 and two subunits primarily responsible for forebrain signal transmission: GluN2A and GluN2B (Tindi *et al.*, 2015). Selective loss of AIDA-1 in the forebrain manifested as a change in NMDAR subunit composition thereby decreasing the abundance of GluN2B-NMDARs and significantly increasing GluN2A-mediated synaptic transmission. Likewise, accumulation of GluN2B in the ER was confirmed as well as the direct association of AIDA-1 with the CASK signaling complex (Tindi *et al.*, 2015). Although

the molecular basis for preferential GluN2B binding has yet to be revealed, Tindi *et al.* (2015) suggested that other signaling molecules could be involved, especially APP and CASK that interact with AIDA-1 and NMDARs (Jeyifous *et al.*, 2009; Tindi *et al.*, 2015). Another study has shown that the CASK/MALS1/Mint1 tripartite signaling complex facilitates transport of NMDARs from endoplasmic reticulum (ER) to the nucleus and elimination of SAP97 and CASK expression causes accumulation of GluN2B in ER (Jeyifous *et al.*, 2009; Tindi *et al.*, 2015).

AIDA-1 PTB was also shown to recognize a non-NPTY sequence in the juxtamembrane region of ephrin A8 (EphA8) receptor tyrosine kinase (Shin *et al.*, 2007). Another study reported that ubiquitinated EphA8 associated with AIDA-1b via SAM domains (Kim *et al*., 2010). The heterotypic interaction between SAM domain of EphA2 receptor and SAM1 domain of scaffolding protein Odin has been characterized at the structural level (Mercurio *et al.*, 2012). Encoded by the *ANKS1A* gene, Odin is ubiquitously expressed in most tissues including brain (Park *et al.*, 2015) and has a strikingly similar domain organization as AIDA-1 and high sequence similarity (i.e. Odin PTB to AIDA-1 PTB - 81% identity). Odin has been implicated in the of EphA receptor signaling (Zhong, *et al.,* 2011; Mercurio *et al.*, 2012) and regulation of epidermal growth factor receptor (EGFR) (Park *et al.*, 2015). Experimental evidence suggests that Odin, in conjunction with 14-3-3 protein complexes, downregulates growth factor signaling via regulation of receptor tyrosine kinase endocytosis (Zhong, *et al.,* 2011). In contrast to Odin that expressed in nearly all mammalian cell lines, AIDA-1 could perform a brain-specific function in receptor tyrosine kinase signaling regulation. However, to date, the

biological significance of AIDA-1–EphA8 interaction remains unknown, and no functional relation to Odin has been established.

1.5.4. Connection of ANKS1b to the Human Diseases

A number of genetic studies have linked the AIDA-1 gene *ANKS1B* to schizophrenia (Purcell *et al.*, 2014; Snyder & Gao, 2013) and autism spectrum disorders (ASDs) (Pinto *et al.*, 2014; M. Uddin *et al.*, 2014). A systematic exome sequencing study in patients with schizophrenia revealed *de novo* mutations in *ANKS1B* (Fromer *et al.*, 2014; Purcell *et al.*, 2014). Single nucleotide polymorphisms (SNP) in *ANKS1B* and rare copy number variations had been observed in ASD affected patients (Pinto *et al.*, 2014). Furthermore, *ANKS1B* comprises exons critical for brain development: *de novo* mutations in these exons are enriched in autism (Uddin *et al.*, 2014). ASDs have been strongly associated with NMDAR dysfunctional synapses and GluN2B-containing NMDARs in particular (Tindi *et al.*, 2015). Moreover, GluN2B-containing NMDARs have been linked to AβPP processing regulation and Alzheimer's disease pathogenesis (Mota, Ferreira, & Rego, 2014). Tindi *et al.* (2015) hypothesize that mutations in the AIDA-1 gene may contribute to neuropsychiatric disorders through alternations in the GluN2B-NMDARs functions. The alternative splicing leading to domination of GluN2B-inactive AIDA-1 isoform is another possible cause yet to be investigated.

Continuous attenuation of synaptic activity involves extensive protein transcription and degradation regulation, and, therefore, requires complex nuclear signaling. Although the precise molecular mechanisms are not fully understood, AIDA-1

multiple protein partnerships including AD associated protein APP, Cajal body protein Coilin, EphA8 receptor, PSD95 and NMDA receptor, proteins of CASK scaffold, and, quite possibly, other unrevealed ligands, indicate the wide range of AIDA-1 orchestrated functions in the brain and, perhaps, therapeutic possibilities for human neurodegenerative disorders yet to be uncovered.

1.6. Scaffolding Proteins Caskins and Their Emerging Neuronal Functions

Multidomain protein Caskin is a relatively recent addition to the scaffolding group of proteins; it was originally discovered as a brain-specific protein found both in vertebrates and invertebrates (Tabuchi *et al.*, 2002). Two mammalian homologs, *CASKIN1* and *CASKIN2*, encoded on chromosome 16 and 17 respectively, are classic members of multidomain scaffolding proteins. Both have a similar domain organization: The amino-terminal part is composed of a series of ankyrin repeats followed by an SH3 and two SAM domains in close juxtaposition, the carboxy-terminal half consists of low complexity, proline-rich sequences (Balázs *et al.*, 2009) ending with a conserved 25 aa. segment of unknown function. All of these modules are recognized in the literature as protein-protein interaction domains (Krauss, 2008) and with the exception of the proline-rich segment, demonstrate substantial (~70%) sequence conservation (Tabuchi *et al.*, 2002) in Caskins. Furthermore, tandem SAM domains are well-known for their oligomerization potential based on their surface complementarity.

1.6.1. Caskin-Specific Scaffold as Part of Ca2+/Calmodulin-Associated Ser/Thr Kinase (CASK) Pathway

Originally, Caskin1 and 2 were named by homology, for the ability to interact with the Membrane Associated Guanylate Kinase (MAGUK) protein CASK (Ca2+/calmodulin-associated Ser/Thr kinase), a plasma membrane scaffolding protein and transcriptional co-regulator (Ojeh *et al.*, 2008; LaConte & Mukherjee, 2013). However, only Caskin1, but not Caskin2, is capable of binding CASK. The CASK/Mint1 and Velis tripartite complex is functionally linked to calcium-mediated signaling, actin microfilament assembly, and communication through the neurexin-neuroligin synaptic adhesion junctions (Borg *et al.*, 1999; LaConte & Mukherjee, 2013; Tabuchi *et al.*, 2002). The structure of CASK is composed of an N-terminal Ca2+/calmodulin-dependent protein kinase II (CaMKII) domain and PDZ domain, a central SH3 domain and finally a guanylate kinase homology domain at the C-terminus (Ojeh *et al.*, 2008; LaConte & Mukherjee, 2013; Tabuchi *et al.*, 2002). A consensus sequence, ExIWVxR, located between the SH3 domain and SAM1 of Caskin1 was identified as minimal binding motif recognized by CaMKII of CASK (Stafford *et al.*, 2011a; 2011b; Tabuchi *et al.*, 2002). This consensus peptide sequence, termed CID (CASK interaction domain), is conserved in both Mint1 and Caskin1, but not in Caskin2 (Stafford *et al.*, 2011a) (**Figure 1.7**).

Shortly after, the crystal structure of another scaffolding protein called Liprin-α2, recognized to be involved in synaptogenesis, cell adhesion and cell migration, was reported in complex with CASK (Wei *et al.*, 2011). In the Liprin-α2-CASK-CaMKII complex the C-lobe of CaMK is involved in an extensive interaction network with both

SAM1 and SAM2. An insertion helix is located between the SAMs (namely αL helix) (Wei *et al.*, 2011) where Trp981 serves as a hydrophobic anchor within the determined minimal binding sequence, GNVWVTHE. Therefore, it is clear that although Liprin-α2 employs a different minimal binding motif, all three proteins (Mint1, Caskin1 and Liprin-α2) target the same hydrophobic pocket on CASK and all three use tryptophan as a key docking residue (Stafford *et al.*, 2011a; Wei *et al.*, 2011). The fact that both mammalian CASK/Caskin1 and CASK/Liprin-α2 are not observed in the other CASK orthologues (LaConte & Mukherjee, 2013), and the mammalian Liprin-α2 binding motif is not conserved in *C-elegans*, *Drosophila* and Liprin-α1 isoform, could be a manifestation of evolutionary development in CASK regulated pathways of higher order organisms. In addition to the CaMK domain of CASK, Caskin1 interacts with the intracellular tails of cell surface proteins such as neuronal cell adhesion protein neurexin 1 and Ca2+-regulated vesicle fusion-mediator protein synaptotagmin (Stafford *et al.*, 2011b). A functional connection of the Caskin competitor Mint1 to synaptic vesicle fusion has been also reported (Okamoto & Sudhof, 1997; Olsen *et al.*, 2005). Therefore, the putative model of Caskin1-modulated architecture at presynaptic sites has been proposed, whereby neurexin1-CASK-Caskin1 associations link the CASK/Caskin1/Velis complex to the presynaptic membrane and oligomeric-Caskin1 recruited synaptotagmin docks and guides synaptic vesicles to the synaptic cleft.

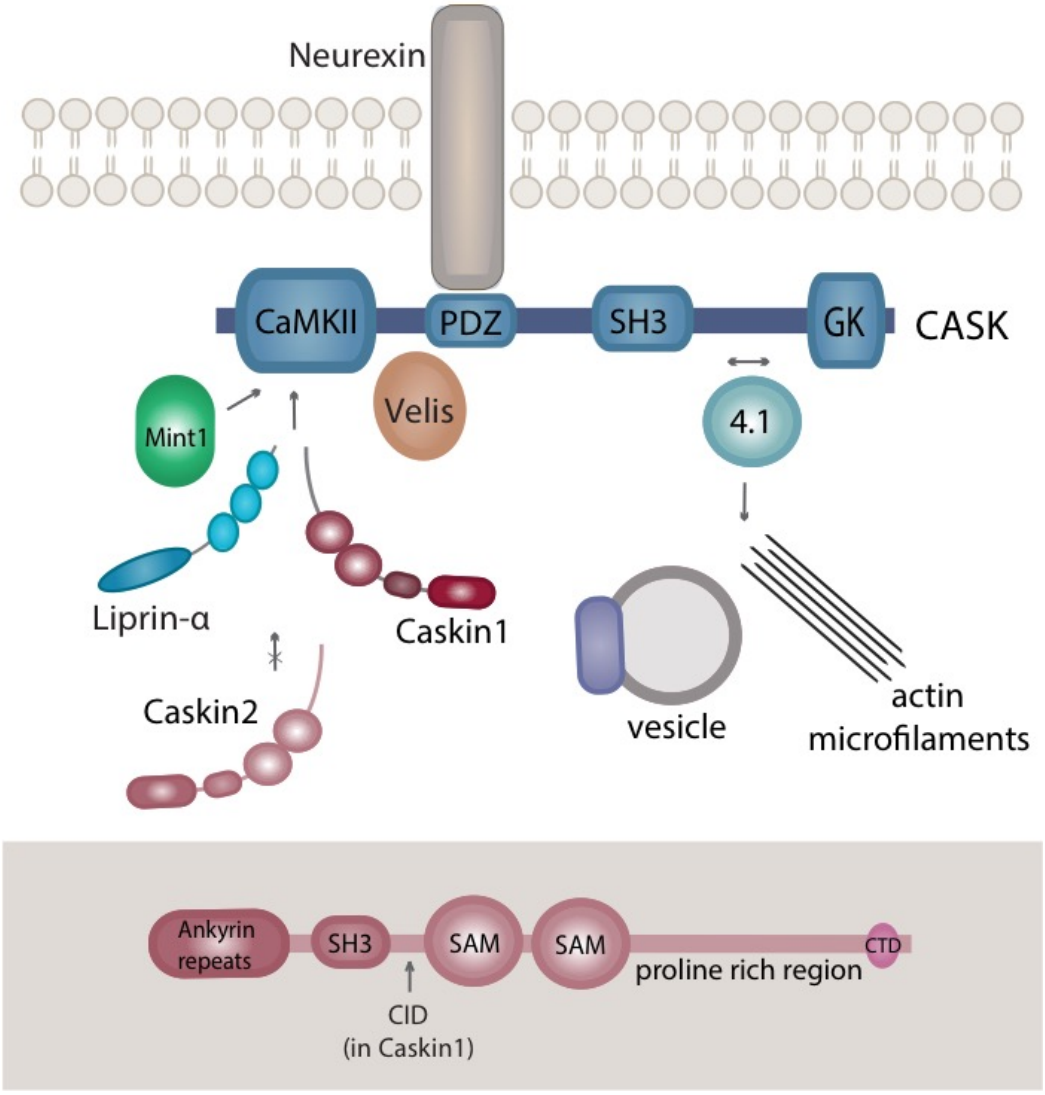


Figure 1.7 – An array of CASK-mediated protein-protein interactions. Cask is recruited to the plasma membrane via its PDZ domain interactions with cytoplasmic regions of cell surface proteins (neurexins, syndecans). The scaffolding proteins Caskin1 Mint1 compete for the same binding sequence within the N-terminal CaMKII domain (Wei *et al.*, 2011), whereas Mint1 and Liprin-α could form a complex with CASK bound to neurexin1 (LaConte *et al.*, 2016). Velis recognizes the sequence between CaMKII and PDZ domain. In addition, protein 4.1 interaction with the C-terminal SH3 and guanylate kinase domains is required for actin filaments assembly. This interaction may also link actin microfilaments to the cytoplasmic tails of cell-surface receptors. CASK C-terminal sequences may support additional intra- and intermolecular interactions. Bottom panel depicts Caskin1/2 domain organization. *Domain annotations:* CaMKII, Ca2+/calmodulin-dependent protein kinase II; *PDZ, domain present in PSD-95; SH3, Src-homology-3, GK, guanylate kinase; SAM, sterile alpha motif; CID, CASK interaction domain; CTD, C-terminal domain.*

A global database search revealed that T-cell lymphoma invasion and metastasis 1 (TIAM1) was the only other protein containing this conserved sequence within EEVIWVRRE peptide that, not unexpectedly, also was able to bind CASK *in vitro* (Stafford *et al.*, 2011a). The TIAM1 protein is responsible for tumor propagation and metastasis (Minard et al., 2004) and a certain overlap in CASK/TIAM1 cellular functions have been recognized (Caruana, 2002; Mertens, Pegtel, & Collard, 2006; Stafford et al., 2011b). A follow-up investigation of Caskin1/Mint1/TIAM1 interaction *in vivo* and its downstream effects is a logical next step.

More recent studies have reported Caskin to have a distinct role in retinal synapses (Anjum *et al.*, 2014) in addition to specializations in brain neuronal synapses (Tabuchi *et al.*, 2002). Punctate co-localization of Caskin1 and CASK was found in a distinct subset of retinal synapses at pre-synaptic sites, which once again supports Caskin in a highly specialized role in organizing scaffolds around particular cell surface receptors (Anjum *et al.*, 2014). This functional outcome is further reinforced by reported Caskin interactions with cytoskeletal adaptor proteins Abi2 (Balázs *et al.*, 2009) and Nck/Dock (Weng *et al.*, 2011) which link actin-based cytoskeleton to cell surface receptors (Anjum *et al.*, 2014; Weng *et al.*, 2011). The SH3 domain of Caskin1 has been shown to be recruited by EphB1 and Nck (Pesti *et al.*, 2012) and subsequently phosphorylated, resulting in a structural change though the biological significance of this is still under investigation. Members of our research group recently contributed with an NMR structure of Caskin2 SH3 domain (Donaldson & Kwan, 2016) and proposed that upon phosphorylation of Y336, the SH3 domain could serve as a suitable ligand for the Crk/Grb2 family of SH2 domains, as

supported by molecular modeling. However, the possibility of the SH3 domain as a non-functional remnant is also anticipated. New studies detected decreased levels of Caskin1 mRNA as a result of intestinal ethanol stress in rats (Middleton *et al.*, 2009) and the study reveals a possible connection of Caskin1 to infantile myoclonic epilepsy through analysis of disease-related co-expression profiles conserved in human and mouse (Ala *et al.*, 2008). These data are opening intriguing possibilities of Caskins' connection to a number of cellular pathways.

1.6.2. Caskin2 Functionally Connected to Leukocyte Common Antigen-Related (LAR) Tyrosine Phosphatase Regulated Pathways

Phosphorylation and dephosphorylation on protein tyrosine residues is the fundamental cell-signaling mechanism maintained by protein-tyrosine kinases (PTKs) and protein-tyrosine phosphatases (PTPs). The Leukocyte common Antigen Related (LAR) family of proteins is a subclass of receptor-like PTPases (RPTPs) in vertebrates represented by three homologs: LAR, PTPσ and PTPδ (Chagnon *et al.*, 2004).

LAR has been implicated in regulation of neurite outgrowth, axonal extension and guidance (in *Drosophila)*, disassembly of cell focal adhesions, maintenance of neuromuscular junctions, nerve regeneration, murine mammary gland development and function and has even been connected to cancer metabolic regulation (Chagnon *et al.*, 2004; Serra-Pagès *et al.*, 1998). The fact that both Liprin-α2 and Caskin2 were shown to interact with LAR may link them to the CASK signaling pathway, on the other hand, some experimental evidence suggests a possibility of a distinct set of functions. The initial

results by Weng *et al.* (2011) regarding the interaction between *Drosophila* orthologs Dlar and Csk2 provided the background for our investigation of human LAR and Caskin2 partnerships. Therefore, it is essential to provide an overview on LAR PTPases from both structural and functional perspectives.

1.6.3. Structure and Regulation via Alternative Splicing and Proteolysis

The extracellular domain of LAR PTPase greatly resembles the domain structure of cell adhesion molecules (CAM) and contains a variable number of Immunoglobulin-like (Ig-like) and fibronectin type III-like (FNIII) repeats followed by a hydrophobic transmembrane region and ending with two cytoplasmic domains, D1 and D2 (Chagnon *et al.*, 2004). Unlike the variable extracellular region, the intracellular domains are remarkably conserved among receptor-like PTPases demonstrating ~84% identity in both vertebrates and invertebrates (Chagnon *et al.*, 2004; Pulido *et al.*, 1995). The membrane proximal domain (D1) is responsible for PTPase enzymatic function and contains a conserved (I/V)HCXAGXGR(S/T)G motif that forms an active site cleft (Tonks, 2003). The central cysteine is directly involved in the nucleophilic attack on the phosphoryl group of the substrate, forming a covalent phosphoenzyme intermediate followed by dephosphorylation (Pot & Dixon, 1992; Serra-Pagès *et al.*, 1998; Zhang, Wang, & Dixon, 1994). Although both domains, D1 and D2, show significant conservation of primary sequence and very similar tertiary structures (Nam *et al.*, 1999) two key amino acid differences (Y/L and D/E) in the substrate coordinating loops change the catalytic site conformation, which renders the D2 domain catalytically inactive and thereby suggests a

regulatory nature (Tsujikawa *et al.*, 2008). The tandem domain crystal structure was solved as a monomer (Nam *et al.*, 1999) at the same time studies demonstrated the inhibitory effect of *cis*-dimerization (or oligomerization) on LAR-RPTP catalytic activity suggesting the structure-directed regulatory role of D2 domain in this catalytic suppression mechanism (Coles, Jones, & Aricescu, 2015; Coles *et al.*, 2014; Wallace *et al.*, 1998). In addition to dimerization, proteolytic cleavage, alternative splicing and glycosylation have been implied to regulate the catalytic activity of LAR RPTPs at distinct synaptic sites and in different tissues (Chagnon *et al.*, 2004; Pulido *et al.*, 1995; Wallace *et al.*, 1998; Zhang & Longo, 1995)

1.6.4. Synaptic Functions of LAR PTPase family

LAR family receptors have been strongly associated with axonal pathfinding, regeneration, synaptogenesis, based largely on genetic and biochemical studies in *Drosophila* and *C. elegans* (Desai *et al.*, 1996; Krueger *et al.*, 1996; Weng *et al.*, 2011b; Xie *et al.*, 2012; Yang *et al.*, 2003; Yeo *et al.*, 1997). It has been proposed that LAR receptor-like protein tyrosine phosphatases link the extracellular matrix (ECM) to developing focal adhesions and promote actin cytoskeleton reorganization (Baker & Macagno, 2010). Remarkably, the axon guidance function of LAR RPTP appears to be preserved across different types of neurons, including motor (Desai *et al.*, 1996; Krueger *et al.*, 1996; Weng *et al.*, 2011), retinal neurons (Clandinin *et al.*, 2001; Hofmeyer & Treisman, 2009; Maurel-Zaffran *et al.*, 2001) and more recently reported somatosensory neurons in a zebrafish study (Wang *et al.*, 2012). The latest intriguing study revealed the

details of LAR regulatory function in circadian pacemaker neuron development. The Lar RNAi knockdown phenotype in *Drosophila* resulted in the elimination of axonal processes from clock neurons and ultimately disruption of activity rhythms (Agrawal & Hardin, 2016).

While precise mechanisms involving the LAR RPTP family proteins remain to be fully characterized, the studies in mammalian systems are especially complicated due to a certain level of cooperation and/or redundancy of functions between LAR, PTPσ and PTPδ (Um & Ko, 2013). Nonetheless, two major synaptic processes mediated by mammalian LAR RPTPs have been highlighted: synaptic assembly and synaptic plasticity. In contrast to studies in invertebrates where LAR was predominantly implicated in presynaptic processes (Xu & Fisher, 2012), mammalian LAR was found to be enriched at excitatory synapses (Dunah *et al.*, 2005) serving at postsynaptic sites. RNA interference (RNAi) knockdowns resulted in a dramatic decrease of excitatory synapses and dendritic spines (Dunah *et al.*, 2005); likewise, disruption of the intracellular protein-binding function of LAR manifested as inhibition of the β-catenin–cadherin complex recruitment at synaptic sites (Brigidi & Bamji, 2011; Um & Ko, 2013). The β-catenin–cadherin complex is involved in axonal development and dendrite arborization (Um & Ko, 2013) and by association with G-protein-coupled receptor kinase-interacting protein 1 (GIT1) linked to AMPA glutamate receptor-mediated synaptic transmission (Brigidi & Bamji, 2011; Ko *et al.*, 2003). Therefore, the role of LAR PTPase AMPA signaling and regulation of synaptic transmission was implied. The growing number of extracellular ligands interacting with Ig domains of LAR, including syndecan, dallylike protein, and laminin-

nidogen complex (Fox & Zinn, 2005; Johnson *et al.*, 2006) endorse LAR as a membrane signal transmitter and synaptic active zone morphogenesis regulator (Stryker & Johnson, 2007).

All three mammalian LAR RPTPs have been reported to interact through their D2 domains with SAM domains of the alpha-liprin family (Serra-Pagès *et al.*, 1998; 1995; Stryker & Johnson, 2007) one of the established determinants of synaptic organization and maturation in *C. elegans* and *Drosophila* (Dai *et al.*, 2006; Spangler & Hoogenraad, 2007). The Liprin-α-LAR complexes specifically localize at focal adhesions (Serra-Pagès *et al.*, 1995). The recruitment of Liprin-α into the Velis/CASK/X11 scaffold (Olsen *et al.*, 2005; Wei *et al.*, 2011), as discussed earlier, and other Liprin-α partnerships with known active zone modulators such as RIM1-α, ERC2 and GIT1 (Ko *et al.*, 2003; Schoch *et al.*, 2002; Stryker & Johnson, 2007) suggest the presence of a Liprin-α-LAR-specific assembly mechanism localizing signaling components at synaptic sites.

Compared to the liprin family, Caskin2 (ckn2) is a more recently identified protein partner of LAR and so far their interaction and its functional requirement for motor axon guidance have been confirmed in Drosophila (Weng *et al.*, 2011). The fact that Liprin-α and Caskin2 cannot bind LAR simultaneously suggests the possibility of the distinct signaling outcomes of these complexes (Weng *et al.*, 2011). Moreover, since both proteins are linked to the CASK signaling pathway and are genetically preserved in both invertebrates and vertebrates, it is reasonable to hypothesize that the mammalian Caskin-LAR specific assembly adds signaling complexity and specificity but its functional

significance remains to be revealed.

The diversity of LAR-RPTPs synaptic protein partnerships including major synaptic scaffolding molecules and transmembrane signaling proteins promote LAR-RPTPs as *bona fide* synaptic adhesion molecules (Han, 2016a; Um & Ko, 2013). In other words, their primary function in neurons as synaptic organizers function has gained substantial support in the literature.

To add to the complexity, besides the functions outlined above, studies had shown that LAR PTPase deficient mice exhibit reduced glucose and insulin levels and impaired mammary gland development (Schaapveld *et al.*, 1997; Yeo *et al.*, 1997). Several studies have revealed the LARσ specific requirement for efficient neuronal regeneration (Fry *et al.*, 2010; Van der Zee *et al.*, 2003; Xie *et al.*, 2001).

1.6.5. Oligomerization Through Tandem SAM Domains

The intrinsic ability of multiple SAM domain proteins to oligomerize is an important requirement often strongly connected to biological function. The tightly bound Caskin1 SAM1-SAM2 unit oligomerizes into long fibrils via contacts between the exposed head and tail surfaces. Electron micrograph images of both Caskin1 and Caskin2 oligomers (**Figure 1.8.**) have a regular shape of rod-like structures although they appear morphologically different.

As this mode of oligomerization has also been observed *in* vivo, Caskin1 may be serving as a railway that links presynaptic vesicles together through an association with synaptotagmin (Stafford *et al.*, 2011b). Oligomerization potential is clearly strategically advantageous for molecular scaffolds in the assembly of signaling complexes at specific cellular sites and signal output amplification. The localized CASK CaMK domain assembly on Caskin1 SAM SAM fibrils resemble "beads on the string," and furthermore, *in vitro* experiments support that the polymer interface indeed facilitates this assembly (Stafford *et al.*, 2011a). The novel oligomerization mode of Caskin2, distinct from Caskin1, is presented in detail in **Chapter 3** of my thesis.

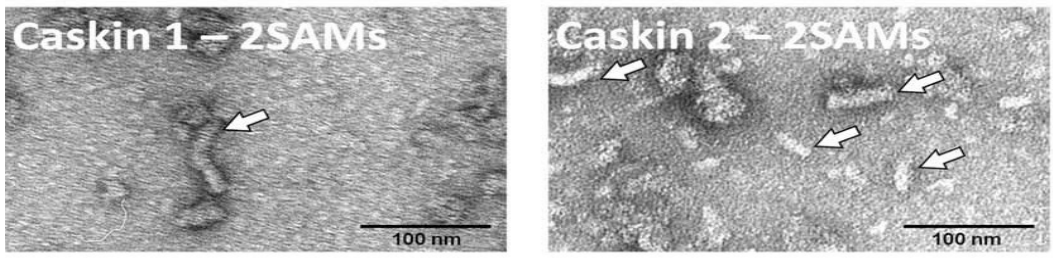


Figure 1.8 – Electron micrographs of Caskin 1 and 2 tandem SAM oligomers expressed as negGFP-hSAM fusions and visualized by negative stain Electron Microscopy (EM) method demonstrate rod-like polymeric structures. Adapted from (Knight *et al.*, 2011).

Shank is the closest protein family with an analogous domain organization. It has the same structural elements as Caskins: ankyrin repeats, SH3 domain, polyproline region, with tandem SAM in Caskins and a single C-terminal SAM domain in Shank. The fact that sequence similarity between Caskin and Shank is not significant indicates that there is no evolutionary conservation between them, but rather the possibility of similar

scaffolding functions. At the postsynaptic sites Shanks are involved in regulation of actin-dependent cytoskeletal remodeling, endocytosis of AMPA receptors, synaptogenesis, synaptic plasticity and signal transmission (Jiang & Ehlers, 2013). Mutations in Shank family genes are connected to neurodevelopmental disorders such as schizophrenia (Gauthier *et al.*, 2010) and autism (Jiang & Ehlers, 2013). The single SAM domain of Shank forms long alpha-helical fibrils cross-linked with zinc ions into a tightly packed highly ordered array (Baron *et al.*, 2006). The most recent study by Arons at al., (2016) proposed the Zn2+ sensitive signaling mechanism at excitatory synapses regulated by Shank3 oligomerization state. They demonstrated that high Zn2+ levels facilitate an assembly of Shank3-based scaffolds that bring Homer, neuroligin, and AMPA receptors into the activated signaling complex, whereas low Zn2+ levels lead to an opposite effect resulting in retention of oligomerized Shank3 aggregates within dendritic spines, and therefore, a weaker synaptic response (Arons *et al.*, 2016). An overall domain organization similarity and polymerization potential suggest that Shank and Caskins serve as signaling protein assembly platforms in the post- and pre-synapse respectively. If oligomerization of Shank3 into fibrils yields dense/static network ultimately reducing the synaptic strength, Caskin2 may alternatively function as a more dynamic net that can be assembled/disassembled rapidly to facilitate release of neurotransmitters at the presynapse.

While the biological significance of a number of other Caskin protein partnerships identified using a yeast two-hybrid screen (Balázs *et al.*, 2009) have yet to be uncovered, one of them is worth mentioning in context of this section: the ubiquitin ligase Siah1. Its

presence among other Caskin ligands could be significant in line with the outcome from a recent global mass spectrometry survey that identified a ubiquitinated lysine residue (K522) in Caskin1 (Wagner *et al.*, 2012) (K535 in Caskin2) located exactly at the oligomerization interface of the SAM-SAM tandem (described in details in **Chapter 3** of this thesis). The presence of the ubiquitination site may suggest the possibility of another regulatory mechanism, orchestrated by ubiquitin, specifically targeting Caskin oligomerization.

1.7. Thesis Overview

An increasing complexity of signaling pathways and cellular networks through vertebrate evolution was a significant driver in protein functional diversity and multiplicity. Molecular modules such as PTB and SAM domains are essential tools used by scaffolding proteins for multiple protein partner recruitment, and hence, regulation of cellular signaling. In addition to serving as protein-protein interaction platforms, SAM domains have a propensity for oligomerization which in itself is a powerful intramolecular signaling regulatory mechanism. The AIDA-1 and Caskin2 structure-function studies described in this dissertation contribute to both a general understanding of protein interaction/oligomerization module properties as well as new insights into their particular functional contributions at level of the neuron.

Over a decade ago AIDA-1 was linked to amyloid-β precursor protein processing essentially earning its name for the ability to bind its cytosol-projected domain. Recent genetic studies connect the AIDA-1 gene ANKS1b with a number of neuropsychiatric disorders including Alzheimer's disease (Mota *et al.*, 2014), schizophrenia (Fromer *et al.*, 2014; Purcell *et al.,* 2014) and autism spectrum disorders (ASDs) (Pinto *et al.*, 2014; Uddin *et al.*, 2015; Uddin *et al.*, 2014). With respect to its role in the amylogenic pathway leading to Alzheimer's disease, the consequences of AIDA-1 interaction with APP-AICD in ABPP processing remain largely unknown. At a minimum, this interaction may serve as a cytoplasmic anchor defining the subcellular localization of AIDA-1 as a result of specific signaling events. This work characterizes the affinity and specificity of the AIDA-1 PTB domain and APP-AICD interaction to provide a comparative analysis at the structural level to the other PTB domain-containing proteins, Fe65 and X11/Mint1, particularly known as AβPP processing and Aβ secretion mediators.

The following section of the dissertation begins with a structural investigation of the Caskin2 protein SAM tandem module. The variety of Caskin2 oligomeric states, monomeric/dimeric at low and oligomeric at high concentrations, were identified and characterized by a combination of structural and biochemical approaches. Such structural diversity could serve as a concentration-dependent mechanism to suppress or amplify low-affinity protein interactions common in signaling pathways. Growing evidence of the mammalian LAR receptor tyrosine phosphatase regulatory functions in complex processes such as axonogenesis, synaptic assembly, and neuronal plasticity, dictates a need for elucidation of the components of LAR-specific signaling pathways. The

confirmation and structural dissection of Caskin2/LAR Homo sapiens homologs partnership were the additional objectives of my studies.

The remarkable abundance of AIDA-1 in various brain compartments; highly specific neuronal localization; multiple protein partnerships of AIDA-1; the self-oligomerization property of Caskin2 collectively indicate that these scaffolding proteins have a potential to orchestrate a range of neuronal processes yet to be fully understood or uncovered. These integrated structural and functional studies hopefully will serve as a platform for further elucidation of their cellular functions.

CHAPTER 2:

SOLUTION STRUCTURE AND PEPTIDE BINDING OF THE PTB DOMAIN FROM THE AIDA-1 POSTSYNAPTIC SIGNALING SCAFFOLDING PROTEIN

The content reported within this chapter has been published in the article listed below:

- Smirnova E., Shanbhag R., Kurabi A., Mobli M., Kwan J.J., Donaldson W.L. (2013) Solution Structure and Peptide Binding of the PTB Domain from the AIDA-1 Postsynaptic Signaling Scaffolding Protein. *PLoS ONE* 8(6): e65605.

Information of the authors' contributions is provided in the corresponding section.

2.1. Introduction

Neurons receive chemical signals through a collection of over four hundred proteins that are organized into a network termed the postsynaptic density (PSD) (Jordan *et al.,* 2004). AIDA-1, a prominent member of the PSD, is edited into at least five isoforms, all of which contain two sterile alpha motif (SAM) domains and a PTB domain (Ghersi *et al.,* 2004). Together, these domains suggest a role for AIDA-1 as a scaffolding molecule that collates proteins at the synapse through multiple protein-protein interactions. Mutations in AIDA-1, consequently impair long term potentiation (LTP), a basic molecular requirement for learning and memory (Jordan *et al.,* 2007). Owing to its

role in many signaling processes, AIDA-1 (located at chromosome 12q23.1) is also known as ANKS1B, ANKS2, cajalin-2 and EB-1.

AIDA-1 derives its name from the ability to bind the carboxy terminal cytoplasmic region of amyloid precursor protein (APP), widely implicated in the development of Alzheimer's disease. AIDA-1 isoforms demonstrate differences in subcellular localization, affinity for APP and effect on the processing of APP to the Aβ40 nonpathologic fragment (Ghersi *et al.,* 2004). While AIDA-1 is predominantly expressed in brain, a related protein, Odin (ANKS1A), with the same domain organization, is more ubiquitously expressed and serves as an adaptor modulating the signaling outcomes of epidermal derived growth factor receptor (EGFR), platelet derived growth factor receptor (PDGFR) and Ephrin A8 receptor tyrosine kinase (Emaduddin *et al.,* 2008).

Previously, the members of our group determined the NMR structure of the AIDA-1 SAM domain tandem and demonstrated that a nuclear localization signal was sequestered at the interface of the two domains (Kurabi *et al.,* 2009). In this study, we have continued a reductionist investigation of a potential AIDA-1 SAM-SAM-PTB domain supramodule by determining the NMR structure of the PTB domain. The structure of the AIDA-1 PTB domain and its ability to bind an NPxY motif in the amyloid precursor protein (APP) cytoplasmic region are similar to the postsynaptic signaling proteins APPL (Mao *et al.,* 2006) and X11/Mint (Matos *et al.,* 2012) and to a lesser extent, Fe65 (Mulvihill *et al.,* 2011). Thus, the nature of signals arising from APP is likely dependent on the context specified by AIDA-1 and the relative affinity of its competitors.

Our initial attempt to perform NMR structural studies on the AIDA-1 PTB domain were hindered by poor solubility regardless of solution conditions chosen. Expression of the PTB domain from Odin (81% identity), presented an even worse case, as this protein fragment could not be refolded from inclusion bodies. A strategy that we pursued to improve the solubility of the AIDA-1 PTB domain involved the progressive substitution of aromatic amino acids that were predicted to be solvent exposed.

2.2. Methods

2.2.1. Cloning, Expression and Protein Purification

A gene fragment encoding the PTB domain (aa. 1043–1195) of human AIDA-1b was PCR amplified with NdeI and EcoRI restriction sites and was subsequently inserted into pET28a (Novagen). The expressed protein contained an amino terminal 6xHis tag and intervening thrombin site. Other PTB domain fragments that lacked either the N-terminal 6xHis tag or the entire affinity tag along with 16 additional unstructured residues were also as insoluble as the fragment chosen for this study. To align the PTB domain described in this study with numerous AIDA-1 isoforms, S1 in the PTB domain structure corresponds to S1045 in AIDA-1b, the longest isoform. A one-liter fermentation in a minimal medium containing 1 g of $^{15}NH_4Cl$ and 4 g of ^{13}C-glucose was sufficient to produce 5–10 mg of purified protein. Purification was achieved by Nickel-NTA affinity chromatography (Qiagen) and gel filtration chromatography on a S-100 HR 16/60 size exclusion column (GE Biosciences). Final buffer conditions were 20 mM Na-phosphate, pH 7.8, 0.15 M NaCl, 0.05% (w/v) NaN_3. Five single aromatic-alanine substitutions

(Y6A, F16A, F24A, Y70A and Y131A) were produced from pET28-AIDA-1-PTB using a service provided by Genscript (Piscataway, NJ). A 6xHis tagged PTB domain variant containing all five substitutions (PTB5M) was produced by DNA2.0 (Menlo Park, CA) by direct gene synthesis in the expression vector, pJExpress401 (T5 promoter plus kanamycin resistance). A 6xHis-tagged, APP-peptide (GYENPTYKFFE) fused to the amino terminus of the AIDA-1 PTB5M mutant with an intervening thrombin site was also synthesized by DNA2.0 in pJExpress401.

2.2.2. Protein Solubility Assessment

Since the objective of the aromatic-alanine substitutions was to improve solubility for a structure determination, ^{15}N-HSQC spectra were used qualitatively. From experience, the wild type AIDA-1 PTB domain was soluble for a least one day at room temperate at a concentration of 0.15 mM thereby permitting experiments to be performed but not to the extent of a structure determination. Each aromatic-alanine substitution mutant was concentrated to 0.15 mM, assessed by NMR and then concentrated until increased resonance line broadening was observed or there was apparent turbidity.

2.2.3. CD Spectroscopy

Far UV circular dichroism (CD) spectra were acquired with a Jasco J-810 instrument at a protein concentration of 50 μM using a rectangular cell with a 0.1 cm path length. Spectra were recorded from 260–200 nm with a scan rate of 50 nm/min and a 1.0 nm bandwidth. A midpoint denaturation temperature (T_m) was determined by heating samples from 20–

90 ºC at 2 ºC/min and monitoring ellipticity at 222 nm.

2.2.4. Protein Binding Studies

Fluorescein isothiocyanate (FITC) labeled peptides spanning portions of APP were produced and purified by CanPeptide (Montreal, QC) for fluorescence anisotropy based binding studies at 25°C using an Agilent Eclipse spectrophotometer equipped with a manual polarizer accessory. Buffer conditions were similar to those used for NMR spectroscopy. Measurements were made under identical conditions and averaged. Anisotropy was calculated from the relationship $(I_{parallel}–GI_{perp})/(I_{parallel}/2GI_{perp})$ and normalized with the blank experiment. The equilibrium dissociation constant (K_d) was calculated by direct fitting the titration curves with a standard two-state relationship using proFit 6.2 (Quantsoft).

2.2.5. Peptide Array

A set of 12-mer peptides on a 150×100 mm cellulose membrane in a 10×30 array was synthesized using the SPOTS method (Frank, 2002) with an Intavis MultiPep instrument. A crude estimate of the peptide content in each spot was made by staining the array with Fast Green FCF. The array was probed with 1 μM of the solubility enhanced 6xHis-PTB5M mutant in PBST (3.2 mM sodium phosphate, 0.5 mM potassium phosphate, 1.3 mM KCl, 135 mM NaCl, 0.1% Tween-20, pH 7.4). Following blocking and washing with 5% skimmed milk and 2.5% bovine serum albumin (Bioshop Canada) in PBST, bound AIDA-1 PTB was identified by incubating the array in a 1∶5000 dilution of horseradish

peroxidase (HRP)-conjugated 6xHis monoclonal antibodies in PBST and developing with a chemiluminescent reagent (Santa Cruz Biotechnology). A complete table of peptides is provided in **Table 2.5**.

2.2.6. NMR Spectroscopy

^{15}N-edited HSQC spectra of the wild type PTB domain, mutants and protein-peptide complexes were acquired at 30 ºC on a Varian 600 MHz NMR spectrometer equipped with a salt tolerant cold probe. The use of a low protein concentration (0.10–0.15 mM) permitted assessment of all protein fragments regardless of intrinsic solubility. Chemical shift assignments on a uniformly^{15}N, ^{13}C labeled sample of PTB5M at 0.8 mM were obtained using a conventional heteronuclear, triple-resonance strategy that incorporated non-uniform sampling for improved resolution and sensitivity. Backbone directed experiments: HNCACB, CBCA(CO)NH, HNCO, HNCACO, side chain directed experiments: H(C)(CO)NH, C(CO)NH, and ^{13}C/^{15}N-edited NOESY spectra were acquired on a Bruker Avance 900 MHz spectrometer equipped with a cold probe. Side chain HCCH-TOCSY, and aromatic HB(CBCG)CD, HB(CBCGCD)CE were acquired at 600 MHz. Protein solutions contained 10% D_2O with the exception of the ^{13}C-edited NOESY dataset in which the PTB5M sample was buffer exchanged into >95% D_2O before data acquisition. Datasets were processed with NMRPipe (Delaglio *et al.,* 1995) or the Rowland Toolkit (Hoch and Stern, 1996) as required and interpreted with CCpNMR Analysis 2 (Fogh *et al.,* 2006). Chemical shift assignments of PTB5M were deposited in the BMRB with the accession code 17934.

Structure Determination: From an initial set of 500 structures calculated with CYANA 3, the top 20 structures were selected with no NOE violations >0.3 Å and no torsion angle violations <5°. This ensemble was then subjected to additional refinement in explicit solvent with a Python script (wrefine.py) supplied with XPLOR-NIH 2.30. The top 15 structures according to lowest refinement energy was deposited as an ensemble in the Protein Data Bank with the accession code 2M38. The ensemble was aligned using MOLMOL 2K1 (Koradi *et al.,* 1996). Structure Comparisons: Cα RMSDs and alignments between the AIDA-1 PTB domain and related proteins were performed with PDBeFold (Krissinel and Henrick, 2004).

Peptide Docking Simulations: Starting from the AIDA-1 PTB5M structure and APP peptide ligand placed in analogous position to that observed in the X11 PTB domain crystal structure (Matos *et al.,* 2012), a two-stage docking simulation, at low resolution (200 structures) and then all-atoms high resolution (100 structures) was performed with FlexPepDock, part of the Rosetta 3.4 software package (Raveh *et al.,* 2011). A low energy structure was selected for analysis.

2.3. Results

Prior to the structure determination, a molecular model of the AIDA-1 PTB domain was made with HOMA (Bhattacharya *et al.,* 2007) using the crystal structure of the X11 PTB domain as the template (Zhang, 1997). Final refinement was performed with FOLDX (Guerois *et al.,* 2002). The surface of the PTB model was scanned for exposed aromatics and compared to a sequence alignment consisting of the PTB domains from

X11, Numb (S.-C. Li *et al.,* 1998) and Fe65 (Radzimanowski *et al.,* 2008). Of the sixteen aromatics in the AIDA-1 PTB domain, Y6, F16, F24, Y70, and Y131 were selected as candidates that were most likely to be surface-exposed (Figure 1a). Thus, by selecting aromatic amino acids (Phe/Tyr/Trp considered equally), we were effectively sampling mutations under sparse conditions that still cover a wide range of surfaces.

Table 2. 1 – Solubilities and thermal denaturation midpoints of the AIDA PTB domain and alanine substitution mutants.

PTB domain	Tm (°C)	Solubility (mM)	Side chain exposure
wild type	62	0.10	N/A
Y6A	65	0.10	exposed
F16A	64	0.20	exposed
F24A	64	0.45	partially exposed
Y70A	64	0.45	exposed
Y131A	64	0.15	exposed
5M	64	0.80	N/A
APP-PTB	72	0.20	N/A
APP-PTB5M	73	0.80	N/A

As shown in **Table 2.1**, the calculated T_m of the mutants was comparable to the wild type PTB domain suggesting that the alanine substitutions did not destabilize the fold. Like the wild type PTB domain, the Y6A and Y131A single mutants could only be concentrated to 0.1 mM before precipitation was observed. The remaining single mutants – F16A, F24A and Y70A – could be concentrated up to 0.5 mM; however, HSQC spectra at these concentrations suffered from line broadening and missing resonances. In contrast, the PTB5M mutant was very soluble at 0.8 mM, with line widths that were comparable to the single mutants acquired at low concentration. Thus, we observed a synergistic effect when multiple aromatic amino acids were substituted with alanine.

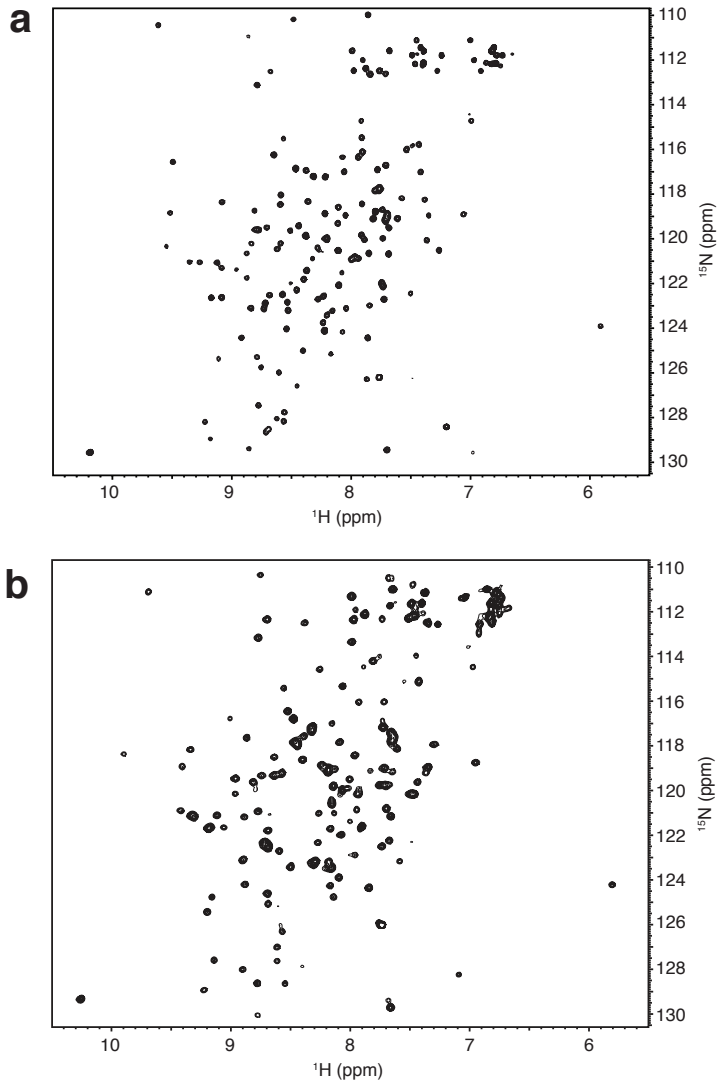


Figure 2. 1 – A comparison of 15N-edited HSQC spectra from the **(a)** AIDA-1 PTB5M protein and the **(b)** AIDA-1 PTB5M protein with an APP binding sequence (GYENPTYKFFE) appended to the N-terminus along with a linker sequence (TLRPPNEATALQ) derived from the native AIDA-1 protein. Both protein concentrations are 0.8 mM.

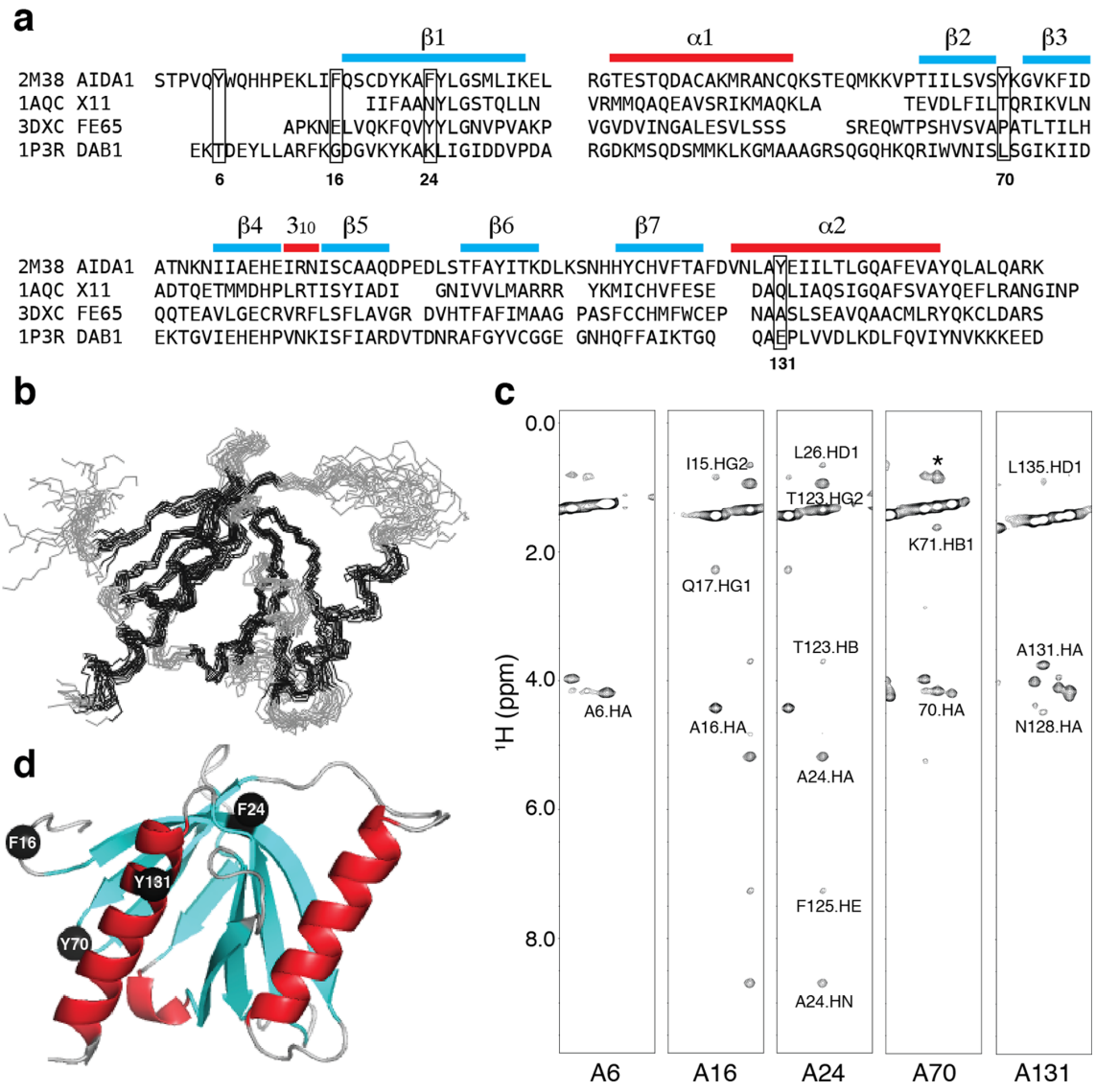


Figure 2. 2 – (a) Sequence alignment of the AIDA-1 PTB domain against the APP binding proteins, Dab1(Yun *et al.,* 2003), X11 (Zhang, 1997) and Fe65 (Radzimanowski *et al.,* 2008). Five aromatic amino acids selected for alanine substitution in AIDA-1 PTB domain are boxed. (b) Backbone atom superposition of top15 structures according to lowest refinement energy. **(c)** Strip plots of a ^{13}C-edited NOESY spectrum at the Cβ chemical shift of each alanine substituted in the PTB5M mutant. An asterisk denotes a resonance not associated with that strip. **(d)** A ribbon representation of the PTB5M model highlighting the positions of the alanine substitutions. Y6A is not shown in the figure as the first 14 amino acids are unstructured and were excluded from the structure calculation.

The impact of the APP ligand on solubility was also investigated. The APP ligand was added exogenously, as a 17-mer peptide and endogenously, by appending the sequence to the amino terminus of the wild type PTB domain. Tethering a peptide ligand to a protein is a useful approach to shift binding kinetics from biomolecular to unimolecular and ensure stoichiometric binding. In either case, addition of the APP ligand enhanced thermostability by 8°C but did not affect solubility. A structural determination of the APP-bound AIDA-1 PTB domain was not pursued because there were fewer HN resonances a ^{15}N-edited HSQC spectrum of the bound PTB domain (~120) versus the free PTB domain (~131) suggesting that ligand and binding cleft were severely line broadened beyond detection **(Figure 2.1).**

While the HSQC spectra of the Y6A, F16A, F24A, Y70A and Y131A PTB domains were all qualitatively similar in terms of chemical shifts and line widths, the F24A mutant spectrum was least similar to the other four mutant spectra under closer inspection suggesting that A24 could be making more structural contributions than the other alanine substitutions. Before the structure was determined (an ensemble of structures is shown in **Figure 2.2b**), we assessed the surface exposure of each aromatic-alanine substitution by examining the NOEs observed from the side chain methyl group. As shown in **Figure 2.2c**, only intramolecular and short range intermolecular NOEs were observed at A6, A16 and A70, suggesting that these methyl groups were significantly solvent-exposed. This was certainly the case for A6 as the chemical shift assignments indicated that the first 15 amino acids of the PTB domain were unstructured. Long-range NOEs

were observed between the methyl group of A24 in β1 and the side chains of the adjacent β-strand (β7), specifically, the aromatic ring of F125 and the side chain of T123. The portion of the β-sheet in which substitution A24 resides was deemed to be resistant to hydrogen exchange as an NOE was observed between the methyl group of A24 and its own backbone amide despite the protein being dissolved in D_2O. Taken together, these observations suggested that A24 was the least surface exposed of the five mutants chosen for the study. Once the structure determination was completed (a cartoon representation is shown in **Figure 2.2d**), these observations were confirmed and the F24A substitution appeared to be accommodated well. A PTB domain variant lacking the F24A substitution was not pursued because APP binding activity was unaffected.

The structure of the AIDA-1 PTB5M mutant was aided substantially from data acquired at high field. A statistical summary is provided in **Table 2.2.** Overall, and as somewhat anticipated, the structure compares favorably to the other PTB domains that bind APP **(Table 2.3)**. The PTB domain family can be divided into three major classes, namely Shc-like, IRS1-like and Dab-like (Margolis *et al.,* 1999; Uhlik *et al.,* 2005): The AIDA-1 PTB domain is a representative of the Dab-like class that binds non-phosphorylated-tyrosine peptides. While essentially complete chemical shift assignments were made, the α1-β2 loop spanning Q51-P62 remains unstructured and consequently dynamic due to a lack of long range NOEs observed throughout the region. The β6-β7 loop spanning K110-H116 also samples more conformations on average, supported by the observation that no resonance assignments could be attributed to N115.

Structural and biochemical investigations of the Fe65 PTB2 domain demonstrated >100-fold difference in affinity between an 11 aa. minimal sequence (K_d=100 μM) and an amino terminally extended 32 aa. (K_d=0.2 μM) (Mulvihill *et al.,* 2011; Radzimanowski *et al.,* 2008). One threonine (T668) in APP located in this extended region is susceptible to phosphorylation and acts as a switch that repartitions the cis and trans states of the adjacent proline (P669) that, in turn, affects the ability of Fe65 to engage its ligand *(Ando* et al*., 2001; Ramelot & Nicholson, 2001).* Titrations of long (APP32) and short (APP17) peptides showed no differences in binding affinity to the AIDA-1 PTB domain suggesting that AIDA-1, like many other PTB domains, binds an NPxY motif with a K_d of ~10 μM **(Figure 2.3** and **Table 2.4)**. As predicted from the NMR structure, a semi-solubilizing Y70A single variant or the fully-solubilizing PTB5M variant had no effect on the affinity of the AIDA-1 PTB domain to APP. An APP peptide bearing a phosphorylated Y687 did not bind the AIDA-1 PTB domain providing further evidence for its inclusion in the Dab-like family. The K_d of the X11 PTB domain with a short APP peptide (14 aa., which is comparable to APP17 used in this study) is 0.3 μM, or over 100× stronger than the AIDA-1 PTB domain (Zhang, 1997). From the perspective of the AIDA-1 PTB domain, though, a lower affinity may not necessarily decrease its occupancy on APP relative to X11 and others, as the effective concentration of AIDA-1 within the PSD is extremely high.

Table 2. 2 – Restraints and statistics for the ensemble of 20 Structures.

NOE restraints	
Total	1127
Intraresidue ($\|i - j\| = 0$)	526
Sequential ($\|i - j\| = 1$)	183
Medium range ($1 < \|i - j\| < 5$)	125
Long range ($\|i - j\| \geq 5$)	293
Additional restraints	
Hydrogen bond distance restraints (HN-N/HN-O)	58
Backbone angle torsional angle restraints	96
RMS deviations from ideality[a]	
Bonds (Å)	0.006±0.000
Angles (°)	0.557±0.021
Improper angles (°)	0.811±0.063
RMS violations	
NOE restraints	0.042±0.003
Dihedral angles (°)	0.051±0.067
Ramachandran analysis for ordered residues[b]	
Most favored regions	90.7%
Additional allowed regions	9.3%
Generously allowed regions	0.0%
Disallowed regions	0.0%
RMSD to average coordinates for ordered residues	
Backbone atoms (Å)	0.7
Heavy atoms (Å)	1.2

[a]As reported by XPLOR-NIH 2.33 using the standard protein force field.
[b]As reported by PROCHECK and selected by PSVS 1.4 for residues 17–51, 62–69, 72–114, 116–145.
doi:10.1371/journal.pone.0065605.t002

Table 2. 3 – Structural similarity of the AIDA-1 PTB domain to related PTB domains that also bind APP.

PDB	Protein	Source	RMSD	Aligned	Identity	Reference
2M38	AIDA1	NMR	0.0 Å	134 aa	100%	this study
1AQC	X11+ APP peptide	X-ray	1.4 Å	109 aa	26%	[17]
1P3R	DAB1	X-ray	1.6 Å	115 aa	27%	[25]
2ELA	APPL1	X-ray	1.7 Å	121 aa	16%	[23]
3DXC	Fe65+ APP peptide	X-ray	1.8 Å	120 aa	20%	[20]

doi:10.1371/journal.pone.0065605.t003

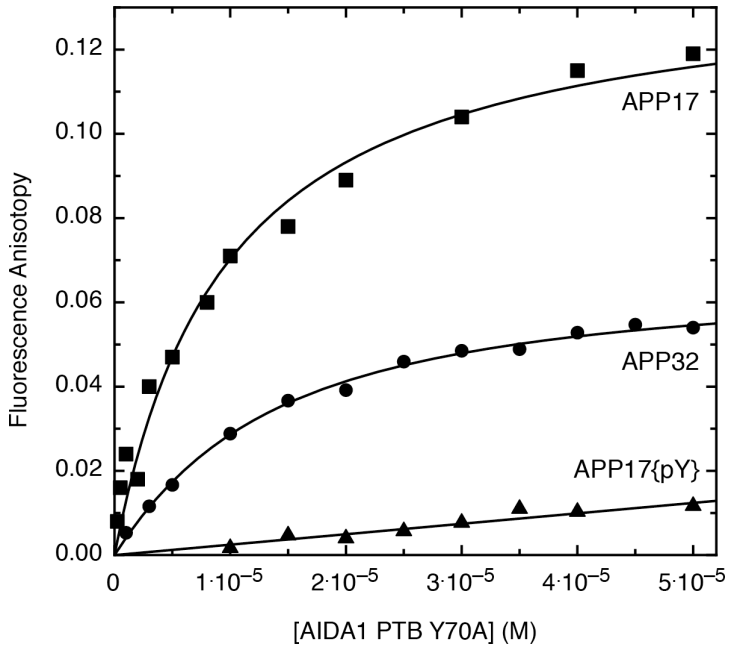


Figure 2. 3 – Titration of FITC-labeled APP peptides with a solubility enhanced mutant (Y70A) of the AIDA-1 PTB domain. Binding was monitored by fluorescence anisotropy. Legend: APP17, a short X11-like binding site; APP32, a longer Fe65-like binding site; APP17{pY}, a short X11-like phosphopeptide. The peptide sequences are listed in **Table 2.4**.

Table 2. 4 – Affinities of APP-derived peptides for two solubility enhanced mutants of the AIDA-1 PTB domain.

Ligand	Peptide sequence	K_d (μM)	
		PTBY70A	PTB5M
APP-32	ggDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN	13.3±1.5 (n = 4)	14.8±1.0 (n = 4)
APP-17	ggQNGYENPTYKFFEQMQN	11.3±1.9 (n = 2)	11.5±1.4 (n = 2)
APP-17{pY}	ggQNGYENPT{pY}KFFEQMQN	ND	>1000(n = 1)

doi:10.1371/journal.pone.0065605.t004

ND: not done.

A comparison of the binding clefts of the AIDA-1, X11 and Fe65 PTB domains is shown in **Figure 2.4**. The cleft of each PTB domain draws contributions from several secondary structures including a short, conserved 3_{10} helix, strands β5/β6 and helix α2. Surveying down the cleft, the first tyrosine of the APP GYENPTYKFFEQ peptide is positioned such that it is predominantly making contacts with G138 and F141 in α2 Analysis of the ensemble of peptides bound to the cleft from the Rosetta based docking simulation identifies an almost equal population of rotamers that place the tyrosine in an analogous position to what is depicted in the X11-PTB/APP complex. The alternative rotamer would contact I134 and L135 in α2 of the AIDA-1 PTB domain. The second tyrosine of the APP GYENPTYKFFEQ peptide is contacted by a different set of amino acids among AIDA-1, X11 and Fe65. In AIDA-1, these residues are N91 in 3_{10} helix and K110 in β6. In X11 and Fe65, there is at least one supporting hydrophobic residue. The first of two consecutive phenylalanines in the APP GYENPTYKFFEQ peptide is supported by a tyrosine in all three PTB domain compared (Y145 in AIDA-1).

A 12-mer SPOT peptide array **(Figures 2.5a** and **2.5b)** was used to survey the amino acid preference of the AIDA-1 PTB domain for APP and APP-like peptides. From an initial window scan of the APP carboxy terminal cytosolic region **(Figure 2.5c)**, a minimal binding sequence of YENPTYKFFE was observed that is consistent with previously described peptide titrations and docking simulations. The minimal binding sequence was then used to exhaustively survey each position in the form of an 'alphabet array' (exhaustive amino acid substitutions at each position in the peptide). The results, summarized in **Figure 2.5d**, present a consensus sequence of YxNxΦYxΨFE where Φ is

a hydrophobic amino acid and Ψ is an aromatic amino acid. Since the requirement for proline in the NPxY motif is not absolute, AIDA-1 has the potential to sample NxxY motifs in receptors such as Ret that guides the development of neurons in the enteric nervous system (L. Li *et al.,* 2006). If this is the case, a higher K_d, and consequently a higher off-rate, would permit more 'handshaking' or sampling of potential protein partners to occur.

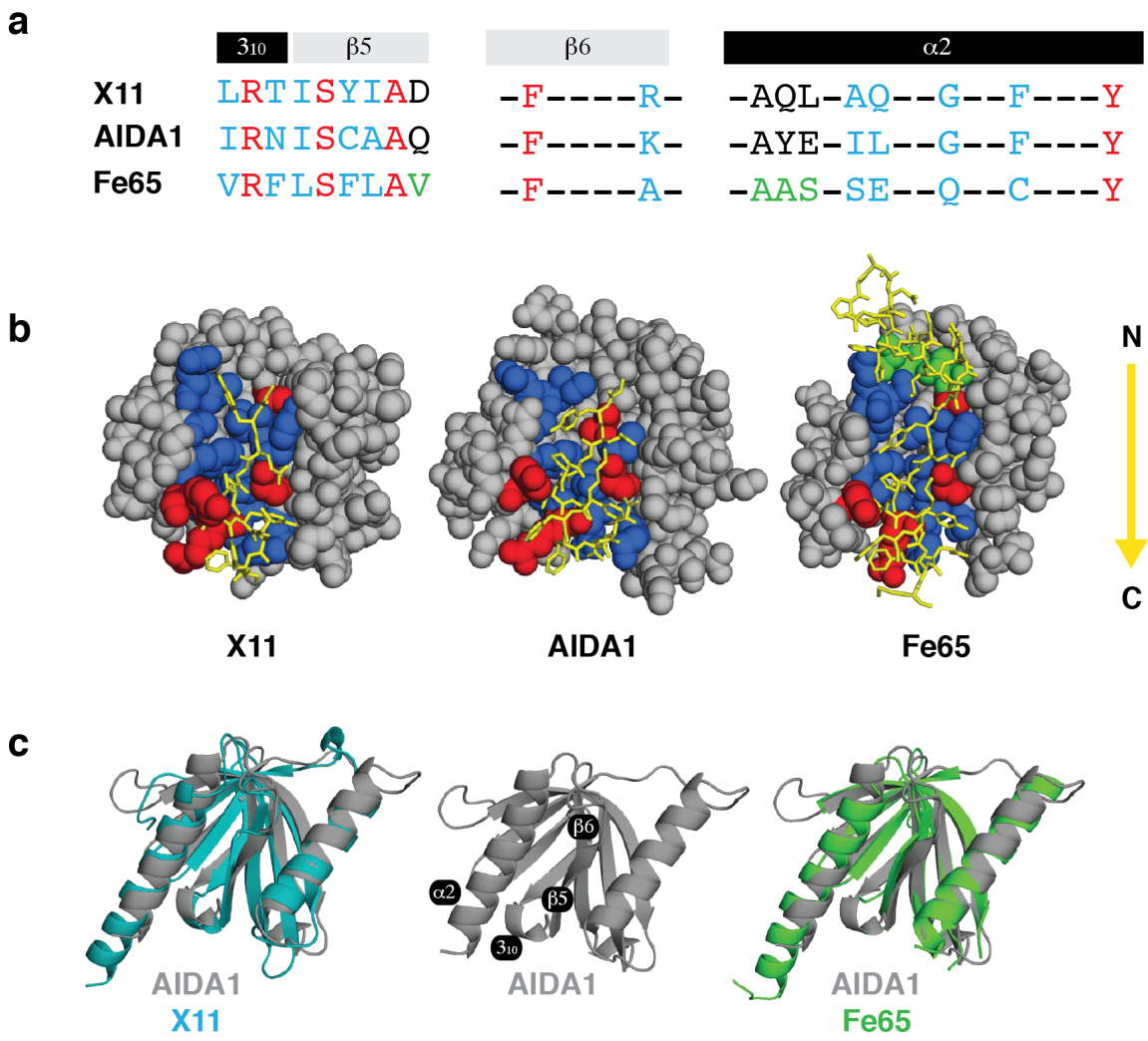


Figure 2. 4 – Interaction of a APP derived peptide (GYENPTYKFFE, shared among all) with the X11, Fe65 and AIDA-1 PTB domains. **(a)** Sequence alignment of amino acids that contribute to the binding cleft; identity in red, homology in blue. The Fe65 PTB domain recognized a longer APP sequence, amino acids that extend its cleft are shown in green. **(b)** APP (yellow, in stick format, N-C direction follows the arrow) interacting with the X11/Fe65 as determined from their respective X-ray structures and with AIDA-1 determined from a molecular docking simulation. **(c)** Backbone alignment of the AIDA-1 (grey), X11 (cyan) and Fe65 (green) PTB domain in the same orientation as **(b)** with the binding cleft facing forward.

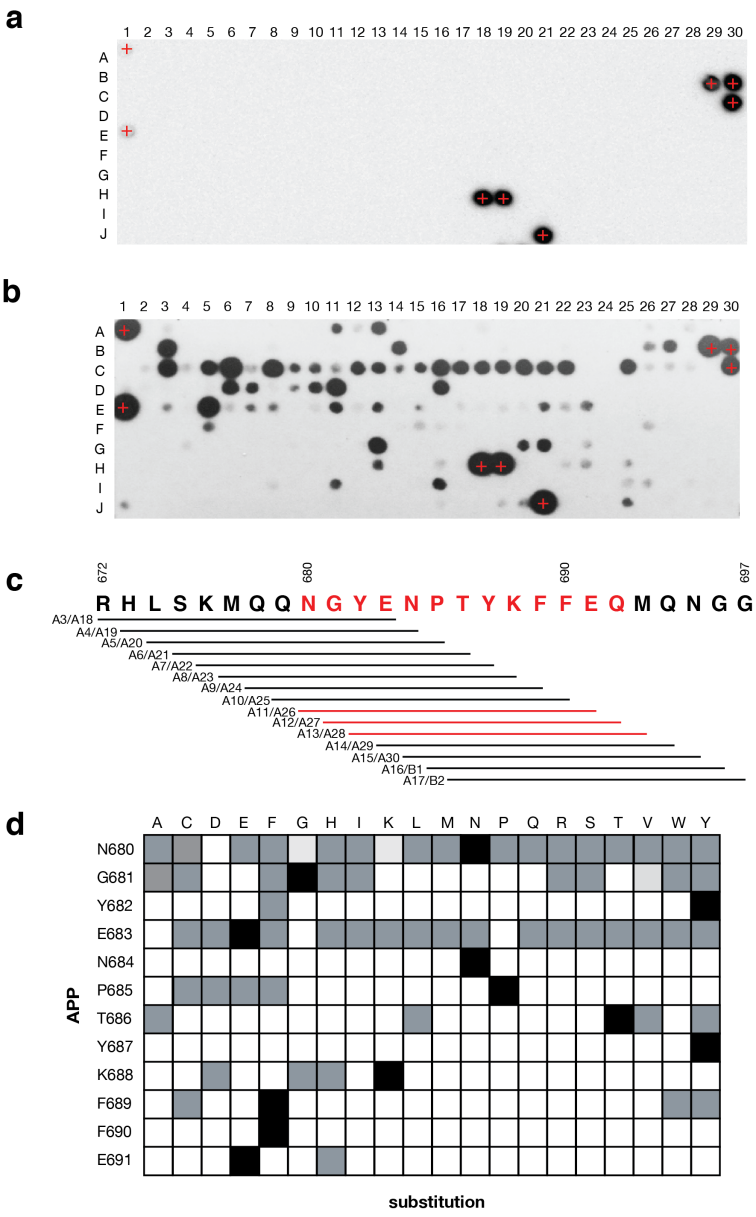


Figure 2. 5 – Amino acid preferences of the AIDA-1 PTB domain for APP determined from a peptide array. A list of peptides on the array are provided in **Table 2.5**. **(a)** The array probed with anti 6xHis mAb only. Positive control 6xHis peptides are identified by a (+). **(b)** The array probed with 6xHis-AIDA-1 PTB domain. **(c)** Sliding window peptide scan of 12-mers spanning aa. 672–697 of APP. Peptides are duplicated on the array; for example, at A3 and A18. Since peptide content per spot can vary, if a signal was observed at the exposure presented it was deemed to be interaction. **(d)** Results of a window scan across the APP C-terminal sequence and an exhaustive positional scan. Grey boxes indicate binding was observed, regardless of signal intensity.

2.4. Discussion

Our initial attempts at biochemical and structural studies of the AIDA-1 PTB domain were precluded by poor solubility. As a result, we made five aromatic-alanine substitutions. While individual substitutions were helpful, it was the combination of all five substitutions that increased solubility to extent that an NMR structure determination was possible. In addition to the solution structure of the AIDA-1 PTB domain, we have determined that its affinity for unphosphorylated APP is moderate relative to similar APP binding proteins such as X11/Fe65 for which dissociation constants of <1 μM have been observed. This difference in affinity may be advantageous for AIDA-1 to participate in signaling contexts beyond APP. From a peptide array study, we determined that the consensus sequence is less stringent (NxxY versus NPxY) for others in the same Dab-like class of PTB domains. Thus, at the neuronal synapse, AIDA-1 could serve as a versatile collator and convenor of signaling events arising from the NMDA receptor, and possibly others.

The structural studies have revealed how the PTB domains of X11 (Matos *et al.,* 2012) and Talin (Goksoy *et al.,* 2008) are autoinhibited by flanking sequences. The AIDA-1-APP interaction is antagonized by a short 26 aa. sequence specified by exon14 in some isoforms through an unknown mechanism (Ghersi *et al.,* 2004). The sequence itself, rich in hydrophobic amino acids, does not resemble the NPxY motif suggesting that regulation of the AIDA-1 PTB domain may be occurring by non-competitive binding. Further structural and biochemical studies of AIDA-1 may lead to the discovery of the selective modification of some neuronal signaling pathways while sparing others. Fine control of

signaling pathways may be one strategy to improve preventive and anti-progression therapies of Alzheimer's disease.

Table 2. 5 – A complete list of 12-mer peptide sequences on the APP peptide array presented in Figure 2.4.

601	A 1	H-H-H-H-H-H-H-H-H-H-H-H	635	B 5	N-G-A-A-N-P-T-Y-K-F-F-E
603	A 3	R-H-L-S-K-M-Q-Q-N-G-Y-E	636	B 6	N-G-Y-A-A-P-T-Y-K-F-F-E
604	A 4	H-L-S-K-M-Q-Q-N-G-Y-E-N	637	B 7	N-G-Y-E-A-A-T-Y-K-F-F-E
605	A 5	L-S-K-M-Q-Q-N-G-Y-E-N-P	638	B 8	N-G-Y-E-N-A-A-Y-K-F-F-E
606	A 6	S-K-M-Q-Q-N-G-Y-E-N-P-T	639	B 9	N-G-Y-E-N-P-A-A-K-F-F-E
607	A 7	K-M-Q-Q-N-G-Y-E-N-P-T-Y	640	B10	N-G-Y-E-N-P-T-A-A-F-F-E
608	A 8	M-Q-Q-N-G-Y-E-N-P-T-Y-K	641	B11	N-G-Y-E-N-P-T-Y-A-A-F-E
609	A 9	Q-Q-N-G-Y-E-N-P-T-Y-K-F	642	B12	N-G-Y-E-N-P-T-Y-K-A-A-E
610	A10	Q-N-G-Y-E-N-P-T-Y-K-F-F	643	B13	N-G-Y-E-N-P-T-Y-K-F-A-A
611	A11	N-G-Y-E-N-P-T-Y-K-F-F-E	644	B14	Q-G-Y-E-N-P-T-Y-K-F-F-E
612	A12	G-Y-E-N-P-T-Y-K-F-F-E-Q	645	B15	N-Q-Y-E-N-P-T-Y-K-F-F-E
613	A13	Y-E-N-P-T-Y-K-F-F-E-Q-M	646	B16	N-G-Q-E-N-P-T-Y-K-F-F-E
614	A14	E-N-P-T-Y-K-F-F-E-Q-M-Q	647	B17	N-G-Y-Q-N-P-T-Y-K-F-F-E
615	A15	N-P-T-Y-K-F-F-E-Q-M-Q-N	648	B18	N-G-Y-E-Q-P-T-Y-K-F-F-E
616	A16	P-T-Y-K-F-F-E-Q-M-Q-N-G	649	B19	N-G-Y-E-N-Q-T-Y-K-F-F-E
617	A17	T-Y-K-F-F-E-Q-M-Q-N-G-G	650	B20	N-G-Y-E-N-P-Q-Y-K-F-F-E
618	A18	R-H-L-S-K-M-Q-Q-N-G-Y-E	651	B21	N-G-Y-E-N-P-T-Q-K-F-F-E
619	A19	H-L-S-K-M-Q-Q-N-G-Y-E-N	652	B22	N-G-Y-E-N-P-T-Y-Q-F-F-E
620	A20	L-S-K-M-Q-Q-N-G-Y-E-N-P	653	B23	N-G-Y-E-N-P-T-Y-K-Q-F-E
621	A21	S-K-M-Q-Q-N-G-Y-E-N-P-T	654	B24	N-G-Y-E-N-P-T-Y-K-F-Q-E
622	A22	K-M-Q-Q-N-G-Y-E-N-P-T-Y	655	B25	N-G-Y-E-N-P-T-Y-K-F-F-Q
623	A23	M-Q-Q-N-G-Y-E-N-P-T-Y-K	656	B26	A-G-Y-E-N-P-T-Y-K-F-F-E
624	A24	Q-Q-N-G-Y-E-N-P-T-Y-K-F	657	B27	C-G-Y-E-N-P-T-Y-K-F-F-E
625	A25	Q-N-G-Y-E-N-P-T-Y-K-F-F	658	B28	A-A-A-A-A-A-A-A-A-A-A-A
626	A26	N-G-Y-E-N-P-T-Y-K-F-F-E	659	B29	H-H-H-H-H-H-H-H-H-H-H-H
627	A27	G-Y-E-N-P-T-Y-K-F-F-E-Q	660	B30	A-A-A-H-H-H-H-H-H-A-A-A
628	A28	Y-E-N-P-T-Y-K-F-F-E-Q-M	661	C 1	D-G-Y-E-N-P-T-Y-K-F-F-E
629	A29	E-N-P-T-Y-K-F-F-E-Q-M-Q	662	C 2	E-G-Y-E-N-P-T-Y-K-F-F-E
630	A30	N-P-T-Y-K-F-F-E-Q-M-Q-N	663	C 3	F-G-Y-E-N-P-T-Y-K-F-F-E
631	B 1	P-T-Y-K-F-F-E-Q-M-Q-N-G	664	C 4	G-G-Y-E-N-P-T-Y-K-F-F-E
632	B 2	T-Y-K-F-F-E-Q-M-Q-N-G-G	665	C 5	H-G-Y-E-N-P-T-Y-K-F-F-E
633	B 3	A-A-Y-E-N-P-T-Y-K-F-F-E	666	C 6	I-G-Y-E-N-P-T-Y-K-F-F-E
634	B 4	N-A-A-E-N-P-T-Y-K-F-F-E	667	C 7	K-G-Y-E-N-P-T-Y-K-F-F-E

668	C 8	L-G-Y-E-N-P-T-Y-K-F-F-E
669	C 9	M-G-Y-E-N-P-T-Y-K-F-F-E
670	C10	N-G-Y-E-N-P-T-Y-K-F-F-E
671	C11	P-G-Y-E-N-P-T-Y-K-F-F-E
672	C12	Q-G-Y-E-N-P-T-Y-K-F-F-E
673	C13	R-G-Y-E-N-P-T-Y-K-F-F-E
674	C14	S-G-Y-E-N-P-T-Y-K-F-F-E
675	C15	T-G-Y-E-N-P-T-Y-K-F-F-E
676	C16	V-G-Y-E-N-P-T-Y-K-F-F-E
677	C17	W-G-Y-E-N-P-T-Y-K-F-F-E
678	C18	Y-G-Y-E-N-P-T-Y-K-F-F-E
679	C19	W-G-Y-E-N-P-T-Y-K-F-F-E
680	C20	Y-G-Y-E-N-P-T-Y-K-F-F-E
681	C21	N-A-Y-E-N-P-T-Y-K-F-F-E
682	C22	N-C-Y-E-N-P-T-Y-K-F-F-E
683	C23	N-D-Y-E-N-P-T-Y-K-F-F-E
684	C24	N-E-Y-E-N-P-T-Y-K-F-F-E
685	C25	N-F-Y-E-N-P-T-Y-K-F-F-E
686	C26	N-G-Y-E-N-P-T-Y-K-F-F-E
687	C27	N-H-Y-E-N-P-T-Y-K-F-F-E
688	C28	N-I-Y-E-N-P-T-Y-K-F-F-E
689	C29	N-K-Y-E-N-P-T-Y-K-F-F-E
690	C30	A-A-A-H-H-H-H-H-H-A-A-A
691	D 1	N-L-Y-E-N-P-T-Y-K-F-F-E
692	D 2	N-M-Y-E-N-P-T-Y-K-F-F-E
693	D 3	N-N-Y-E-N-P-T-Y-K-F-F-E
694	D 4	N-P-Y-E-N-P-T-Y-K-F-F-E
695	D 5	N-Q-Y-E-N-P-T-Y-K-F-F-E
696	D 6	N-R-Y-E-N-P-T-Y-K-F-F-E
697	D 7	N-S-Y-E-N-P-T-Y-K-F-F-E
698	D 8	N-T-Y-E-N-P-T-Y-K-F-F-E
699	D 9	N-V-Y-E-N-P-T-Y-K-F-F-E
700	D10	N-W-Y-E-N-P-T-Y-K-F-F-E
701	D11	N-Y-Y-E-N-P-T-Y-K-F-F-E
702	D12	N-G-A-E-N-P-T-Y-K-F-F-E
703	D13	N-G-C-E-N-P-T-Y-K-F-F-E
704	D14	N-G-D-E-N-P-T-Y-K-F-F-E
705	D15	N-G-E-E-N-P-T-Y-K-F-F-E
706	D16	N-G-F-E-N-P-T-Y-K-F-F-E
707	D17	N-G-G-E-N-P-T-Y-K-F-F-E
708	D18	N-G-H-E-N-P-T-Y-K-F-F-E
709	D19	N-G-I-E-N-P-T-Y-K-F-F-E
710	D20	N-G-K-E-N-P-T-Y-K-F-F-E
711	D21	N-G-L-E-N-P-T-Y-K-F-F-E
712	D22	N-G-M-E-N-P-T-Y-K-F-F-E
713	D23	N-G-N-E-N-P-T-Y-K-F-F-E
714	D24	N-G-P-E-N-P-T-Y-K-F-F-E
715	D25	N-G-Q-E-N-P-T-Y-K-F-F-E
716	D26	N-G-R-E-N-P-T-Y-K-F-F-E
717	D27	N-G-S-E-N-P-T-Y-K-F-F-E
718	D28	N-G-T-E-N-P-T-Y-K-F-F-E
719	D29	N-G-V-E-N-P-T-Y-K-F-F-E
720	D30	N-G-W-E-N-P-T-Y-K-F-F-E
721	E 1	A-A-A-H-H-H-H-H-H-A-A-A
722	E 2	A-A-A-A-A-A-A-A-A-A-A-A
723	E 3	N-G-Y-E-N-P-T-Y-K-F-F-E
724	E 4	N-G-Y-A-N-P-T-Y-K-F-F-E
725	E 5	N-G-Y-C-N-P-T-Y-K-F-F-E
726	E 6	N-G-Y-D-N-P-T-Y-K-F-F-E
727	E 7	N-G-Y-E-N-P-T-Y-K-F-F-E
728	E 8	N-G-Y-F-N-P-T-Y-K-F-F-E
729	E 9	N-G-Y-G-N-P-T-Y-K-F-F-E
730	E10	N-G-Y-H-N-P-T-Y-K-F-F-E
731	E11	N-G-Y-I-N-P-T-Y-K-F-F-E
732	E12	N-G-Y-K-N-P-T-Y-K-F-F-E
733	E13	N-G-Y-L-N-P-T-Y-K-F-F-E
734	E14	N-G-Y-M-N-P-T-Y-K-F-F-E
735	E15	N-G-Y-N-N-P-T-Y-K-F-F-E
736	E16	N-G-Y-P-N-P-T-Y-K-F-F-E
737	E17	N-G-Y-Q-N-P-T-Y-K-F-F-E
738	E18	N-G-Y-R-N-P-T-Y-K-F-F-E
739	E19	N-G-Y-S-N-P-T-Y-K-F-F-E
740	E20	N-G-Y-T-N-P-T-Y-K-F-F-E
741	E21	N-G-Y-V-N-P-T-Y-K-F-F-E
742	E22	N-G-Y-W-N-P-T-Y-K-F-F-E
743	E23	N-G-Y-Y-N-P-T-Y-K-F-F-E
744	E24	N-G-Y-E-A-P-T-Y-K-F-F-E
745	E25	N-G-Y-E-C-P-T-Y-K-F-F-E

746	E26	N-G-Y-E-D-P-T-Y-K-F-F-E
747	E27	N-G-Y-E-E-P-T-Y-K-F-F-E
748	E28	N-G-Y-E-F-P-T-Y-K-F-F-E
749	E29	N-G-Y-E-G-P-T-Y-K-F-F-E
750	E30	N-G-Y-E-H-P-T-Y-K-F-F-E
751	F 1	N-G-Y-E-I-P-T-Y-K-F-F-E
752	F 2	N-G-Y-E-K-P-T-Y-K-F-F-E
753	F 3	N-G-Y-E-L-P-T-Y-K-F-F-E
754	F 4	N-G-Y-E-M-P-T-Y-K-F-F-E
755	F 5	N-G-Y-E-N-P-T-Y-K-F-F-E
756	F 6	N-G-Y-E-P-P-T-Y-K-F-F-E
757	F 7	N-G-Y-E-Q-P-T-Y-K-F-F-E
758	F 8	N-G-Y-E-R-P-T-Y-K-F-F-E
759	F 9	N-G-Y-E-S-P-T-Y-K-F-F-E
760	F10	N-G-Y-E-T-P-T-Y-K-F-F-E
761	F11	N-G-Y-E-V-P-T-Y-K-F-F-E
762	F12	N-G-Y-E-W-P-T-Y-K-F-F-E
763	F13	N-G-Y-E-Y-P-T-Y-K-F-F-E
764	F14	N-G-Y-E-N-A-T-Y-K-F-F-E
765	F15	N-G-Y-E-N-C-T-Y-K-F-F-E
766	F16	N-G-Y-E-N-D-T-Y-K-F-F-E
767	F17	N-G-Y-E-N-E-T-Y-K-F-F-E
768	F18	N-G-Y-E-N-F-T-Y-K-F-F-E
769	F19	N-G-Y-E-N-G-T-Y-K-F-F-E
770	F20	N-G-Y-E-N-H-T-Y-K-F-F-E
771	F21	N-G-Y-E-N-I-T-Y-K-F-F-E
772	F22	N-G-Y-E-N-K-T-Y-K-F-F-E
773	F23	N-G-Y-E-N-L-T-Y-K-F-F-E
774	F24	N-G-Y-E-N-M-T-Y-K-F-F-E
775	F25	N-G-Y-E-N-N-T-Y-K-F-F-E
776	F26	N-G-Y-E-N-P-T-Y-K-F-F-E
777	F27	N-G-Y-E-N-Q-T-Y-K-F-F-E
778	F28	N-G-Y-E-N-R-T-Y-K-F-F-E
779	F29	N-G-Y-E-N-S-T-Y-K-F-F-E
780	F30	N-G-Y-E-N-T-T-Y-K-F-F-E
781	G 1	N-G-Y-E-N-V-T-Y-K-F-F-E
782	G 2	N-G-Y-E-N-W-T-Y-K-F-F-E
783	G 3	N-G-Y-E-N-Y-T-Y-K-F-F-E
784	G 4	N-G-Y-E-N-P-A-Y-K-F-F-E
785	G 5	N-G-Y-E-N-P-C-Y-K-F-F-E
786	G 6	N-G-Y-E-N-P-D-Y-K-F-F-E
787	G 7	N-G-Y-E-N-P-E-Y-K-F-F-E
788	G 8	N-G-Y-E-N-P-F-Y-K-F-F-E
789	G 9	N-G-Y-E-N-P-G-Y-K-F-F-E
790	G10	N-G-Y-E-N-P-H-Y-K-F-F-E
791	G11	N-G-Y-E-N-P-I-Y-K-F-F-E
792	G12	N-G-Y-E-N-P-K-Y-K-F-F-E
793	G13	N-G-Y-E-N-P-L-Y-K-F-F-E
794	G14	N-G-Y-E-N-P-M-Y-K-F-F-E
795	G15	N-G-Y-E-N-P-N-Y-K-F-F-E
796	G16	N-G-Y-E-N-P-P-Y-K-F-F-E
797	G17	N-G-Y-E-N-P-Q-Y-K-F-F-E
798	G18	N-G-Y-E-N-P-R-Y-K-F-F-E
799	G19	N-G-Y-E-N-P-S-Y-K-F-F-E
800	G20	N-G-Y-E-N-P-T-Y-K-F-F-E
801	G21	N-G-Y-E-N-P-V-Y-K-F-F-E
802	G22	N-G-Y-E-N-P-W-Y-K-F-F-E
803	G23	N-G-Y-E-N-P-Y-Y-K-F-F-E
804	G24	N-G-Y-E-N-P-T-A-K-F-F-E
805	G25	N-G-Y-E-N-P-T-C-K-F-F-E
806	G26	N-G-Y-E-N-P-T-D-K-F-F-E
807	G27	N-G-Y-E-N-P-T-E-K-F-F-E
808	G28	N-G-Y-E-N-P-T-F-K-F-F-E
809	G29	N-G-Y-E-N-P-T-G-K-F-F-E
810	G30	N-G-Y-E-N-P-T-H-K-F-F-E
811	H 1	N-G-Y-E-N-P-T-I-K-F-F-E
812	H 2	N-G-Y-E-N-P-T-K-K-F-F-E
813	H 3	N-G-Y-E-N-P-T-L-K-F-F-E
814	H 4	N-G-Y-E-N-P-T-M-K-F-F-E
815	H 5	N-G-Y-E-N-P-T-N-K-F-F-E
816	H 6	N-G-Y-E-N-P-T-P-K-F-F-E
817	H 7	N-G-Y-E-N-P-T-Q-K-F-F-E
818	H 8	N-G-Y-E-N-P-T-R-K-F-F-E
819	H 9	N-G-Y-E-N-P-T-S-K-F-F-E
820	H10	N-G-Y-E-N-P-T-T-K-F-F-E
821	H11	N-G-Y-E-N-P-T-V-K-F-F-E
822	H12	N-G-Y-E-N-P-T-W-K-F-F-E
823	H13	N-G-Y-E-N-P-T-Y-K-F-F-E

824	H14	N-G-Y-E-N-P-T-Y-A-F-F-E
825	H15	N-G-Y-E-N-P-T-Y-C-F-F-E
826	H16	N-G-Y-E-N-P-T-Y-D-F-F-E
827	H17	N-G-Y-E-N-P-T-Y-E-F-F-E
828	H18	A-A-A-H-H-H-H-H-H-A-A-A
829	H19	G-S-H-H-H-H-H-H-G-S-S-A
830	H20	A-A-A-A-A-A-A-A-A-A-A-A
831	H21	N-G-Y-E-N-P-T-Y-F-F-F-E
832	H22	N-G-Y-E-N-P-T-Y-G-F-F-E
833	H23	N-G-Y-E-N-P-T-Y-H-F-F-E
834	H24	N-G-Y-E-N-P-T-Y-I-F-F-E
835	H25	N-G-Y-E-N-P-T-Y-K-F-F-E
836	H26	N-G-Y-E-N-P-T-Y-L-F-F-E
837	H27	N-G-Y-E-N-P-T-Y-M-F-F-E
838	H28	N-G-Y-E-N-P-T-Y-N-F-F-E
839	H29	N-G-Y-E-N-P-T-Y-P-F-F-E
840	H30	N-G-Y-E-N-P-T-Y-Q-F-F-E
841	I 1	N-G-Y-E-N-P-T-Y-R-F-F-E
842	I 2	N-G-Y-E-N-P-T-Y-S-F-F-E
843	I 3	N-G-Y-E-N-P-T-Y-T-F-F-E
844	I 4	N-G-Y-E-N-P-T-Y-V-F-F-E
845	I 5	N-G-Y-E-N-P-T-Y-W-F-F-E
846	I 6	N-G-Y-E-N-P-T-Y-Y-F-F-E
847	I 7	N-G-Y-E-N-P-T-Y-K-A-F-E
848	I 8	N-G-Y-E-N-P-T-Y-K-C-F-E
849	I 9	N-G-Y-E-N-P-T-Y-K-D-F-E
850	I10	N-G-Y-E-N-P-T-Y-K-E-F-E
851	I11	N-G-Y-E-N-P-T-Y-K-F-F-E
852	I12	N-G-Y-E-N-P-T-Y-K-G-F-E
853	I13	N-G-Y-E-N-P-T-Y-K-H-F-E
854	I14	N-G-Y-E-N-P-T-Y-K-I-F-E
855	I15	N-G-Y-E-N-P-T-Y-K-K-F-E
856	I16	N-G-Y-E-N-P-T-Y-K-L-F-E
857	I17	N-G-Y-E-N-P-T-Y-K-M-F-E
858	I18	N-G-Y-E-N-P-T-Y-K-N-F-E
859	I19	N-G-Y-E-N-P-T-Y-K-P-F-E
860	I20	N-G-Y-E-N-P-T-Y-K-Q-F-E
861	I21	N-G-Y-E-N-P-T-Y-K-R-F-E
862	I22	N-G-Y-E-N-P-T-Y-K-S-F-E
863	I23	N-G-Y-E-N-P-T-Y-K-T-F-E
864	I24	N-G-Y-E-N-P-T-Y-K-V-F-E
865	I25	N-G-Y-E-N-P-T-Y-K-W-F-E
866	I26	N-G-Y-E-N-P-T-Y-K-Y-F-E
867	I27	N-G-Y-E-N-P-T-Y-K-F-A-E
868	I28	N-G-Y-E-N-P-T-Y-K-F-C-E
869	I29	N-G-Y-E-N-P-T-Y-K-F-D-E
870	I30	N-G-Y-E-N-P-T-Y-K-F-E-E
871	J 1	N-G-Y-E-N-P-T-Y-K-F-F-E
872	J 2	N-G-Y-E-N-P-T-Y-K-F-G-E
873	J 3	N-G-Y-E-N-P-T-Y-K-F-H-E
875	J 5	N-G-Y-E-N-P-T-Y-K-F-K-E
876	J 6	N-G-Y-E-N-P-T-Y-K-F-L-E
877	J 7	N-G-Y-E-N-P-T-Y-K-F-M-E
878	J 8	N-G-Y-E-N-P-T-Y-K-F-N-E
880	J10	N-G-Y-E-N-P-T-Y-K-F-Q-E
881	J11	N-G-Y-E-N-P-T-Y-K-F-R-E
882	J12	N-G-Y-E-N-P-T-Y-K-F-S-E
883	J13	N-G-Y-E-N-P-T-Y-K-F-T-E
885	J15	N-G-Y-E-N-P-T-Y-K-F-W-E
887	J17	N-G-Y-E-N-P-T-Y-K-F-F-A
888	J18	N-G-Y-E-N-P-T-Y-K-F-F-C
889	J19	N-G-Y-E-N-P-T-Y-K-F-F-D
890	J20	N-G-Y-E-N-P-T-Y-K-F-F-E
891	J21	A-A-A-H-H-H-H-H-H-A-A-A
892	J22	A-A-A-A-A-A-A-A-A-A-A-A
893	J23	N-G-Y-E-N-P-T-Y-K-F-F-F
894	J24	N-G-Y-E-N-P-T-Y-K-F-F-G
895	J25	N-G-Y-E-N-P-T-Y-K-F-F-H
896	J26	N-G-Y-E-N-P-T-Y-K-F-F-I
897	J27	N-G-Y-E-N-P-T-Y-K-F-F-K
898	J28	N-G-Y-E-N-P-T-Y-K-F-F-L
899	J29	N-G-Y-E-N-P-T-Y-K-F-F-M
900	J30	N-G-Y-E-N-P-T-Y-K-F-F-N
901	K 1	N-G-Y-E-N-P-T-Y-K-F-F-P
902	K 2	N-G-Y-E-N-P-T-Y-K-F-F-Q
904	K 4	N-G-Y-E-N-P-T-Y-K-F-F-S
905	K 5	N-G-Y-E-N-P-T-Y-K-F-F-T
908	K 8	N-G-Y-E-N-P-T-Y-K-F-F-Y

CHAPTER 3:

TANDEM SAM DOMAINS DRIVE THE DYNAMIC OLIGOMERIZATION OF THE CASKIN2 NEURONAL SCAFFOLDING PROTEIN

All of the content reported within this chapter has been published in the article listed below:

- Smirnova E., Kwan J.J., Siu R., Gao X., Zoidl G., Demeler B., Saridakis V., and Donaldson L.W. (2016) "A New Mode of SAM Domain Mediated Oligomerization Observed in the Caskin2 Neuronal Scaffolding Protein." *Cell Communication and Signaling*, 1–14.

Information of the authors' contributions is provided in the corresponding section.

3.1. Introduction

Caskin2 and its mammalian homolog Caskin1, are multidomain proteins that share the same overall organization (Tabuchi *et al.,* 2002). The amino terminal half of both proteins consist of protein-protein interaction modules, namely six ankyrin repeats, an SH3 domain, and two SAM domains (**Figure 3.1**). The carboxy terminal half consists of low complexity, proline-rich sequences (Balázs *et al.,* 2009) ending with a conserved 25 aa. segment of unknown function. The Caskins are named for their ability to interact with CASK (calcium / calmodulin-dependent serine kinase), a MAGUK protein that is implicated in a number of neurological conditions including autism and X-linked mental

retardation (Hsueh, 2006; K. Chen and Featherstone, 2011; Corvin, 2010). Only one homolog, *Ckn*, is observed in the *Drosophila* genome (Weng *et al.,* 2011) and no homologs are observed in *C. elegans* suggesting that from an evolutionary perspective, multiple mammalian Caskins may have arisen to promote a more comprehensive set of signaling circuits. In Caskin1, the CASK interaction domain (CID) is located between the SH3 and SAM1 domains and facilitates direct contact with the calmodulin kinase catalytic domain of CASK. The CID is also present in the scaffolding protein, X11/Mint (Stafford *et al.,* 2011). The CID, however, is not present in Caskin2 rendering it unable to bind CASK (Tabuchi *et al.,* 2002). Thus, despite their organizational similarity, Caskin1 and Caskin2 may have diverged with respect to their scaffolding functions in neurons, their structures and their protein partners.

Sterile Alpha Motif (SAM) domains are well represented in the human genome reflecting the versatility of this compact, five-helix fold to facilitate protein-ligand interactions that include other proteins, nucleic acids and lipids (Qiao, 2005). The most prevalent partners of SAM domains are, in fact, other SAM domains leading to a variety of homotypic and heterotypic SAM-SAM interactions in transcription factors (Qiao *et al.,* 2004; Qiao *et al.,* 2006) and neuronal signaling protein assemblies (Baron *et al.* 2006; Bourgeron, 2009; Harada *et al.,* 2008). Because SAM domains generally employ two complementary surfaces, homotypic interactions may produce not only dimers, but also assemblies of SAM domains polymers to highlight the considerable molecular weight they can attain (Knight *et al.,* 2011).

The Caskin1 and Caskin2 tandem SAM domains were first identified to self-associate during an electron microscopy based survey which sought to identify new SAM domain mediated polymers (Knight *et al.,* 2011). Later high resolution X-ray studies revealed that the Caskin1 SAM1-SAM2 tandem self-associated into helical fibrils (Stafford *et al.,* 2011). Two roles have been proposed for this architecture at presynaptic sites. First, oligomers of Caskin1 could link and concentrate cell-adhesion proteins including Ephrin B1 and CASK-associated neurexin. Second, oligomers of Caskin1 could form a tether by which a stream of vesicles loaded with chemical transmitters could be guided via synaptogamin to the synaptic cleft (Stafford *et al.,* 2011).

A crystal structure demonstrating that the tandem SAM domains of Caskin2 form an oligomer that is distinct from Caskin1 is reported in this study. By analytical ultracentrifugation, a dissociation constant describing the monomer-dimer equilibrium of the SAM tandem was observed to be in the micromolar range, a favorable concentration in the cell for tuning oligomerization and opening up the possibility for additional regulation by post-translational modifications and protein partners. An EGFP-tagged Caskin2 SAM1-SAM2 protein expressed in neuroblastoma cells formed punctae consistent with high order oligomers while a structure-directed surface mutant was distributed diffusely. In support of the structural distinction between Caskin1 and Caskin2, the punctae were morphologically different. This study provides a foundation to begin exploring the effect of protein partnerships and post-translational modifications that direct the oligomeric state of Caskin2 and consequently, its function in neurons, possibly apart from the processes directed by Caskin1.

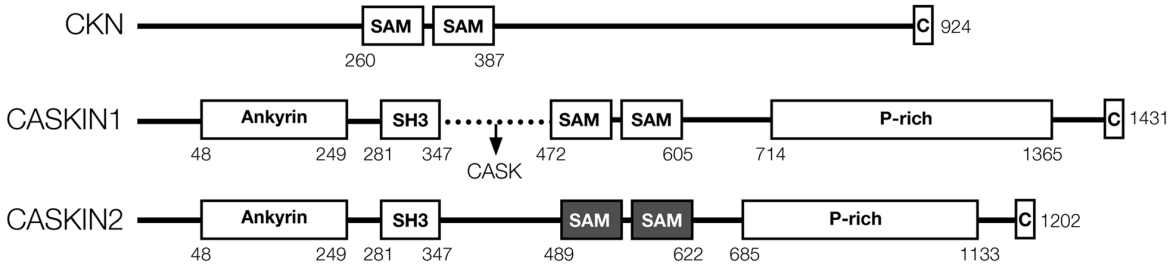


Figure 3. 1 – Conservation of the tandem SAM domains among three neuronal signaling scaffolding proteins, *Drosophila* Ckn, human Caskin1, and human Caskin2. The location of the binding site in Caskin1 for the scaffolding protein CASK is shown by an arrow. The Caskin2 SAM tandem described in this study is shaded *grey*.

3.2. Materials and Methods

3.2.1. Cloning

The human Caskin2 SAM1-SAM2 tandem (483-634; Uniprot Q8WXE0) and individual SAM1 (483-549) and SAM2 (550-634) domains were amplified by PCR from a human cDNA and inserted into the *BamHI* / *XhoI* restriction sites of pET28 (Novagen) followed by transformation into *E. coli* BL21:DE3 to produce a 6xHis tagged protein. Five Caskin2 mutants, G537D, K540E, G537D/K540E, L589E, and G607D were made using the Quikchange method (Agilent). An EGFP fusion protein to the wild type Caskin2 SAM1-SAM2 tandem and G537D/K540E mutant was prepared by inserting a suitable PCR product into the *XhoI* / *KpnI* restriction sites of pEGFP-N1 (Clontech). A similar approach was used to make EGPF-tagged Caskin1 SAM1-SAM2 (470-613; Uniprot Q8WXE9) and a G520D/K523E mutant using a synthetic Caskin1 gene fragment (GenScript).

3.2.2. Expression and Protein Purification

Isotopic labeling of Caskin2 SAM1, SAM2, and SAM1-SAM2 for NMR spectroscopy was achieved by a 1.0 L fermentation in a minimal medium containing 1 g $^{15}NH_4Cl$ as the sole nitrogen source and/or 3 g of ^{13}C-glucose as the sole carbon source. Proteins for X-ray crystallographic studies were expressed in a minimal medium with the addition of 50 mg/L of selenomethionine 15 min before induction. Cell pellets were dissolved in T300 buffer (20 mM Tris-HCl, 300 mM NaCl, 0.05 % NaN_3) and lysed by French press. Highly purified protein was obtained from a two-step purification involving Nickel-NTA affinity

chromatography (Qiagen), followed by gel filtration chromatography (Sephacryl-100, HiLoad 16/60; GE Life Sciences). The final buffer for NMR analyses was phosphate buffered saline (PBS; 20 mM sodium phosphate, pH 7.8, 0.15 M NaCl, 0.05% (w/v) NaN_3. Crystallographic screening was performed with proteins in T300 buffer.

3.2.3. Cell Culture, Transient Transfection and Immunoblotting

Neuroblastoma 2a (Neuro2a) cells (Olmsted *et al.,* 1970) were maintained using standard growth conditions and used for expression and localization studies as described in (Prochnow *et al.,* 2009). 30,000 cells were seeded onto 13 mm glass cover slips in 24 well plates and 200-400 ng plasmid DNA transfected using Effectene reagent as recommended by the manufacturer (Qiagen). Whole cell protein lysates from transfected Neuro2a cells collected 48 hours post-transfection were separated by 10% SDS-PAGE and transferred to 0.2 μm Hybond-ECL nitrocellulose membrane (GE Life Sciences) for immunodetection. Primary antibodies were diluted 1:1000 (rabbit anti-GFP; Santa Cruz) and 1:20000 (mouse anti-β-actin; Sigma-Aldrich). Secondary antibodies (LI-COR Biosciences) were diluted 1:20000 (donkey anti-rabbit IRDye680LT) or 1:20000 (goat anti-mouse IRDye800CW). Signals were detected using the Odyssey Infrared Imaging System (LI-COR Biosciences).

3.2.4. Confocal Microscopy

Transfected cells were fixed with 4% paraformaldehyde for 20 min at room temperature, washed with PBS, counterstained with DAPI and mounted for imaging. Samples were

visualized using a Zeiss LSM 700 confocal microscope with a Plan-Apochromat 63x/1.4 Oil DIC M27 objective and the ZEN 2010 program to control all hardware parameters. Images were collected by line averaging (4x) at high resolution (2048x2048 pixel) using single planes or *z*-stacks. Images were exported and further processed using ImageJ. For deconvolution, the point-spread function was calculated using the Gaussian PSF 3D and Iterative 3D Deconvolve software plugins in ImageJ. Images were combined in Adobe Photoshop for presentation.

3.2.5. Analytical Ultracentrifugation

Sedimentation velocity (SV) experiments were performed with a Beckman Optima XL-I at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies at the University of Texas Health Science Center at San Antonio. SV data were analyzed with UltraScan-III (Gorbet *et al.,* 2015) All calculations were performed on the XSEDE UltraScan Science Gateway using high-performance computing resources at the Texas Advanced Computing Center, at the San Diego Supercomputing Center, and at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio. All measurements were made in 20 mM sodium phosphate buffer, pH 7.8, supplemented with 0.15 M or 0.3 mM NaCl. The experimental data were collected in intensity mode at 20˚C, and at 50,000 rpm, using standard epon-charcoal two-channel centerpieces. Hydrodynamic corrections for buffer density, viscosity and partial specific volume were made as implemented in UltraScan-III, except when equilibrium constants were fitted to whole boundary models. In those cases, the monomer molar mass, which is

known, was held constant, and the partial specific volume was floated to account for the variability in partial specific volume under different salt concentrations. The experimental data were first modeled with solutions of the Lamm equation (Cao and Demeler, 2005), which are fitted to experimental data by two-dimensional spectrum analysis (Brookes, Cao, and Demeler, 2009) using meniscus fitting and simultaneous time- and radially invariant noise removal (Demeler, 2001). Noise corrected data were further analyzed by the enhanced van Holde-Weischet method (Demeler and van Holde, 2004). This approach provides diffusion corrected sedimentation coefficient distributions, providing clear evidence for the presence of heterogeneity, and for identifying reversible mass action reactions. Quantitative equilibrium constants were obtained by fitting analytical ultracentrifugation sedimentation velocity (AUC-SV) experiments by genetic algorithm analysis of as described in (Demeler *et al.,* 2010). 95% confidence intervals were determined by Monte Carlo analysis (Demeler and Brookes, 2007).

3.2.6. NMR Spectroscopy

All experiments were performed with either uniformly ^{15}N-labeled, or ^{13}C,^{15}N-labeled samples, as required. Assignment of the G537D/K540E mutant at 0.8 mM was achieved by a conventional triple resonance strategy (HNCACB, CBCACONH, HNCO, HNCACO) acquired at 310 K with non-linear sampling on a Bruker Avance 950 MHz NMR spectrometer at the Imaging and Characterization Core Laboratory (KAUST). Datasets were processed with a combination of NMRpipe (Delaglio *et al.,* 1995) and istHMS (Hyberts *et al.,* 2012) and interpreted with CCPN Analysis (Skinner *et al.,* 2015).

Backbone ^{15}N relaxation experiments at a protein concentration of 0.3 mM were acquired on a Bruker Avance 700 MHz NMR spectrometer at the York University Life Sciences Building Central Facility. A longitudinal ^{15}N T_1 relaxation rate was determined by acquiring 2D spectra with delays of 200, 400, 600, 800, 1000, and 1200 ms. A transverse ^{15}N T_2 relaxation rate was determined by acquiring 2D spectra with delays of 17, 34, 51, 68, 85, 102, 136, and 170 ms. In both cases, spectra were processed and peaks integrated with NMRPipe and then fit to a single exponential function with LMquick (Farrow *et al.,* 1994). A rotational correlation time (τ_c) was calculated from the average T_1/T_2 ratio (Kay, Torchia, and Bax, 1989). From the correlation time, a molecular weight was estimated according to the linear relationship τ_c = MW * 0.433 + 0.775 published at the University of San Diego NMR Center (*http://sopnmr.ucsd.edu/biomol-tools.htm*).

3.2.7. X-Ray Crystallography

Crystals of selenomethionine labeled Caskin2 SAM1-SAM2 were obtained by hanging drop vapor diffusion at 4 ˚C with equal parts of a 0.6 mM protein solution in T300 buffer and reservoir solution containing 0.1 M Tris pH 7.5, 2.4 M sodium formate, 5 mM DTT. After 24 hours, mature crystals were cryoprotected with the same crystallization solution containing 15% glycerol and flash frozen in liquid nitrogen prior to diffraction experiments. A diffraction dataset using the single anomalous dispersion method at the peak wavelength was acquired at the Canadian Light Source beam line 08B1-1 with a Rayonix MH300HE area detector (Grochulski *et al.,* 2012). All data were processed using XDS (Kabsch, Kabsch, and IUCr, 2010). The calculated Matthews coefficient

(Kantardjieff and Rupp, 2003) of 4.43 $Å^3$/Da suggested the presence of one molecule in the asymmetric unit leading to a solvent content of 72%. Phasing, density improvement, solvent flattening and refinement was performed with Phenix (Adams *et al.,* 2010). Six selenium sites were identified and an initial model was produced with AutoSol. From it, a partial model containing 113 of 166 amino acids was achieved with AutoBuild. This model was completed by successive cycles of refinement using Phenix-Refine and manual rebuilding in Coot (Emsley and Cowtan, 2004). Rigid body refinement and secondary structure restraints were applied throughout the refinement process. In the final refinement stages, target weight optimization was performed. No water molecules were added. Structural analysis was performed with MolProbity (V. B. Chen *et al.,* 2009) and PROCHECK (Laskowski *et al.,* 1993). Backbone RMSD was calculated with SSM (Krissinel and Henrick, 2004).

3.3. Results

3.3.1. The SAM Domains of Caskin2

Prior to the structural studies, sequence alignments and secondary structure predictions were performed to define the boundaries of each five-helix SAM domain. These boundaries were experimentally established through the production of pure, ^{15}N-labeled SAM1, SAM2 and SAM1-SAM2 proteins for NMR spectroscopy. At room temperature, SAM2 appeared to be folded due to the excellent dispersion and uniform resonance intensities observed in ^{1}H,^{15}N-HSQC spectra (**Figure 3.2**). SAM1, however, demonstrated the spectral characteristics of a partially unfolded protein with fewer than

expected resonances and limited chemical shift dispersion. Upon cooling the SAM1 protein to 5 ˚C and reacquiring spectra, a greater number of upfield and downfield resonances were observed suggesting that SAM1 was stabilized at low temperature. The $^{1}H,^{15}N$-HSQC spectrum of SAM1-SAM2 was not the straightforward addition of the SAM1 and SAM2 spectra suggesting the two domains were coupled. Throughout the course of these studies, we noted that SAM1-SAM2 had a strong tendency to oligomerize as evidenced by increased spectral line widths at concentrations greater than 50 μM and was affected by temperature and ionic strength.

3.3.2. Crystal Structure of The SAM1-SAM2 Tandem

Serendipitously, we observed microcrystal formation during the concentration of the Caskin2 SAM1-SAM2 tandem preparations for NMR spectroscopy at high salt concentration (0.5 M NaCl). The salt dependence on crystallization was explored by a sparse matrix screen of crystallization conditions. The structure was subsequently solved at 2.75 Å resolution from a SAD dataset acquired at the Canadian Light Source Synchrotron (**Table 3.1**). A single SAM domain tandem was observed in the asymmetric unit. From a survey of the crystal contacts, the minimal biological unit was assigned to a dimer, which then repeated as a large oligomer.

Despite having ~60% sequence identity to the Caskin1 SAM tandem, we observed a different oligomeric architecture in the crystal structure of the Caskin2 SAM tandem (Stafford *et al.,* 2011). Since each SAM domain bears a complementary head and tail surface, a tandem can interact with itself, as in the case of Caskin1, to form a tight unit

which we will call a compact monomer. The unoccupied head and tail surfaces, in turned, can link compact monomers in both directions to produce long fibrils (**Figure 3.3**). In contrast, the Caskin2 SAM tandem presented here forms a domain swapped dimer where SAM1 interacts with SAM2 of a second molecule and *vice versa*. Since each SAM domain in the dimer has an available interaction surface, the Caskin2 SAM domain oligomer has the potential to form a branched oligomer in contrast to the linear assembly observed for Caskin1 (**Figure 3.3**).

The intra-SAM domain contacts within one dimer and inter-SAM domain contacts between dimers follow a head-to-tail type interaction that has been observed in many homo- and heterotypic SAM-SAM structures including, but not limited to AIDA-1 (Kurabi *et al.,* 2009), ANKS3/ANK6 (Leettola *et al.,* 2014), Ste11/Ste50 (Kwan *et al.,* 2006), LEAFY (Sayou *et al.,* 2016), Liprin-α/Liprin-β (Wei *et al.,* 2011), Ph/Scm (Kim and Kim, 2005), Shank3 (Baron *et al.,* 2006), Ship2/EphA2 (Lee *et al.,* 2012; Leone, Cellitti, and Pellecchia, 2008), TEL (Kim *et al.,* 2001), and Yan/Mae (Qiao *et al.,* 2004).

The head interaction surface of SAM2, located on the opposite side of this small globular domain, draws contributions from helices 2, 3, and 4. The tail interaction surface of SAM1 draws contributions nearly exclusively from helix 5. A detailed view of the head and tail surfaces of Caskin2 SAM1 and SAM2 are presented in **Figure 3.4** and follow the same coloring scheme as **Figure 3.3**, for clarity.

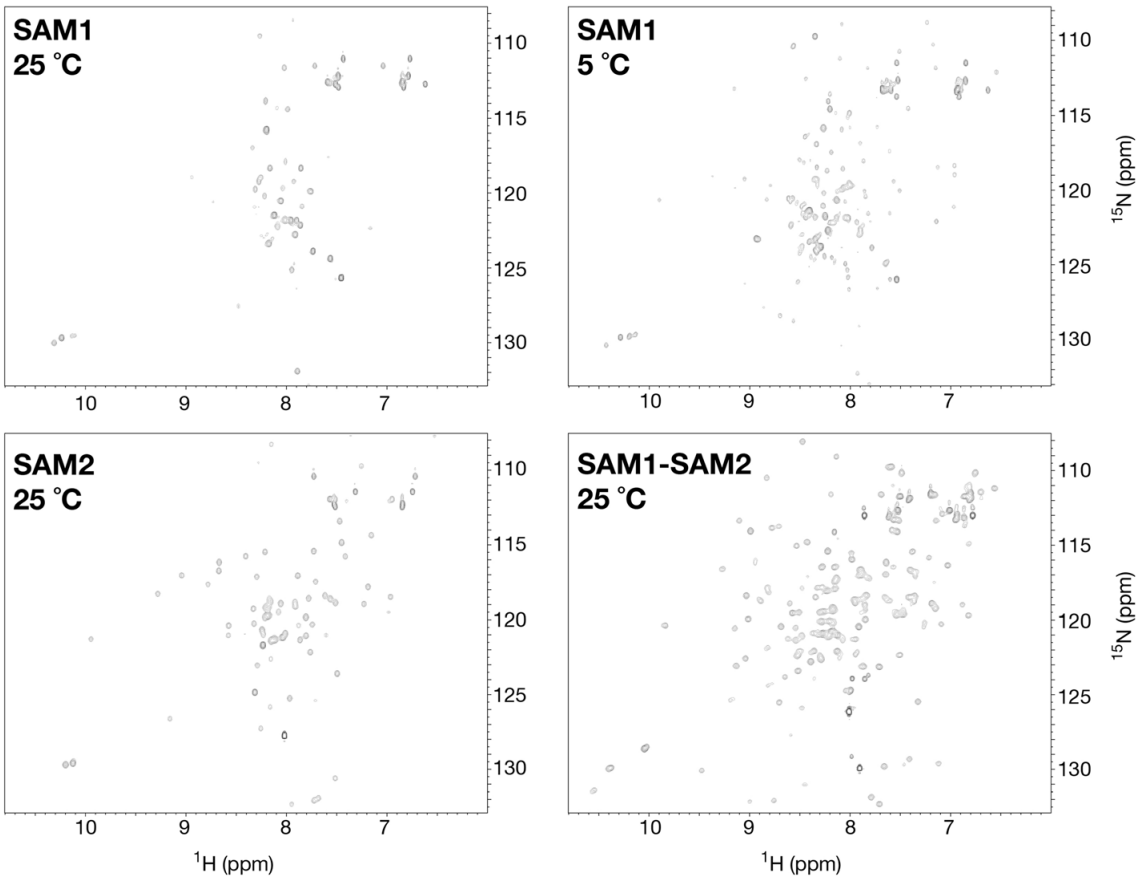


Figure 3. 2 – In isolation, Caskin2 SAM1 and SAM2 demonstrate different thermostabilities. ^{1}H-^{15}N HSQC spectra acquired at 700 MHz at a protein concentration of 100 μM in PBS buffer supplemented with 10 % D_2O. At 25 °C, SAM1 appears to be partially unfolded as the spectrum shows poor amide resonance dispersion as well as fewer resonances than expected. When the SAM1 sample is reacquired at 5 °C, more resonances are apparent. In contrast, the spectrum of SAM2 suggests that it is folded at 25 °C. The spectrum of the SAM1-SAM2 tandem is not an addition of the individual SAM1 and SAM2 spectra suggesting an interaction between the two domains.

Table 3. 1 – Data collection and refinement statistics.

Data collection	
Space group	P6522
Cell dimensions (Å)	96.4, 96.4, 119.2
Wavelength (Å)	0.97912
Reflections	377 578 (38 328)
Unique reflections	9004 (881)
Multiplicity	41.9 (43.5)
R-merge (%)	7.7 (0.86)
$< I / \sigma(I) >$	56.8 (4.68)
Completeness (%)	100.0
Refinement	
Resolution	48.51-2.75
Reflections	9004
R_{work}	0.2449
R_{free}	0.2649
Protein atoms	1100
Protein residues	140
Water molecules	0
RMSD bond lengths (Å)	0.012
RMSD angles (˚)	1.269
Ramachandran statistics	
Most favored (%)	92.14
Additional allowed (%)	7.86
Disallowed (%)	0.0

Values in parentheses correspond to the highest resolution shell (2.85-2.75 Å)

While the SAM-SAM head-to-tail interaction is predominantly hydrophobic, ionic contacts serve an important role at the intramolecular SAM-SAM interface of the dimer and the intermolecular SAM-SAM interface between dimers. Specifically, ionic contacts were observed between D527 / K610 and D516 / K611 at the intramolecular SAM1-SAM2 interface and between H538 / D585 and K540 / D592 at the intermolecular SAM1-SAM2 interface. A more extensive ionic contact network was observed in the AIDA-1

neuronal scaffolding protein SAM tandem; a consequence of a highly basic nuclear localization signal being buried at the SAM-SAM interface. Ionic contacts also help the SHIP2 SAM domain discern its bona fide EphA1 and EphA2 SAM protein partners from other closely related SAM domains such as EphB2 (Lee *et al.,* 2012).

A hydrophobic network with contributions from W554 and Y558 and peripheral support from L555, together serve to restrain the linker in one conformation in the crystal structure (**Figure 3.5**). These contacts, in turn, may limit the freedom that the two pairs of SAM domain have in solution. Near the linker, an ionic contact (E565 / R618) from the SAM domains across the dimer interface further add to the compactness of the assembly. It is worth noting that in Caskin1, Y558 is replaced by H542 and E565 is replaced by V549. Thus, both the hydrophobic and ionic contacts are not preserved in the linker and may contribute to the different types of oligomers observed. Finally, in this assessment of the linker region, we wish to emphasize that the sole conformation of the linker in the crystal structure should be interpreted with the caveats that it exhibited the highest B-factors in the refined model along with diminished electron density quality from an examination of an omit map that provides an unbiased assessment of the experimental data (**Figure 3.6**).

To test if the two Caskin2 SAM domains could bind each other independently, a ^{1}H-^{15}N HSQC reference spectrum of uniformly ^{15}N-labeled SAM2 at 100 μM was initially acquired, followed by the addition of unlabeled (^{14}N) SAM1 at a 1:1 stoichiometric ratio and reacquisition of the spectrum. From an examination of the overlaid spectra, only a

few minor peak changes were observed in stark contrast to the spectrum of the tethered SAM1-SAM2 protein presented in **Figure 3.4.** Thus, this experiment suggests that the two SAM domains must be tethered to interact with each other, with the linker potentially playing an active role in their association.

3.3.3. Mutational Analysis of The SAM Domain Interfaces

Consistent with the majority of SAM domain protein NMR and crystal structures solved to date, a head-to-tail type interaction facilitates SAM-SAM contacts within the dimer and throughout the oligomer. As a result, mutants on this surface can be designed that break one type of contact, intra- or intermolecular, while preserving the other. The substitution mutants described in this section are highlighted in **Figure 3.5.**

The tail surface is comprised of residues from the beginning of helix 5. Within helix 5, a glycine plays a critical role because the absence of a side chain at this position permits the close approach of the helix backbone to the head surface of the opposing SAM domain. In Caskin2 SAM1 and SAM2, these glycines are G537 and G607, respectively. According to the crystal structure, a substitution at G537 is predicted to preserve the dimer interface and inhibit oligomerization. Likewise, a substitution at G607 is predicted to decouple the SAM domains within the dimer leading to an open monomer similar to what was observed in the asymmetric unit of the crystal structure. Consistent with these predictions, an isotopically ^{15}N-labeled G537D mutant was more soluble than the wild type SAM tandem and demonstrated an ^{1}H-^{15}N HSQC spectrum with excellent dispersion while an isotopically labeled G607D mutant demonstrated poor solubility and was only

partially folded by a qualitative comparison of ^{1}H-^{15}N spectra with the wild type protein.

Using the same NMR survey employed for the G537D mutant, a modest increase in solubility was also observed for a K540E mutant. This substitution is located one helical turn down from the previously described G537D mutant. The combination of the two tail substitutions, expressed as a G537D/K540E double mutant, produced synergistic increase in solubility. This double mutant permitted solution NMR studies to be performed at a high protein concentration (0.8 mM, ~15 mg/mL) and temperature (37 ˚C). Furthermore, the favorable solution characteristics of the G537D/K540E double mutant made an interesting counterpoint to the wild type protein for additional *in vitro* and *in vivo* studies.

Given our success at breaking intermolecular contacts between dimers at the tail surface of SAM1, we also investigated an L589E mutant that was predicted to break contacts between dimers at the head surface of SAM2. ^{1}H-^{15}N HSQC spectra of isotopically labeled preparations of the L589E mutant presented the characteristic spectral dispersion of a folded and coupled SAM tandem but suffered from limited solubility similar to what we observed for the individual G537D and K540E mutants.

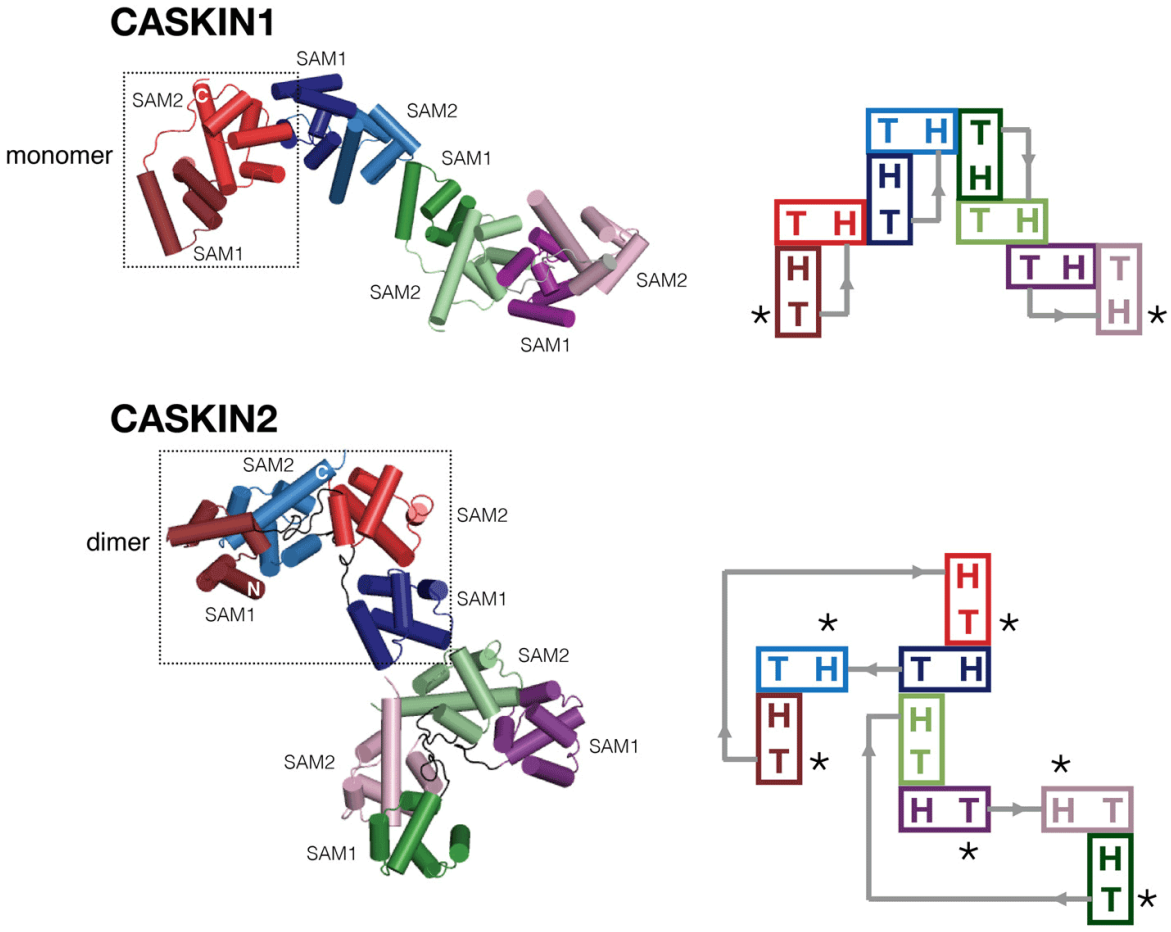


Figure 3. 3 – A comparison of the Caskin1 (PDB: 3SEI) and Caskin2 SAM domain tandem oligomers. Each SAM1-SAM2 tandem is colored individually, with SAM1 represented by a darker shade. The repeating unit is boxed. All of the intra- and intermolecular SAM domain interactions shown follow a head-to-tail type interaction. The head surface is derived from helices 2, 3 and 4 while the tail surface is predominantly derived from helix 5. To the right of each structure is a schematic illustrating the interactions between head and tail surfaces. An asterisk denotes available head and tail surfaces. Note that the Caskin1 oligomer can only grow as a fibril in both directions while the Caskin2 oligomer can form a branched structure.

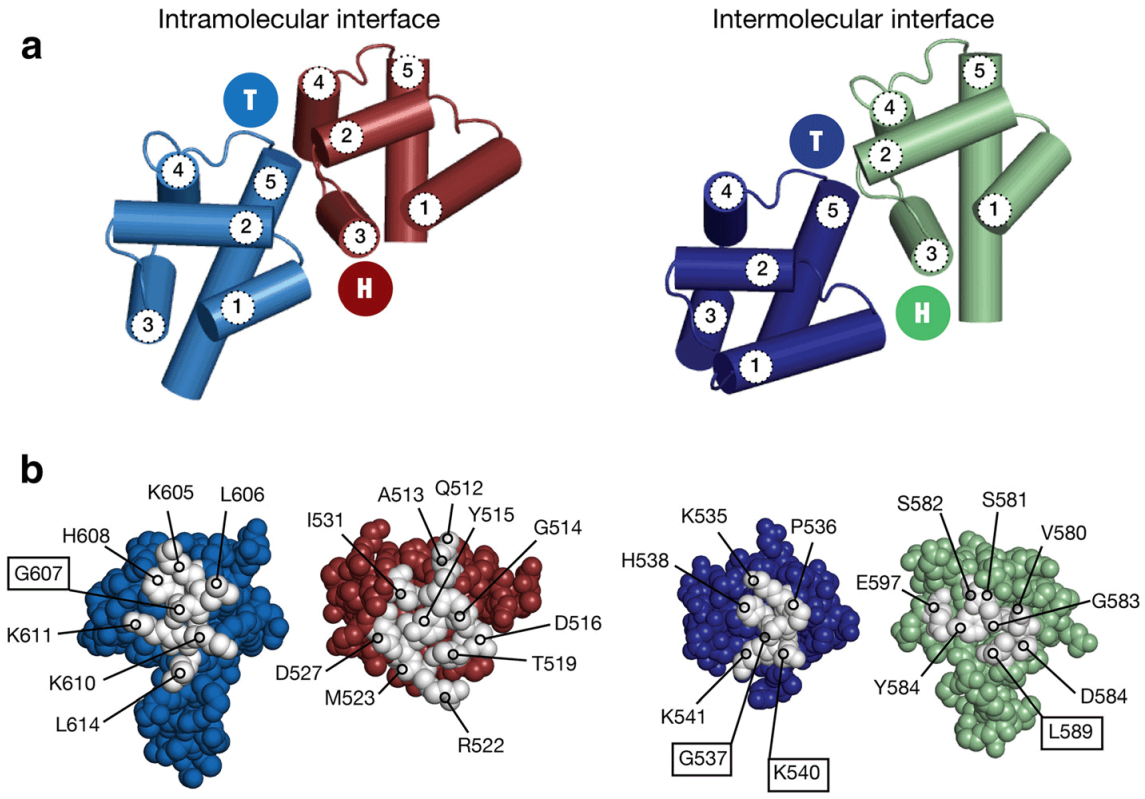


Figure 3. 4 – Detailed view of the complementary surfaces of the Caskin2 SAM tandem, following the same color scheme as **Figure 3.3. (a)** Cartoon representation of the five helices comprising each SAM domain, and the location of the complementary head (H) to tail (T) surfaces. The head surface is formed by contributions from helices 2, 3, and 4. The tail surface is formed by contributions mainly from helix 5. **(b)** Intermolecular contacts between at the intra- and intermolecular head and tail surfaces are labeled. Boxes indicated amino acids selected for mutagenesis.

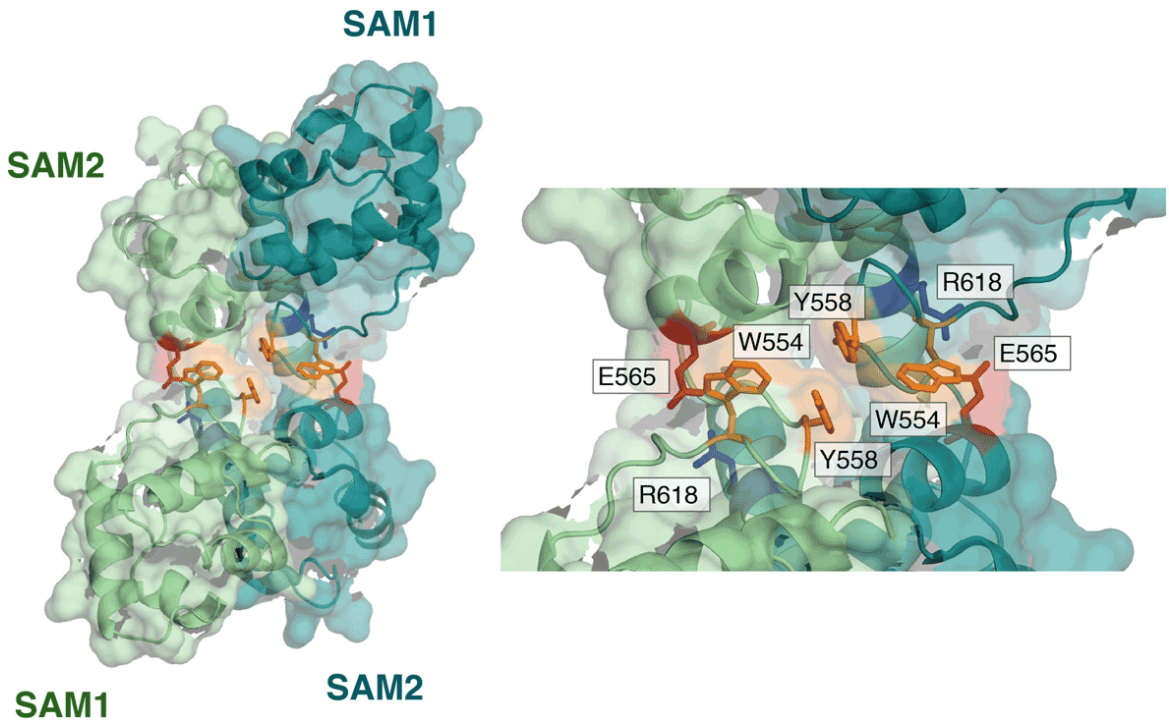


Figure 3. 5 – The linker interface in the Caskin2 SAM tandem dimer. As observed in the crystal structure of the domain swapped dimer, SAM1 and SAM2 are restrained by intra- and intermolecular hydrophobic interactions between W554 and Y558 (*orange*) in the linker region. This central interface is further defined by an ionic interaction between E565 (*red*) and R618 (*blue*).

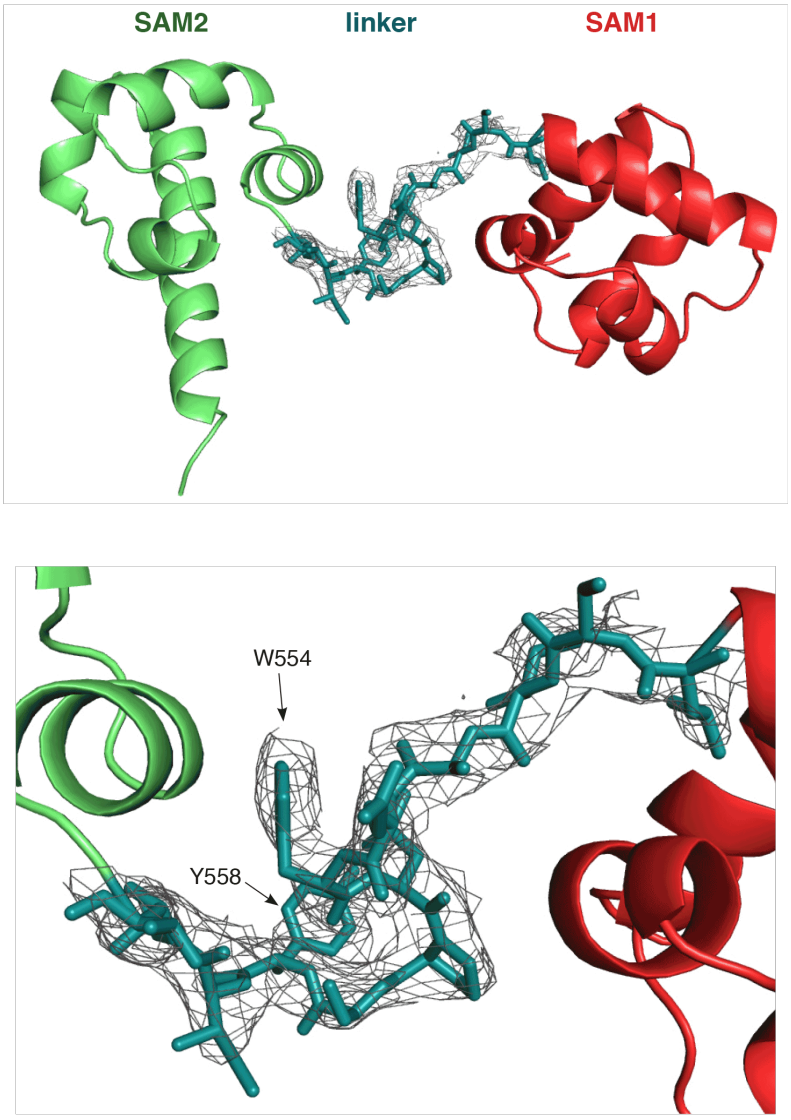


Figure 3. 6 — Omit map of the Caskin2 SAM domain tandem linker region. A composite 2mFo-DFc omit map covering the linker (549-561) was calculated with Phenix 1.10.1 (iterative removal of phase bias simulated annealing for low resolution structures). The map shown was contoured at 1.0σ. The SAM domains and the linker are colored separately for reference. Selected aromatic amino acids are labeled for reference.

3.3.4. Structural Features of the G537D/K540E Double Mutant

Since the G537D/K540E mutant was very soluble, a uniformly $^{15}N/^{13}C$ labeled sample was produced and studied by NMR methods. We had confidence that the SAM domain interactions were preserved because the ^{1}H-^{15}N HSQC spectra of the double mutant (at 0.8 mM) and wild type protein (at 0.01 mM) were superimposable **(Figure 3.7**). From a set of conventional heteronuclear experiments acquired at 950 MHz, some backbone (HN, CA, CB, C) chemical shift assignments could not be made likely due to hydrogen exchange occurring at 37 ºC and pH 7.8 (**Figure 3.8**). Six of the thirteen backbone amide resonances in the linker region between the SAM domains could not be assigned suggesting it could be experiencing motions in the intermediate (μs-ms) timescale exacerbated by hydrogen exchange. Thus the data in solution appear to suggest that the linker in the G537D/K540E mutant is more flexible in contrast to the single conformation that observed in the wild type SAM tandem crystal structure.

In the crystal structure, the C-terminal segment of helix 5 in SAM2 extends to G658 and makes contacts with a similar segment in another dimer. From the solution NMR studies of the G537D/K540E double mutant that suppresses oligomerization, helix 5 is shorter, ending instead at L652 as evidenced by the absence of strong sequential backbone HN(*i*,*i+1*) NOEs from this position onwards. Building upon the G537D/K540E framework, a Δ620 C-terminal deletion mutant was expressed and ^{15}N uniformly labeled.

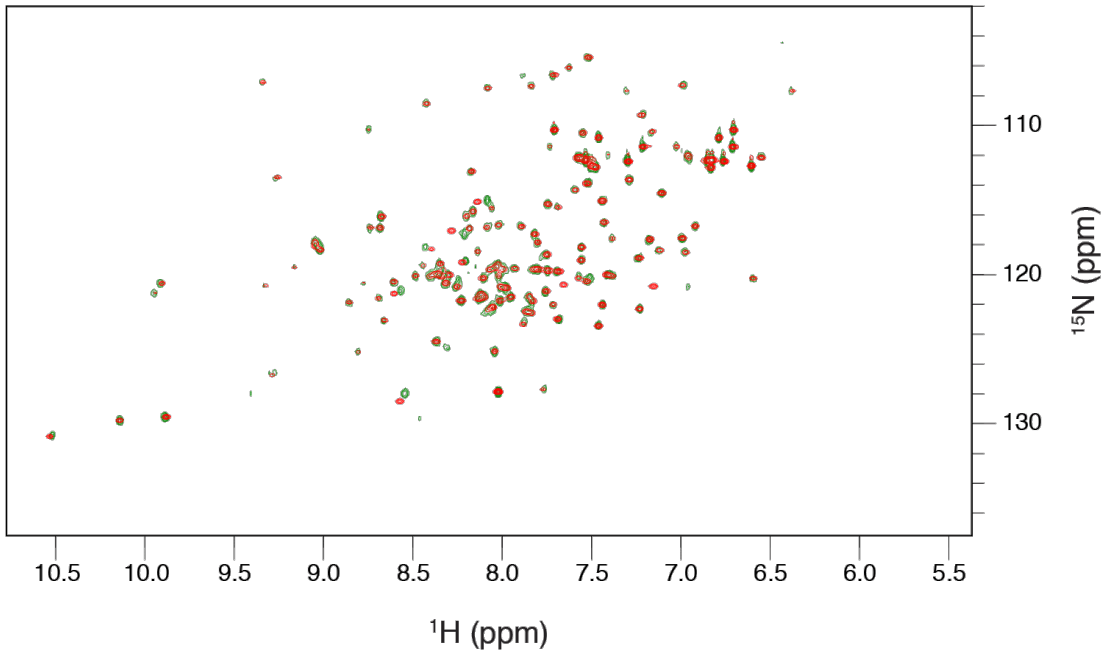


Figure 3.7 — Comparison of wild type and double mutant SAM tandem proteins by NMR spectroscopy. Overlay of ^{1}H-^{15}N HSQC spectra acquired at 298 K of the wild type protein (green; 10 µM, 20 mM sodium phosphate pH 7.8, 300 mM NaCl) and G537D/K540E double mutant (red; 800 µM, 20 mM sodium phosphate pH 7.8, 150 mM NaCl).

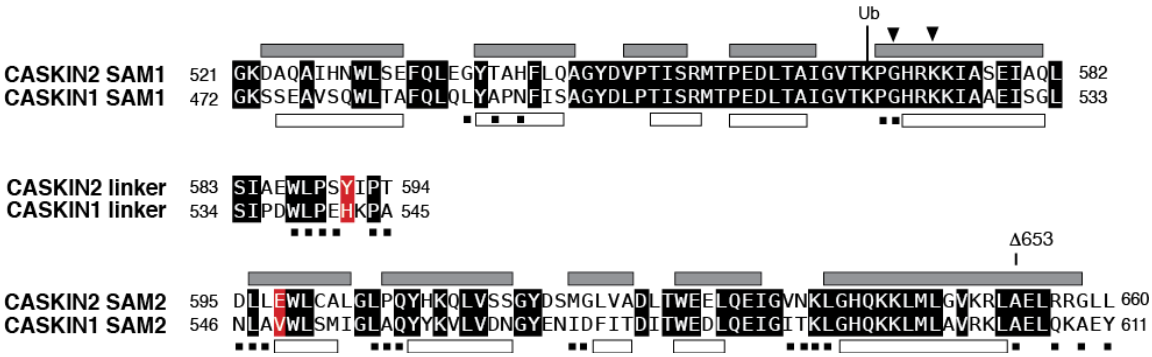


Figure 3. 8 – Secondary structure of the Caskin2 SAM tandem by NMR and X-ray methods. Closed and open rectangles denote the five helices in each SAM domain. Triangles denote two surface exposed amino acids (G537, K540) whose substitution suppressed oligomerization. *Black* squares denote amino acids that could not be assigned in 950 MHz NMR spectra of a Caskin2 G537D/K540E double mutant. Ub denotes a ubiquitin site observed from a global proteomics survey of Caskin1 (Wagner *et al.,* 2012). Δ653 identifies the site of a C-terminal truncation of the G537D/K540E double mutant to delimit the boundary of helix 5 in solution. In the sequence comparison with Caskin2, *black* boxes denote sequence similarity. Two *red* boxes denote differences between Caskin2 and Caskin1 that are predicted to reduce hydrophobic and ionic contacts at, and in the vicinity of, the linker region.

The ^{1}H-^{15}N HSQC of this deletion mutant was virtually indistinguishable from the parent G537D/K540E confirming that helix 5 is not only shorter in the oligomerization-suppressed double mutant, and that the region from residue 653 onwards does not make any significant contributions to the SAM1-SAM2 fold.

A series of ^{15}N T_1 and T_2 relaxation rate measurements were made on a ^{13}C,^{15}N labeled sample of the G537D/K540E double mutant at high concentration (0.8 mM) at 25 ºC. From the spectra, 49 non-overlapping resonances corresponding to structured regions of protein were selected for further analysis with an average ^{15}N T_1 and T_2 rates of 1.55 ± 0.15 s and 0.064 ± 0.002 s, respectively. From the T_1/T_2 ratios of each observation, an average correlation time of 13.2 ± 1.1 ns was determined. In terms of molecular weight, this correlation time corresponds to an isotropically tumbling protein of 29 ± 2 kDa. To put this observation in context, the monomeric molecular weight of the 6xHis tagged SAM1-SAM2 tandem is 20.4 kDa, and if the unstructured amino- and carboxy termini are ignored, the remaining 140 aa. contribute 15.7 kDa. Thus, the correlation time suggests that the SAM tandem in solution has characteristics of a protein assembly larger than a monomer, upwards to a dimer.

3.3.5. Monomer-Dimer Equilibria of the Wild Type SAM Tandem and an Oligomerization Suppressed Double Mutant

The oligomerization state of the wild type Caskin2 SAM tandem is affected by temperature, protein concentration and ionic strength. In our early NMR studies, a transition to the oligomer occurred at concentrations near the practical limit of the

technique (~50 μM) leading us to pursue a structure-directed G537D/K540E double mutant that was resistant to oligomerization. However, we observed differences in the linker and helix 5 of SAM2 leading us to consider the possibility that the mutations affected the equilibrium between the monomeric and dimeric states. A solution of the G537D/K540E double mutant structure was not pursued because the observed correlation time suggested that there could be two indistinguishable states — a compact monomer similar to Caskin1 crystal structure (Stafford *et al.,* 2011) and a dimer similar to the Caskin2 crystal structure presented in this study.

To complement and extend these initial observations at high concentrations, a series of analytical ultracentrifugation / sedimentation velocity (AUC-SV) experiments were performed at two low concentrations (10 μM and 34 μM) and two ionic strengths (150 and 300 mM NaCl). AUC-SV is particularly well suited for studying mass action driven reversible associations and detecting subtle changes in thermodynamic behavior.

Representative diffusion corrected sedimentation profiles shown in **Figure 3.9** clearly demonstrate that the wild type Caskin2 SAM tandem responds to mass action, while the G537D/K540E double mutant does not. In other words, as the concentration increases, oligomerization of the wild type SAM tandem increases and the diffusion corrected sedimentation distributions shift towards higher values. As we initially observed during crystallization trials, salt concentration was also observed to promote oligomerization in the analytical ultracentrifuge, with the highest protein and salt concentrations producing an additive effect. Quantitative K_d values and anisotropy information were obtained by fitting AUC-SV experiments from the 34 μM experiments

to reversibly self-associating monomer-dimer equilibrium models using a genetic algorithm (Demeler *et al.,* 2010). Only the higher concentration experiments were fitted, since these experiments cover a larger concentration range and therefore contain more signal, covering both monomer and dimer species with higher confidence. Since AUC-SV experiments produce a moving boundary which extends from zero concentration to the loading concentration (34 μM, in this case), reversibly self-associating systems will display a reaction boundary where the ratio of monomer to oligomer changes from 100% monomer near zero concentration towards increasing amounts of the oligomeric species at the higher loading concentration. Fitting the entire reacting boundary shape with finite element solutions of the Lamm equation for reacting systems (Cao and Demeler, 2008) then permits an accurate determination of the equilibrium constant. All AUC-SV experiments produced excellent fits with RMSD values comparable to the more degenerate 2DSA fits. The K_d values for all four measurements are summarized in **Table 3.2**. These results clearly show that the K_d determined for the double mutant far exceeded the loading concentration, suggesting essentially monomeric composition. The K_d of the double mutant at 300mM NaCl concentration is approximately three-fold higher than the loading concentration, indicating that even under high salt conditions there is only negligible self-association. Thus, the AUC-SV study provides convincing evidence that the oligomerization deficient mutant G537D/K540E at a 34 μM concentration and below is a compact monomer.

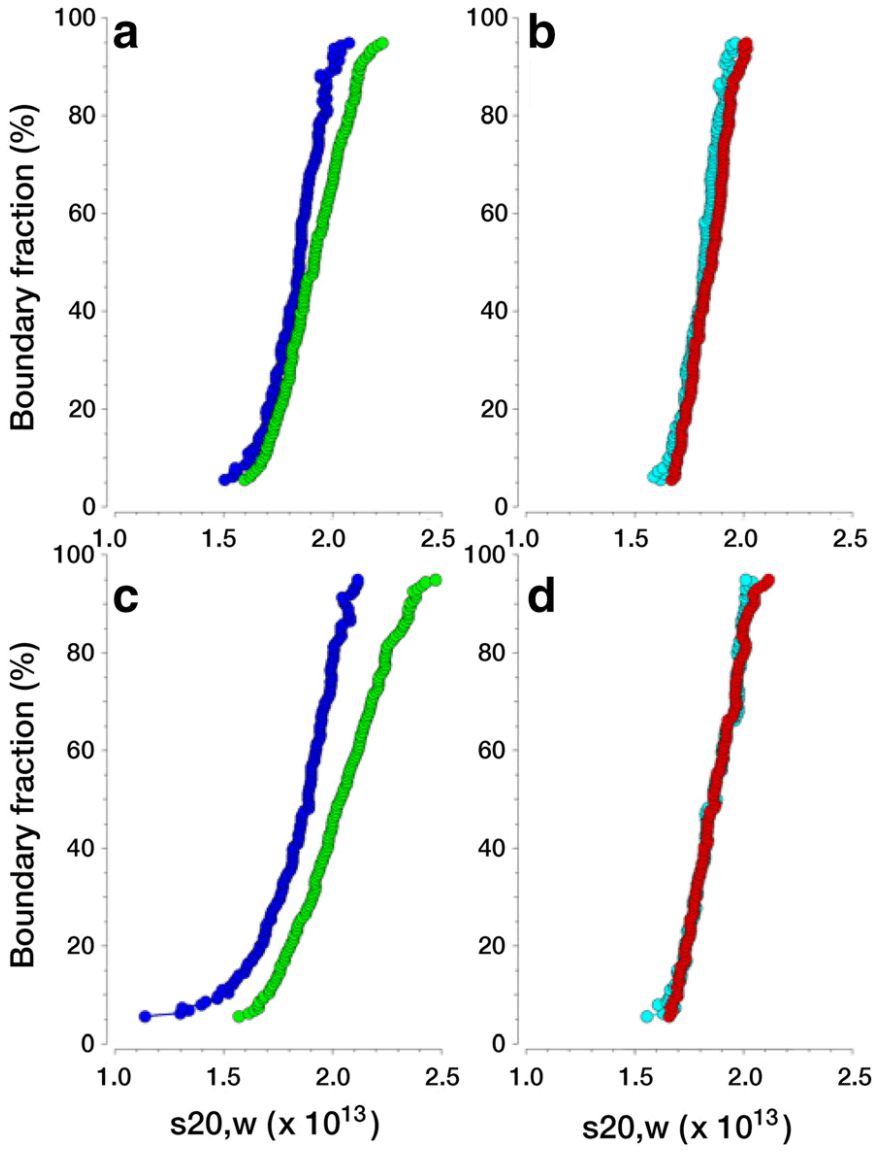


Figure 3. 9 – Van Holde - Weischet integral G(s) sedimentation coefficient distributions for Caskin2 at 10 μM (wild type, *blue*; G537D/K540E double mutant, *cyan*) and 34 μM (wild type, *green*; G537D/K540E double mutant, *red*) loading concentrations. Panels **(a)** and **(b)** were measured at 150 mM NaCl, while panels **(c)** and **(d)** were measured at 300 mM NaCl. A shift in sedimentation coefficient for higher concentrations indicates reversible mass action. This effect is only seen for the wild type, not for the double mutant. Furthermore, the effect is enhanced at higher ionic strength, indicative of a decrease in K_d for the wild type. These results indicate that the double mutant only exists in a monomeric form at low concentration, while the wild type SAM tandem dimerizes and is more sensitive to changes in ionic strength.

Table 3. 2 – Monomer-dimer equilibrium constants for wild type Caskin2 SAM1-SAM2 and an oligomerization-inhibited double (G537D/K540E) at two NaCl concentrations.

	150 mM NaCl	300 mM NaCl
Wild type		
K_d (μM)	52.9 (49.3, 56.5)	27.6 (26.8, 28.3)
φ (monomer)	1.25	1.27
φ (dimer)	1.28	1.25
G537D/K540E		
K_d (μM)	n.d.	99.8 (98.9,100.7)
φ (monomer)	1.14	1.02

Values in parentheses represent the 95 % confidence intervals obtained from a genetic-algorithm Monte Carlo analysis. φ represents the anisotropy of the molecule, with a value close to 1.0 indicating a more compact and globular structure, while increasingly larger values reflect more extended shapes. Since the dimer concentration of the G537D/K540E mutant is negligible, the anisotropy of the dimer was not calculated. A K_d for the G537D/K540E mutant at 150 mM NaCl could not be detected (n.d.) because the sample was essential monomeric.

A comparison of the anisotropy values from the AUC-SV analysis indicate that the monomeric and dimeric forms of the wild type SAM tandem are similarly compact (**Table 3.2**). The anisotropy values also indicated that G537D/K540E double mutant monomer is slightly more compact than the wild type SAM tandem reinforcing the observations from the NMR investigation that the double mutant and wild type SAM tandems have structural differences.

3.3.6. Expression of the Caskin1 and Caskin2 SAM Domain Tandems in Neuro2a Cells

To begin understanding how oligomerization may affect cellular processes, we transfected EGFP fusions of wild type Caskin2 SAM tandem (EGFP-WT) and the non-oligomerizing G537D/K540E mutant (EGFP-G537D/K540E) into Neuro2a cells. We chose to express only the SAM tandems to visualize their effect independently from the other protein interaction domains (ankyrin and SH3) in the amino terminal region and other unknown interaction motifs in the carboxy terminal region of Caskin2. From a series of micrographs analyzed, one representative set is shown in **Figure 3.10a**. Both EGFP-WT and EGFP-G537D/K540E were observed throughout the cell, including the nucleus. Nuclear localization by diffusion is possible since the molecular weight of the EGFP-Caskin2 SAM tandem is ~50 kDa. While the fluorescence distribution was relatively uniform for the G537D/K540E mutant, fluorescence was concentrated in dense punctae for the wild type protein. The same expression assay under similar conditions was performed with the EGFP-tagged Caskin1 SAM tandem and an analogous double mutant G520D/K523E to the Caskin2 G537D/K540E double mutant features in this study (**Figure 3.10b**). While we did not perform an *in vitro* study to confirm that the Caskin1 mutant was oligomerization-suppressed, it is worth noting that a Caskin1 G520E single mutant described in (Stafford *et al.,* 2011) was sufficient on its own.

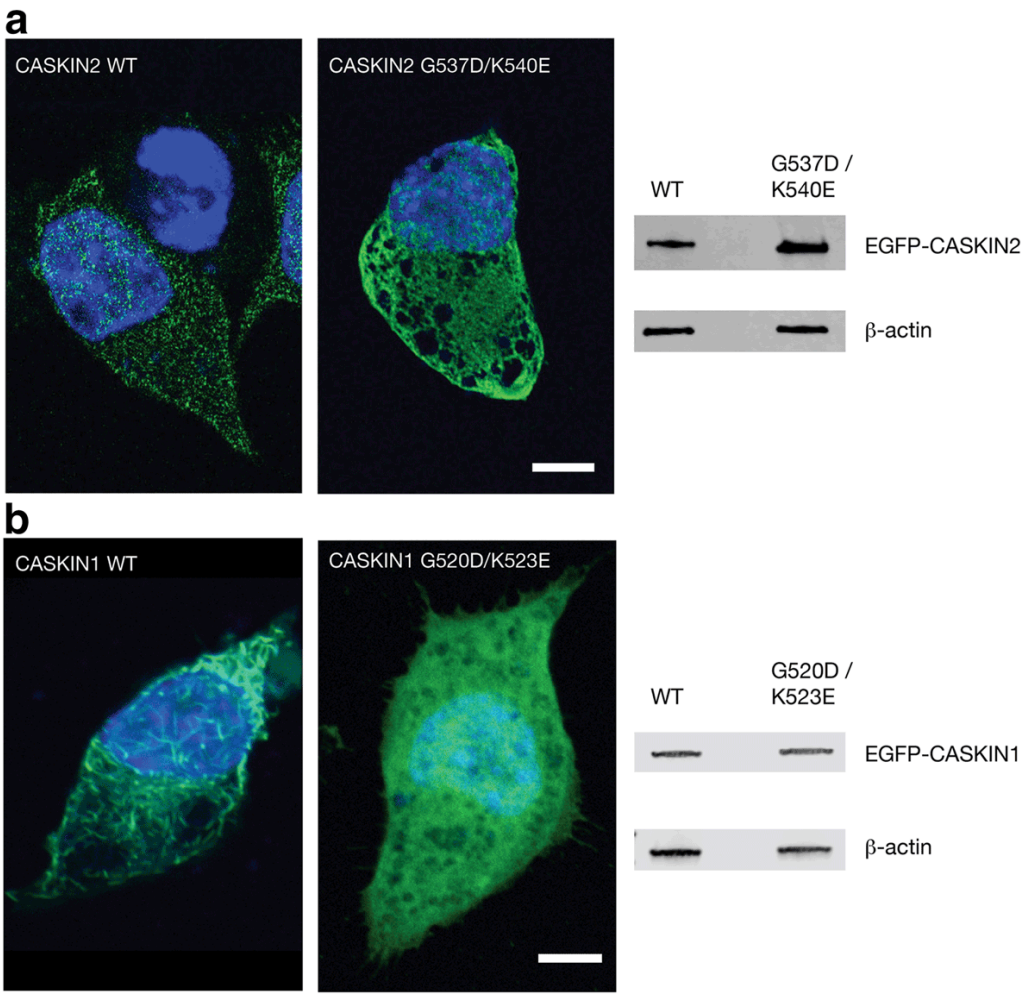


Figure 3. 10 – Caskin2 and Caskin1 SAM domain expression in Neuro2a cells. Images were made 48 h after transient transfection with **(a)** Caskin2-EGFP (wild type) and mutant Caskin2 (G537D/K540E)-EGFP **(b)** Caskin1-EGFP (wild type) and mutant Caskin1 (G520D/K523E)-EGFP plasmids. The *green* fluorescence demonstrates distinct protein distributions for the wild type and mutant proteins. Counterstaining with DAPI (*blue*) reveals that the subcellular distribution of wild type and mutant proteins in the cytoplasm and nucleus is indistinguishable. Scale bar: 5 μm. Western blots of cell lysates demonstrating expression of EGFP-Caskin2 and EGFP-Caskin1 proteins probed with monoclonal anti-EGFP antibodies presented on the left side of the panels. The blots were re-probed with monoclonal anti-β-actin antibodies as a loading control.

Consistent with our observations for the Caskin2 SAM tandem, the Caskin1 SAM tandem double mutant was distributed throughout the cytoplasm and nucleus, and the wild type Caskin1 SAM tandem formed punctae. The punctae, however, were distinct from the Caskin2 SAM tandem, appearing not as condensed speckles, but rod-like structures throughout the cell. In summary, this *in vivo* expression study is consistent with our observations from crystallography — the Caskin2 and Caskin1 SAM domains self-associate differently and consequently present a different oligomeric architecture. The differences in the morphology of the aggregates cannot be explained by variances in concentration since both SAM tandems were expressed at approximately the same levels as an actin control.

3.4. Discussion

We have presented data from a set of complimentary sources (X-ray crystallography, NMR spectroscopy, analytical ultracentrifugation, and *in vivo* expression) demonstrating that the Caskin2 SAM tandem experiences concentration and salt dependent oligomerization. While the Caskin2 SAM domain crystal structure presents a series of head-to-tail contacts that are typical for most self-associating SAM domains, the manner in which the oligomer is organized as a repeating set dimers is new and distinct from Caskin1.

Analysis of the AUC-SV data suggests that the wild type Caskin2 SAM tandem is in a reversible monomer-dimer equilibrium at low concentrations (10-34 μM). In contrast,

the oligomerization suppressed G537D/K540E double mutant is essentially monomeric with the dimeric form only beginning to become apparent at high (>500 μM) concentrations. This difference between the wild type and double mutant proteins is qualitatively apparent in the magnitude of the shifts and shapes in sedimentation distributions. From these data and prior knowledge of the system, a quantitative approach using discrete reversible monomer-dimer equilibrium models were justified to determine a K_d of the wild type SAM tandem from the SV data directly at two ionic strengths. Consistent with the crystallization conditions, ionic strength enhanced dimerization for the wild type SAM tandem and to a much lesser extent for the G537D/K540E double mutant.

The K_d of the Caskin2 SAM tandem is well suited to the anticipated levels of the protein at synaptic sites and is within the realm of other signaling domains such as SH3 and WW domain that must rapidly sample ligands to fulfill their biological functions. At low concentrations, Caskin2 in its monomeric or dimeric form could serve as a classical adaptor bringing protein partners together (**Figure 3.11**). Furthermore, dimeric Caskin2 may help activate associated proteins that depend upon dimerization. At higher concentrations, oligomeric Caskin2 could provide the increased avidity to concentrate and amplify low affinity protein partnerships that would otherwise be suppressed.

Along with concentration, ionic strength can contribute to oligomerization, although it is unclear if cation fluxes associated with neuronal signaling are sufficient to serve a regulatory role. Pursuing this idea, we did observe a series of hydrophobic contacts

in the central portion of the linker region supported by ionic contacts from the nearby SAM domains. In an analogous interaction mode to Caskin2 domain-swapped dimer, hydrophobic interactions dominate in Byr2-SAM / Ste4-SAM (Ramachander *et al.,* 2002; Ramachander and Bowie, 2004) and Ste11 / Ste50 (Kwan *et al.,* 2006) heterodimer interface with peripheral support from charged / polar residues. Likewise, Ph (Kim *et al.,* 2002), TEL (Kim *et al.,* 2001) and Yan (Qiao *et al.,* 2004) homo-oligomers and Ph / Scm (Kim and Kim, 2005) hetero-oligomers assemble around a central hydrophobic cluster in the central head-tail interface supported by number of peripheral electrostatic interactions. An examination of the Caskin2 and Caskin1 sequences suggests that these contacts would not be preserved thereby providing a possible explanation for why the minimal repeating unit of Caskin1 is a compact monomer while the minimal repeating unit of Caskin2 is a dimer. Furthermore, since ionic contacts are involved, salt concentration and pH may also serve a role at mediating Caskin2 dimerization and oligomerization. Indeed, as salt concentration increases, the K_d describing the monomer-dimer equilibrium increases. The precise effects of salt concentration could be complex as charges are screened and hydrophobic effects begin to predominate.

The high concentration of sodium formate used to promote crystallization represents the extreme effect where the protein is essentially salted out of solution. While our investigation was limited to only two NaCl concentrations (0.15 and 0.3 M) and one pH, we refer the reader to a survey of the EphA2 / SHIP2 SAM domain heterodimer for a comprehensive perspective of ionic interactions using NMR methods and molecular modeling (Lee *et al.,* 2012).

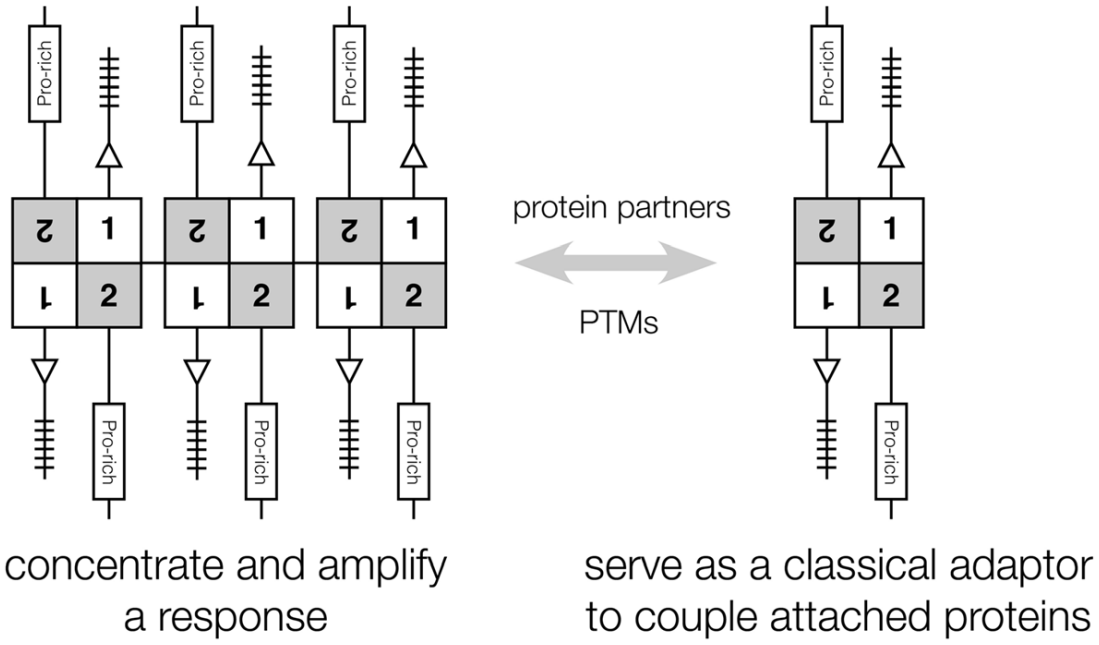


Figure 3. 11 – Signaling consequences of dimerization and oligomerization by the tandem SAM domains of Caskin2. In its oligomeric form, Caskin2 would provide a vast array of interaction sites with enhanced avidity for many proteins through its available intrinsically unstructured regions and ankyrin repeats. In its dimeric form, Caskin2 could fulfill a classic adaptor role bringing together protein partners that potentially depend on dimerization themselves for coupling and activation.

EGFP-tagged Caskin2 SAM1-SAM2 protein was observed as punctae when expressed in Neuro2a cells. This distinctive pattern is very similar to what has been reported for the oligomeric form of the Dishevelled DIX domain (Fiedler *et al.,* 2011; Schwarz-Romond *et al.,* 2007). Although they were less apparent, punctae were also observed in micrographs of GFP-tagged diacylglyercol kinase d1 (DGKδ1), facilitated by the Zn(II) dependent oligomerization of its single SAM domain (Harada *et al.,* 2008; Knight *et al.,* 2010). Mutations in the DGKδ1SAM domain that either abolished Zn(II) binding or inhibited oligomerization resulted in disappearance of punctae and translocation of DGKδ1 to the plasma membrane. Supplementing the structural study of the Caskin1 SAM tandem, transfections of GFP-tagged full length Caskin1 were performed in HEK293 cells with the majority of fluorescence observed in the cytoplasm along with some higher intensity speckles near the nucleus (Stafford *et al.,* 2011). To enable a consistent comparison with the results presented in this study, the Caskin1 SAM domain tandem and oligomerization inhibited double mutant were expressed in Neuro2a cells under similar conditions (vector, fluorescent reporter, protein levels) as the Caskin2 SAM tandem. As shown in the micrographs, there was a striking difference in the morphology of the aggregates. Taken together with the crystal structures, the Caskin2 and Caskin1 SAM domains appear to oligomerize differently *in vitro* and *in vivo*. The biological consequences of this difference may reflect the divergent roles that each protein plays in the neuron.

If Caskin2 oligomerization is an essential aspect of its neuronal signaling function, it stands to reason that there should be ways to regulate oligomerization that supersede solution conditions such as protein concentration, pH, and divalent ion concentration (Gundelfinger *et al.,* 2006; Knight *et al.,* 2010). Post-translational modifications and protein partner binding (Qiao *et al.,* 2004) offer targeted opportunities to affect the oligomerization process, by repressing the formation of oligomers or by facilitating the disassembly of oligomers. While no biological process that regulates Caskin2 has been identified to date, a global mass spectrometry survey identified a ubiquitinated lysine (K536) in Caskin1 at the same oligomerization surface as the G537D/K540E mutants described in this study (Wagner *et al.,* 2012). Ubiquitination may possibly block the formation of oligomers in an analogous manner that has been reported for the *Dishevelled* DIX domain (Madrzak *et al.,* 2015) and incorporate this neuronal signaling scaffolding protein into other signaling and translation pathways in the neuron.

In conclusion, sterile alpha motif (SAM) domains are versatile protein-protein interaction modules. Using the Caskin2 scaffolding protein as a focus of this investigation, we have demonstrated that its SAM domain tandem is able to sample monomeric, dimeric, and oligomeric states. Given the structural distinctiveness of these states, Caskin2 has the potential to support many different functions in neuronal signaling circuits.

CHAPTER 4:

CASKIN2 PARTNERSHIP WITH THE LEUKOCYTE COMMON ANTIGEN RELATED PROTEIN TYROSINE PHOSPHATASE RECEPTOR (LAR)

4.1. Introduction

Neurons generate new synapses through complex neuronal adhesion events that involve the assembly of receptors, channels, pre- and postsynaptic signaling scaffolds. Stabilization and maturation of newly created synaptic sites greatly rely on the cell adhesion molecules (Um & Ko, 2013). In addition to serving as a platform for neuronal connections, synaptic adhesion molecules are key participants in the development of long-term potentiation and synaptic plasticity (Um & Ko, 2013). Two studies functionally link the LAR receptor tyrosine phosphatase protein family to neuronal system development in both vertebrates and invertebrates (Chagnon, Uetani, & Tremblay, 2004; Xu & Fisher, 2012). Most recently, their role as major synaptic adhesion molecules has been proposed (Han, Jeon, Um, & Ko, 2016; Um & Ko, 2013). Identification and characterization of intra- and extracellular ligands of LAR PTPs is crucial for understanding the outcomes of distinct signaling pathways and the connections between them.

The LAR PTP family includes three vertebrate (LAR, PTPδ and PTPσ) and two fly homologs (Dlar and DPTP69D). Weng *et al*. (2011) reported that *Drosophila* Csk directly interacts with Dlar via an N-terminal SAM domain and connected this partnership to a motor axon pathfinding. The same group had demonstrated via a yeast two-hybrid assay that both the tandem SAM domain of mouse Caskin ortholog as well as the full-length protein preferentially interacts with two murine LAR receptor family members, LAR and PTPσ, but not with PTPδ. Other multidomain proteins of Liprin family have been reported to interact with the D2 domain of LAR through their triple-SAM domain module (Astigarraga *et al*., 2010; Serra-Pagès *et al*., 1995; Serra-Pagès, Medley, Tang, Hart, & Streuli, 1998; Spangler & Hoogenraad, 2007; Stryker & Johnson, 2007). Liprin/LAR complexes have been functionally linked to presynaptic active zone organization and synaptic maturation in *C. elegans* and *Drosophila* (Dai *et al*., 2006; Patel *et al*., 2006; Spangler & Hoogenraad, 2007; Taru & Jin, 2011). The fact that Liprin-α and Ckn cannot simultaneously bind LAR (Weng *et al*., 2011a) suggests that they might target a common or overlapping binding site on LAR. The competitive nature of Caskin/Liprin:LAR interactions suggests the possibility of very distinct signaling outcomes of these complexes.

The research presented here demonstrates the evidence that the interaction is preserved between human orthologs of LAR and Caskin2. Experimental evidence (Spot blot analysis, *in vitro* co-immunoprecipitation) suggests that it takes place at the second, catalytically inactive, domain of LAR and the SAM2 domain of Caskin2. The interaction

between wild type Caskin2 SAM1 SAM2 and LARD1D2 was further characterized by fluorescence anisotropy methods establishing a stoichiometric ratio (1:1) with a dissociation constant of 1.1 ± 0.4 μM (n = 4). Following *in vitro* studies, proteins were co-expressed in neuroblastoma cells (Neuro2a) and visualized by confocal microscopy demonstrated strong co-localization between EGFP-Caskin2 SAM1 SAM2 and dsRED-LARD1D2/D2 fluorescent fusion constructs.

4.2. Materials and Experimental Procedures

4.2.1. Peptide array synthesis and experimental procedures

An array of 12-mer peptides was synthesized and anchored on the cellulose membrane as previously described in the SPOT method (Frank, 2002) with an Intavis MultiPep synthesizer. As a synthesis quality control check of synthesis, a crude estimate of the peptide content in each spot was made by staining with Fast Green FCF dye. The peptide array membrane was washed three times in phosphate-buffered saline with Tween-20 buffer (PBS/T buffer composition: 3.2 mM Na_2HPO4, 0.5 mM KH_2PO_4, 1.3 mM KCl, 135 mM NaCl, and 0.1% Tween 20, pH 7.4) prior to blocking with Blocking Buffer (5% milk, 2.5% BSA in PBS/T) at 4 ˚C overnight. The membrane was probed with 1μM of protein of interest in PBS (1 hr incubation at RT on the orbital shaker) and followed by removal of unbound protein by three PBS/T washing steps and pre-blocking with Blocking Buffer before the final 1 hour incubation with horseradish peroxidase (HRP)-conjugated antibody diluted to 1:5000 in Blocking Buffer. The membrane was washed five times in PBS/T before the development with ECL chemiluminescence reagents (ECL

Western blotting analysis system, Amersham) and imaging on the X-ray film. In cases where membrane was re-probed, it was regenerated with 6M guanidinium hydrochloride: 0.5M imidazole solution.

4.2.2. Co-IP and immunoblotting

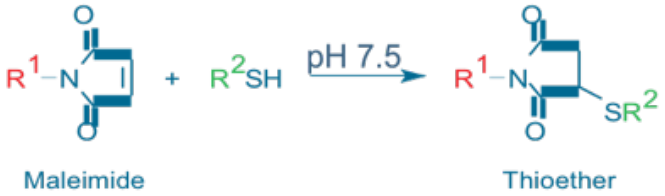
Immunoprecipitation experiments with FLAG-H_6-LARD1D2 and His_6-Caskin SAM1 SAM2 and GST-tagged single SAM domain constructs were performed in two modes using GST sepharose beads (Clontech) and anti-FLAG-Ab-conjugated magnetic beads (Sigma-Aldrich). The batch binding technique was generally used for both types of pull-downs with the difference that the centrifugation was used in the case of GST beads and a magnetic separator for anti-FLAG beads for separation steps. Twenty microliters of packed gel volume of anti-FLAG M2 (equivalent to ~40μL of 50% bead suspension) and 250µL of GST bead suspension per reaction according to manufacturer's recommended protocols were used. Prior to the protein binding step, the resin was washed three times with 10x packed gel volumes of TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7.4) buffer. After removing the buffer, the appropriate protein was added to the beads and allowed to bind by inverting for 20 minutes at room temperature. Beads were separated from supernatants (gentle centrifugation at 3000 rpm or in magnetic separator) following by three washes with TBS buffer. The appropriate presumptive interacting protein was added to the beads (either purified or freshly made pre-cleared protein lysate) and allowed to bind by inverting for 1hr at RT. The supernatants were removed and the beads were washed with TBS buffer at least three times. After removing the buffer, protein complexes

were eluted using 150-200µl of glutathione buffer (pH 7.8) or 2xSDS loading dye from GST and anti-FLAG resin respectively. Samples were separated on a 4-20% SDS-PAGE gel (BioRad) and transferred to PVDF or nitrocellulose membrane using the Trans-Blot Turbo transfer system (BioRad) and the Trans-Blot Turbo mini transfer packs (BioRad) for immunodetection. The membrane was pre-blocked overnight at 4 ºC in Blocking Buffer (5% milk, 2.5% BSA in PBS/T) to minimize the nonspecific signal background. Primary antibodies, Anti-H_6-rabbit (Santa Cruz) and anti-GST-mouse (Sigma-Aldrich), were used at the 1:4000 dilution, followed by probing with secondary antibodies for detection; goat-anti-mouse IRDye800CW and donkey-anti-rabbit IRDye680LT (LI-COR Biosciences) at 1:20000 dilutions. Signals for IRDyes were detected using the Odyssey Infrared Imaging System (LI-COR Biosciences).

4.2.3. Fluorescence anisotropy titrations

Labeling method: A maleimide dye conjugation reaction was performed in accordance with the Invitrogen recommended protocol "Thiol-Reactive Probes" in PBS buffer (pH 7.5) The stock solution of the maleimide dye BODIPY® FL (*N*-(2-Aminoethyl Maleimide, B10250) (Thermo Fisher Scientific) was freshly prepared before labeling at a concentration of 0.5mg/μL in anhydrous DMSO. Purified Caskin SAM2/SAM1-SAM2 protein was prepared in phosphate buffered saline (PBS) buffer: 20 mM sodium phosphate, pH 7.5, 0.15 M NaCl. At pH 7.5 the cysteine free thiol (sulfhydryl) group readily reacts with the maleimide dye forming a stable thioether bond.

Labeling chemistry reaction:



The dye was added to the protein at 20-fold molar excess and incubated for at least 2 hours at RT protected from light. The reaction was quenched with an excess of glutathione and protein was dialyzed overnight at 4 ºC into the assay buffer (PBS, pH 7.8) to eliminate the excess of unbound dye and glutathione. The protein conjugate was stored at 4 ºC protected from light.

Fluorescence titration procedure: Labeled 6xHis-Caskin SAM SAM and SAM2 (at 0.5 μM) and unlabeled 6xHis-LARD1D2 at a concentration range of 38 nM to 19 μM were used for fluorescence anisotropy binding studies at 25 ºC using an Agilent Eclipse spectrophotometer equipped with a manual polarizer accessory. Excitation/emission signals were measured at 496/516 nm wavelengths with an averaging time of 0.25 sec with three reading replicates for each sample. Anisotropy was calculated from the relationship $(I_{parallel}–GI_{perp})/(I_{parallel}/2GI_{perp})$ and normalized with the blank experiment and to the maximum fluorescence change (F/F_{max}). The final plot was generated using the Prism 7 graphing program (GraphPad Software Inc.). The equilibrium dissociation constant (K_d) was calculated by fitting the averaged data set (n=4) directly using nonlinear least squares analysis plugin within the Prism 7 program.

4.2.4. Cloning and site-directed mutagenesis, protein expression and purification

EGFP and dsRed constructs were sub-cloned from Caskin2 SAM1-SAM2 tandem in pET28 (483-634; Uniprot Q8WXE0) (Novagen) and LARD1D2 in pD441-NF (DNA 2.0). The fusion proteins were prepared by inserting a suitable PCR product into the XhoI/KpnI restriction sites of the pEGFP-N1 and pDsRed-monomer (Clontech) expression vectors. Followed by the transformation into DH5α cells, amplification and DNA extraction was performed for further use in mammalian cells. The C-terminal deletion ΔCSS[G70D/K73E] in pET28b plasmid was generated by site-directed mutagenesis from the Caskin2 SAM1-SAM2 G537D/K540E (pET28b) clone using the Quikchange Lightning kit (Agilent Technologies). All plasmid constructs were sequence verified (TCAG DNA Sequencing Facility, Hospital for Sick Children, Toronto). pET28b plasmids were transformed into *E. coli* BL21:DE3 for 6xHis-tagged protein expression. Transformed *E. coli* BL21:DE3 were grown at 37 ºC in LB (or M9 minimal media for NMR samples) until the exponential phase (OD_{600} of 0.7-0.8). Protein expression was induced with the addition of 1 mM IPTG and further incubation for another 3 hours at 37 ºC. In the case of His_6-FLAG-LARD1D2, the highest levels of expression were achieved in Arctic express DE3 cells with an overnight induction at 13 ºC. Cell pellets were dissolved in T300 buffer (20 mM Tris-HCl, 300 mM NaCl, 0.05 % NaN_3) and lysed by French press. Highly purified protein was obtained from a two-step purification involving Nickel-NTA affinity chromatography (Qiagen), followed by gel filtration chromatography (Sephacryl-100, HiLoad 16/60; GE Life Sciences). GST-tagged proteins

were purified using a glutathione sepharose column (Sigma-Aldrich) with elution buffer containing 10mM glutathione, pH 7.8 in T300 buffer. All bacterial expression constructs included a thrombin protease cleavage sequence located N-terminally of the protein sequence. Proteins were liberated from the GST tag when required depending on the experiment, reverse purified and dialyzed into the final buffer. Protein purity was assessed by SDS-PAGE chromatography. The final buffer for NMR analyses was phosphate buffered saline (PBS): 20 mM sodium phosphate, pH 7.8, 0.15 M NaCl, 0.05 % (w/v) NaN_3.

4.2.5. Cell culture, transient transfection, coexpression and immunoblotting

Neuroblastoma 2a (Neuro2a) cells (Olmsted *et al*., 1970) were maintained using standard growth conditions and used for expression and localization studies as described in (Prochnow, Hoffmann, Dermietzel, & Zoidl, 2009). Cells were seeded in 24-well plates or glass-bottom dishes (MatTek Corporation, Ashland, MA, USA) and transfected with 200 ng (single transfection) or 400ng (double transfection) endotoxin-free plasmid DNA, using the Effectene transfection protocol (Qiagen Inc., Valencia, CA, USA). Whole cell protein lysates from transfected Neuro2a cells collected 48 h post-transfection were separated by 10 % SDS-PAGE and transferred to 0.2 μm Hybond-ECL nitrocellulose membrane (GE Life Sciences) for immunodetection. Primary antibodies were diluted 1:1000 (rabbit anti-GFP; Santa Cruz) and 1:20000 (mouse anti-β-actin; Sigma-Aldrich). Secondary antibodies (LI-COR Biosciences) were diluted 1:20000 (donkey anti-rabbit IRDye680LT) or 1:20000 (goat anti- mouse IRDye800CW). Signals were detected using

the Odyssey Infrared Imaging System (LI-COR Biosciences).

4.2.6. Confocal microscopy

Transfected cells were fixed with 4% paraformaldehyde for 20 min at room temperature, washed with PBS, counterstained with DAPI and mounted for imaging. Samples were visualized using a Zeiss LSM 700 confocal microscope with a Plan-Apochromat 63x/1.4 Oil DIC M27 objective and the ZEN 2010 program to control all hardware parameters. Images were collected by line averaging (4x) at high resolution (2048x2048 pixel) using single planes or z-stacks. Mander's overlap coefficients were calculated using the built-in co-localization analysis tools in the ZEN 2010 program (Manders & Tyberghein, 1993). Ten random cells from each co-transfection data set were used for overlap coefficient determinations. Images were combined in Adobe Illustrator CC for presentation.

4.3. Results

4.3.1. A search for putative interaction surfaces on Caskin SAM-SAM and LARD2

To identify the interaction surfaces a synthetic peptide blotting method (SPOT blot) was used. A 12-mer peptide array covering the Caskin2 tandem SAM domain amino acid sequence with a three amino acid shift window was probed for interaction with LARD2. Two surface exposed regions on the second SAM domain were detected (**Figure 4.1a**). The first consensus sequence (103)L-G-L-P-Q-Y-H-K(110) highlighted in green (**Figure 4.1b, c)** covers the end of helix 1 and the portion of helix 2 plus the short loop between them, with a basic Lys110 protruding into solution. The second detected

sequence (150)R-L-A-E-L-R-R-G(157), highlighted in a darker orange shade, comprises the core of the helix 5 within SAM2 (**Figure 4.1b, c**). Note that helix 5 is not involved in the dimerization interface and is solvent exposed in either form of Caskin, dimeric or monomeric. In addition to cMyc-His_6-LARD2 protein, the membrane was also tested with GST-LARD2 and GST-LARD1D2 protein constructs revealing the same peptide regions (data not shown). The complementary peptide array was produced with two amino acid sliding sequence window covering the entire sequence of LAR D2 domain (**Figure 4.2**), based on the *Drosophila* ortholog sequence (GenBank ID: AAC47002.1).

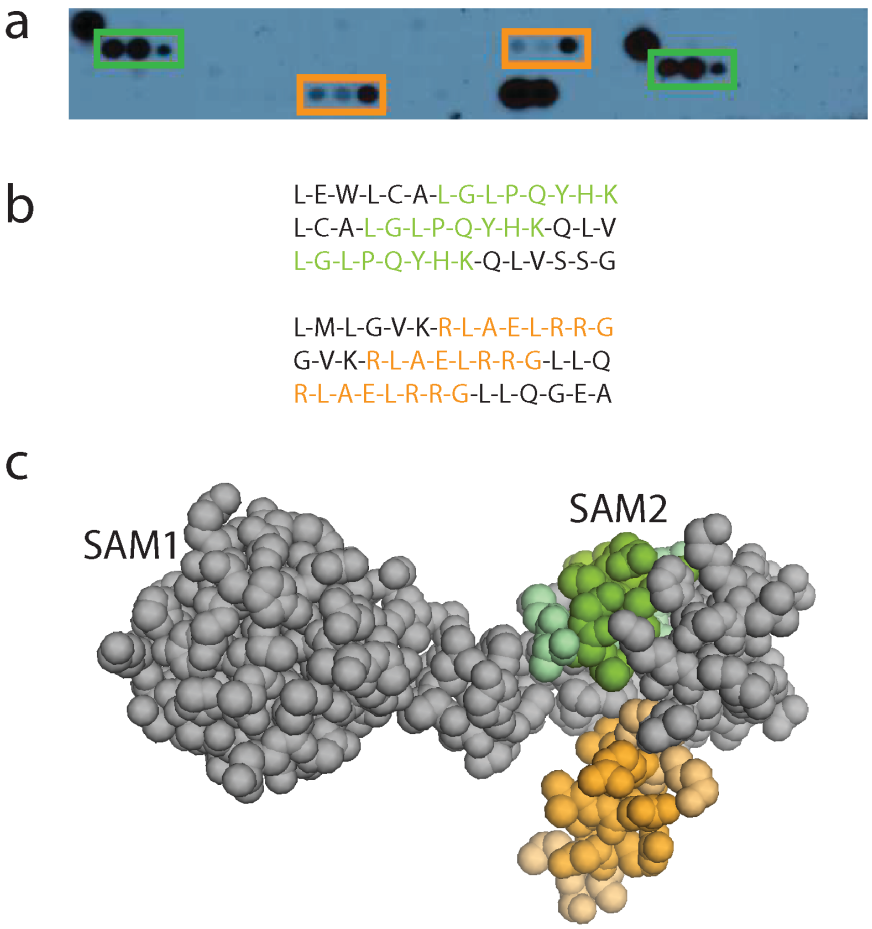


Figure 4. 1 – Peptide SPOT array analysis of the interaction between LARD2 and Caskin SAM1 SAM2. (a) The array comprising SAM1-SAM2 domain sequence of Caskin2 with three amino acid shift window probed with cMyc-His_6-LARD2. The resulting positive peptide sequences detected with His_6-HRP-conjugated antibody listed in **(b)** panel and consistently highlighted in green and orange color in all panels including the spherical molecular representation of SAM tandem structure **(c)** The eight core amino acids identified as consensus sequences are colored in darker shades and six flanking residues are colored in lighter shade of green and orange. Four additional, unboxed, spots are positive control peptides composed of 6xHis flanked by 3xAla on either side.

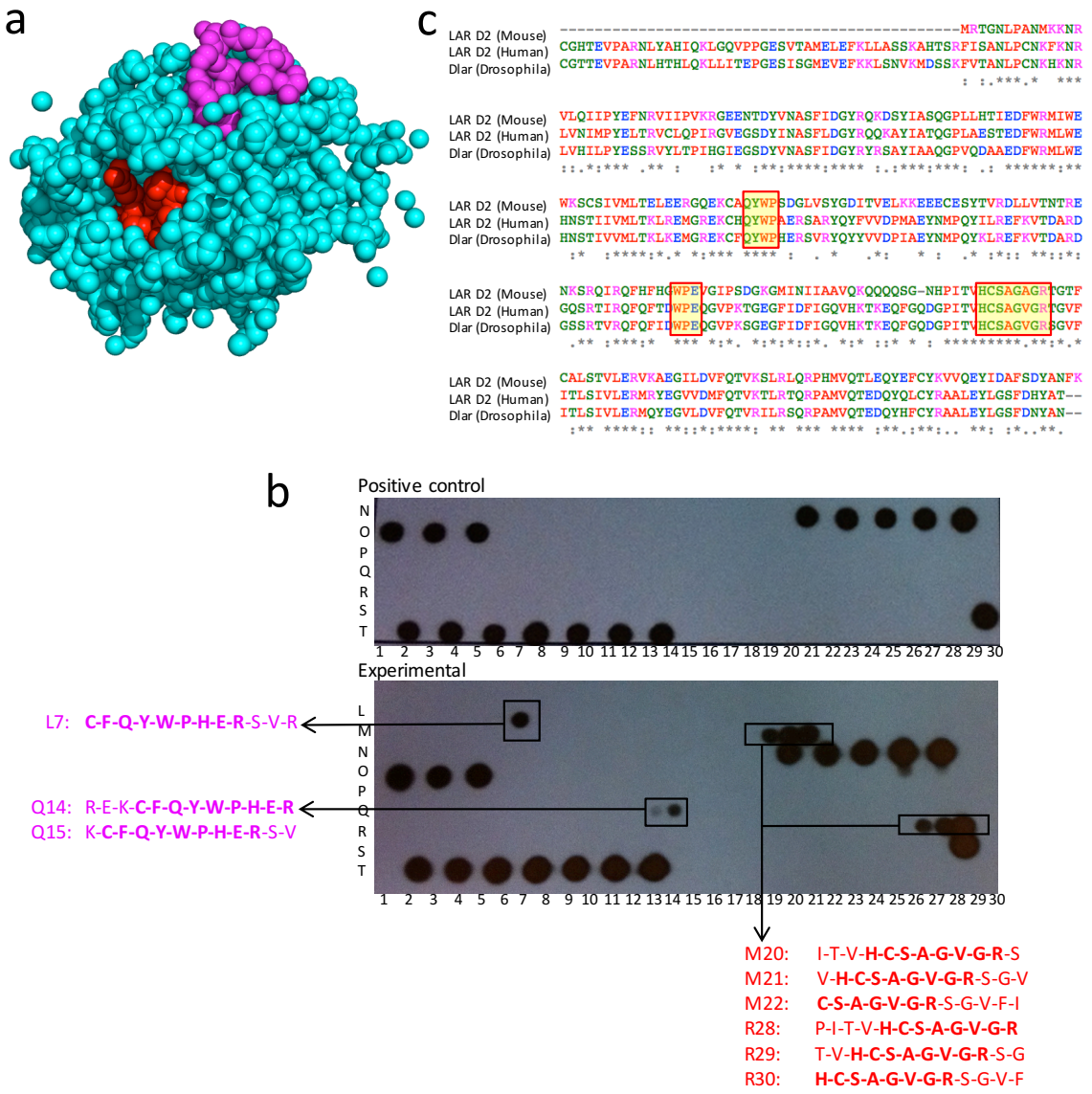


Figure 4. 2 – Peptide SPOT array of LARD2 sequence tested with Caskin SAM1 SAM2. **(a)** Spherical representation of the LAR D2 domain with regions identified by spot blot experiment colored in purple and red consistent with a panel **(b)** showing the LARD2 peptide array probed with His_6-Caskin SAM1 SAM2 detected with His_6-HRP-conjugated antibody as well as a positive control experiment where the array was probed with His_6-HRP antibody only. Resulting consensus sequences H-C-S-A-G-V-G-R (red) and C-F-Q-Y-W-P-H-E-R-S-V-R (purple) are indicated on the blot. **(c)** Sequence alignment of LAR D2 domains from different organisms. Highlighted regions corresponding to fully conserved regions within sequences identified by SPOT blot analysis. Additionally, a WPE-loop of D2 domain is highlighted. It corresponds to WPD-loop associated with (pY)-peptide binding cleft in the catalytically active D1 domain. An alignment was performed using Clustal Omega multiple protein sequence alignment program (http://www.ebi.ac.uk/Tools/msa/clustalo/). UniProtKB sequence IDs: P18052 LAR D2 (Mus musculus), P10586 LAR D2 (Homo sapiens), P16621 Dlar (*Drosophila* melanogaster).

Likewise, two distinct surfaces of the LARD2 domain demonstrated binding affinity for Caskin SAM1 SAM2 wild type. The signal detected for the consensus sequence H-C-S-A-G-V-G-R (colored in red) was relatively stronger than for the other: C-F-Q-Y-W-P-H-E-R-S-V-R (colored purple) **(Figure 4.2a, b)**. Additional experiments with G70A, K73E and G140A mutant versions of the Caskin2 tandem SAM domain produced similar results (data not shown).

The LAR SPOT blot result was followed up with an NMR-based titration of Caskin SAM1 SAM2 with an 11 amino acid-long synthetic peptide, REKCFQYWPHERSVR (CanPeptide, Montreal QC), corresponding to the "purple" SPOT blot sequence derived from *Dlar* (*Drosophila* ortholog) protein. The peptide was titrated up to 10-fold molar excess into the ^{15}N Caskin2 SAM1-SAM2 protein and monitored by ^{15}N-HSQC spectra (data not shown). The experiment demonstrated moderate signal broadening, but no significant chemical shift changes were observed, suggesting the absence of the interaction or possibility of very weak interaction.

The plausible binding sequence derived from helix five of Caskin SAM2 domain (151)R-L-A-E-L-R-R-G(158) (**Figure 4.1c** - orange**)** was tested as a candidate for LAR D1D2 interaction site. The rationale for choosing this sequence over the other was its location within a surface-exposed finger-like helix. A significant portion of the C-terminal end of helix 5 does not contribute to the SAM1-SAM2 intramolecular interface that stabilizes the domains, neither is it involved in the intermolecular dimerization/oligomerization interface. The C-terminal deletion version was sub-cloned

of more soluble and stable Caskin2 SAM1 SAM2 [G70D/K73E] mutant using Side directed mutagenesis (SDM) method, with truncation ending at -R-L-A*. This last portion of the α-helix is involved in intense hydrogen bonding network and indispensable for SAM2 domain overall structural fold and stability. The new construct was verified by sequencing and named C-term ΔCSS[G70D/K73E]. The IP experiments on GST and FLAG antibody-conjugated agarose beads (data not shown) revealed that the truncated construct ΔCSS[G70D/K73E] was able to recruit the LARD2 protein similarly to the wild type.

In order to further test the hypothesis that D2 domain of LAR is recognized by SAM2 domain of Caskin2 *in vitro* immuno-precipitation (IP) experiments were applied. FLAG-H_6-LARD1D2 was immobilized on highly specific FLAG beads and tested with single and tandem Caskin2 SAM domain constructs. Selected experiment presented in **Figure 4.3.** The reverse pull-down experiments were performed with immobilized GST-tagged Caskin2 protein constructs incubated with LARD1D2 and confirmed the SAM2 domain of Caskin2 as the interaction domain (**Figure 4.4**). The appropriate negative control experiments were performed to demonstrate the absence of nonspecific binding of LAR to GST-tag or GST-conjugated resin. Additional co-IPs to test LARD1D2 against various mutant versions of Caskin2 with amino acid alternations targeting the oligomerization interface: CSSG70A, CSSK73E, CSSG140A and CSS[G70D/K73E] were performed. All assessed mutants were shown to co-elute with LARD1D2 (data not shown) leading to the conclusion that Caskin2 oligomerization interface is not involved in the interaction.

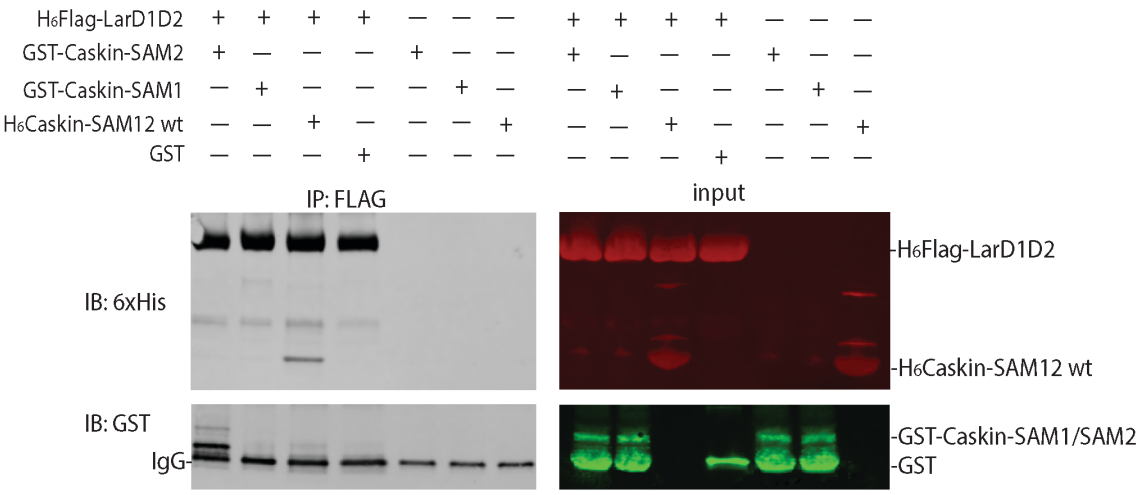


Figure 4. 3 – FLAG Immunoprecipitation. IP experiments with purified recombinant FLAG-H_6-LARD1D2 protein immobilized on the anti-FLAG-Ab-conjugated magnetic beads (Sigma-Aldrich) incubated with H_6-Caskin2 SAM1-SAM2 WT (purified) and GST-Caskin2 SAM2/SAM1 protein lysates. Experimental, negative control and loading control samples were transferred to Western blots. Double primary, Anti-H_6-rabbit and anti-GST-mouse antibodies, followed by IR-800-anti-rabbit and IR-600-anti-mouse secondary antibodies respectively were used for the detection and imaging with Odyssey infrared imager. H_6-Caskin2 SAM1-SAM2 WT and GST-Caskin2 SAM2 co-eluted with LARD1D2 protein, whereas GST-Caskin2 SAM1 did not. The low molecular weight band present in all experimental samples (left panel) was attributed to in the FLAG antibody light chain (IgG) that was nonspecifically detected by the infrared IR-600 antibody.

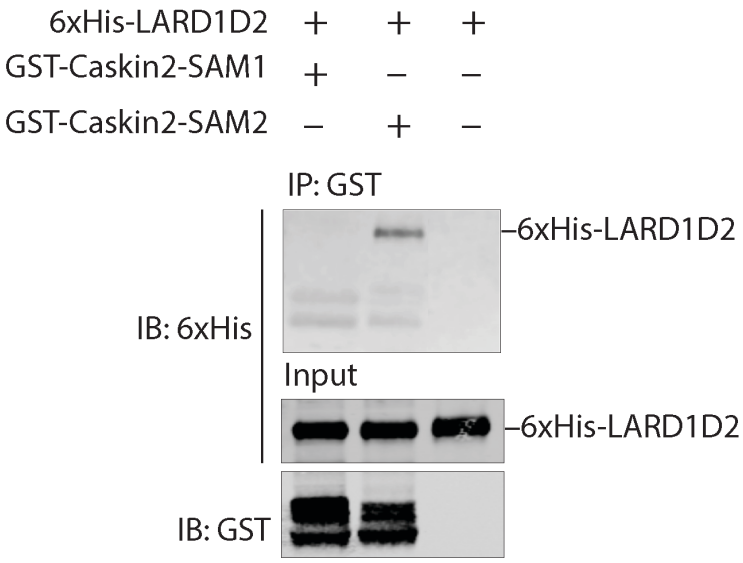


Figure 4. 4 – GST Immunoprecipitation. A complementary reverse IP experiment with GST-Caskin2 SAM2/SAM1 proteins immobilized on the anti-GST-Ab-agarose beads (Clontech) incubated with FLAG-H_6-LARD1D2 protein. Double primary, Anti-H_6-rabbit and anti-GST-mouse antibodies, followed by IR-800-anti-mouse and IR-600-anti-rabbit secondary antibodies were used for the detection and imaging with Odyssey infrared imager. LARD1D2 protein co-eluted with GST-Caskin2 SAM2.

4.3.2. Determination of the reaction kinetic parameters by fluorescence anisotropy

Building on the successful application of the fluorescence polarization spectroscopy method for determination of binding affinities of APP peptides to the PTB domain of AIDA-1 protein (Smirnova *et al*., 2013), I used this approach to study the LAR-Caskin interaction. The smaller protein, Caskin2 SAM1-SAM2 (20 kDa), was used for labeling and detection and LAR D1D2/D2 (60-70 kDa) as a titrant. A maleimide fluorophore was chosen as a probe because it is selective for and makes covalent bonds with exposed, reduced cysteine side chains. The labeling method particularly suitable for Caskin2 since there is only one cysteine in its tandem SAM domain. Titrations were performed using two methods: (1) microplate sampling method using the Synergy Hybrid Reader detection with a sample volume of 10 μL; (2) a Varian Agilent Eclipse spectrophotometer equipped with a manual polarizer accessory that requires 500 μL total sample volume. The titration set plotted as an average of four independent experiments is presented in **Figure 4.5.** The dissociation constant of 1.1 ± 0.4 μM for wild type Caskin2 SAM1-SAM2 interaction with LARD1D2 was calculated using nonlinear regression analysis. Additional LARD1D2 titrations with the maleimide-labeled Caskin2 SAM2, expressed as a single domain, resulted in a similar dissociation constant of 0.8 μM. Therefore, the microplate-based titrations with single SAM2 domain delivered K_d values comparable to a tandem SAM domain interaction with LARD1D2 derived from Varian titration series.

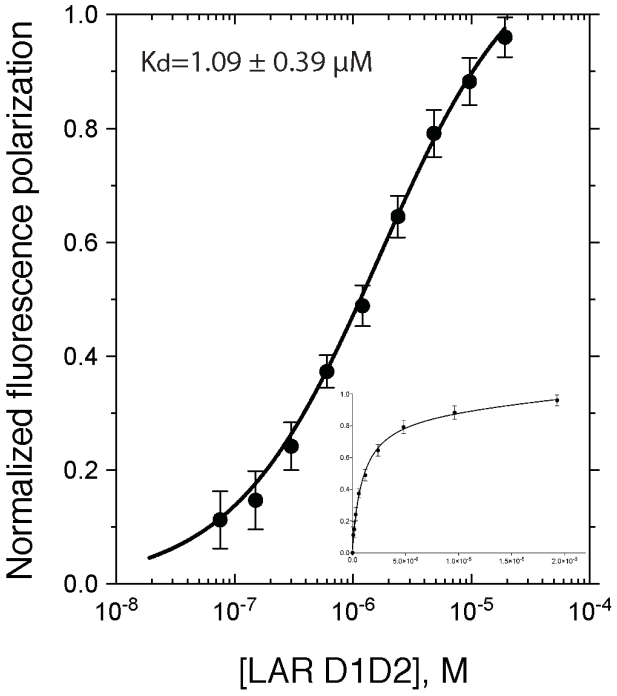


Figure 4.5 – Fluorescence anisotropy binding assay. Normalized fluorescence anisotropy titration curve of LAR D1D2 with maleimide-labeled Caskin2 SAM SAM WT plotted as mean values of four experiments with bars representing SEM errors for each data point. Outer plot: logarithmic scale; inset plot: linear scale. The plot was generated using Prism 7 graphing program (by GraphPad Software Inc.)

However, due to an extremely small sample volume (10 μL) and the high instrument sensitivity, the Synergy Hybrid Reader data sets were suffering from multiple experimental artifacts. Therefore, the final reported kinetic parameters were calculated from data sets obtained from Varian spectrophotometer-based experiments.

4.3.3. Assessment of subcellular distribution and colocalization between LAR and Caskin2 in neuroblastoma cells

An attempt to assess colocalization of Caskin and LAR *in vivo* and possibly determine their subcellular distribution was made by using fluorescent fusions of proteins in the mammalian expression vectors. Single and co-transfections were performed in Neuro2a cells using EGFP wild type Caskin2 SAM tandem (EGFP-CSSwt), the non-oligomerizing mutant EGFP-CSS[G537D/ K540E], dsRED-LARD2 and D1D2 fusion constructs (refer to **Figure 4.6** and **4.7**). Not unexpectedly, the single transfections with Caskin EGFP constructs demonstrated both cytoplasmic and nuclear distribution, whereas dsRED-LARD2/D1D2 were diffusely distributed in the cytoplasm, often with noticeable accumulation in the cell membrane. Since fluorescent fusion proteins were overexpressed as separate protein domains rather than full-length proteins, the EGFP-CSSwt (47 kDa) protein was small enough to diffuse through the nuclear pore (Wang & Brattain, 2007). The degree of co-localization of EGFP-CSSwt and dsRED-LAR single D2 and double D1D2 domain constructs were assessed by calculating Manders overlap coefficients (Manders *et al.,* 1993) (**Figure 4.6c**). Average overlap coefficients of 0.83 and 0.80 for EGFP-CSSwt with dsRED-LARD2/D1D2 respectively demonstrated substantial colocalization in both cases.

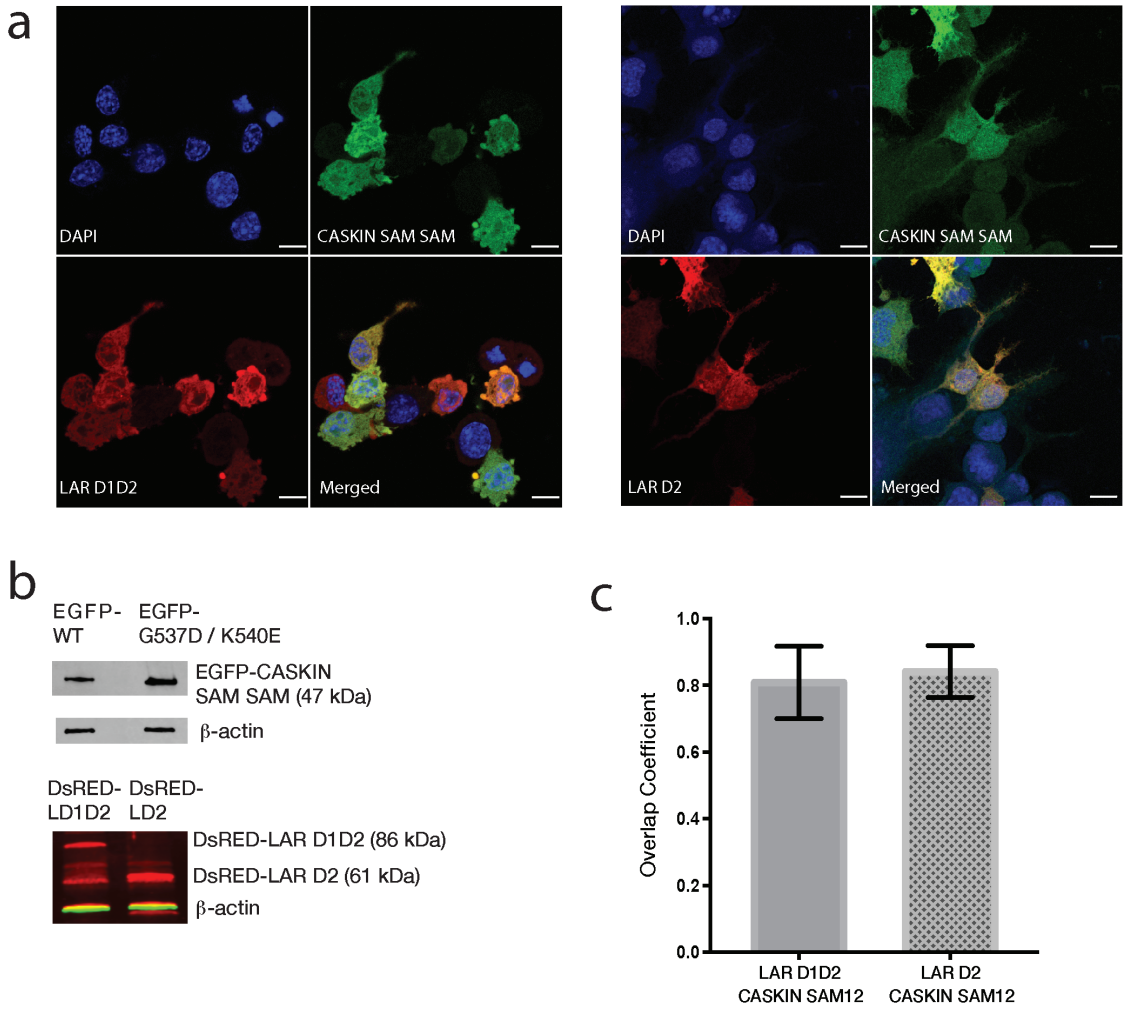


Figure 4.6 –Caskin2 SAM1 SAM2 and LARD1D2/D2 co-expression in Neuro2a cells. Neuro2a cells were transiently transfected with **(a)** pEGFP-N1-CSS w.t. and dsRED-mono-LARD1D2 and **(b)** dsRED-mono-LARD2; blue (DAPI) nucleus staining, green channel – pEGFP, red channel – dsRED-mono; images acquired at the 40x magnification with confocal microscope LSM 700. Scale bar: 10 µm. **(c)** Western blot analyses of whole cell N2A lysates were performed to assess relative expressions of pEGFPN1-Caskin2 and dsRED-mono-LARD1D2/D2 fusion proteins with β-actin as a loading control. Monoclonal anti-EGFP, anti-LAR and anti-β-actin primary and IR-860/800 secondary antibodies were used for detection. **(d)** Overlap coefficients ($1.0 \geq y \geq 0$) generated by the ZEN 2010 program for Caskin2 SAM1 SAM2 WT and LARD1D2/D2 plotted as mean values with SEM error bars using Prism7 program, where n=54 for LARD1D2 and n=58 for LARD2.

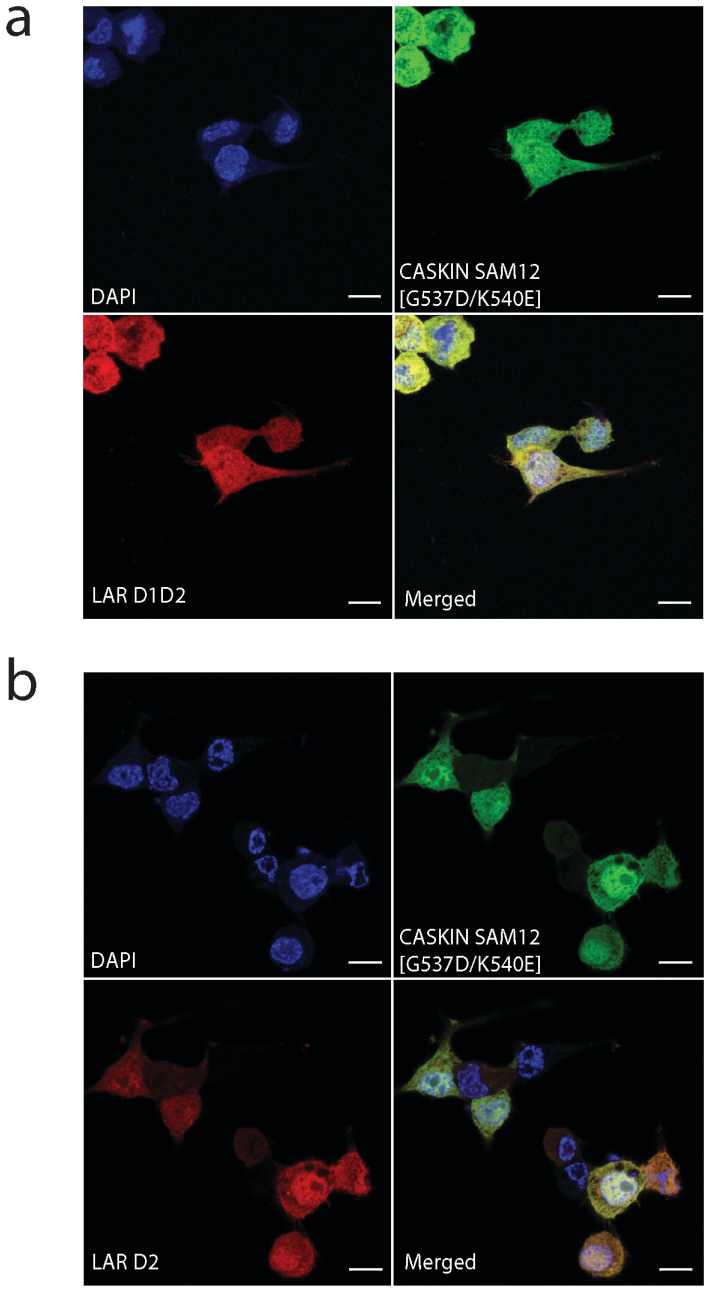


Figure 4.7 – Selected images of the double mutant Caskin2 (EGFP-G537D/ K540E) and LARD1D2/D2 co-expression experiments. Images were acquired at the 40x magnification using a Zeiss LSM 700 confocal microscope. Scale bar: 10µm.

4.4. Discussion

Apart from their direct phosphatase functions, the LAR RPTPs protein family only recently emerged as prospective synaptic cell adhesion molecules. This notion is supported by their high evolutionary conservation, the receptor-like structures and multiple extracellular and intracellular protein partnerships, including a set of synaptic membrane proteins linked to the synaptic development and synaptic adhesion pathways (Han *et al*., 2016; Um & Ko, 2013). A growing number of LAR PTPase family intracellular protein partners have been reported to interact with their membrane distal domains (D2), including Caskin, Liprin-α, Trio, diaphanous-related formins (DRFs), β-catenin, and Abelson tyrosine kinase (Abl)–Enabled (Ena)/vasodilator-stimulated phosphoprotein (VASP) proteins, missing in metastasis protein (MIM-B) (reviewed by Um & Ko, 2013). The proposed LAR-RPTPs-based synaptogenesis model is illustrated in **Figure 4.8.** where LAR-RPTPs have been suggested to act in conjunction with transmembrane receptors and intracellular signaling protein scaffolds, recruiting synaptic vesicles and active zone proteins; shaping the presynaptic differentiation; supporting the synaptic adhesion between dendrites and axons; and eventually transducing various postsynaptic signals.

Studies demonstrate that Caskin1, but not Caskin2 is able to participate in the CASK signaling scaffold (Stafford *et al*., 2011a). In addition, our group had recently shown *in vitro* and *in vivo* that Caskin2 tandem SAM domain oligomerizes differently (Smirnova *et al.*, 2016) from Caskin1 (Stafford *et al.*, 2011b). *Drosophila* Caskin and both

mammalian and *Drosophila* alpha-liprin proteins have been reported to directly interact with the D2 domains of LAR RPTPs via their SAM domain modules in a competitive fashion (Wei *et al.*, 2011; Serra-Pagès *et al.*, 1995; 1998; Stryker & Johnson, 2007). Both partnerships emerge as essential determinants of axon guidance in *Drosophila* (Spangler & Hoogenraad, 2007; Wei *et al.*, 2011). In addition, Liprin-α-LAR complexes recruit a number of known active zone modulators, such as RIM1-α, ERC2 and GIT1 (Ko *et al*., 2003; Schoch *et al*., 2002; Stryker & Johnson, 2007) required for synaptic organization and maturation. Although the biological consequences of LAR/Caskin2-specific signaling scaffold, distinct from Liprin-α and CASK/Velis/Caskin1/Mint has yet to be identified, the existence of two closely related, yet different, human Caskin homologs suggest their different roles in the neuron.

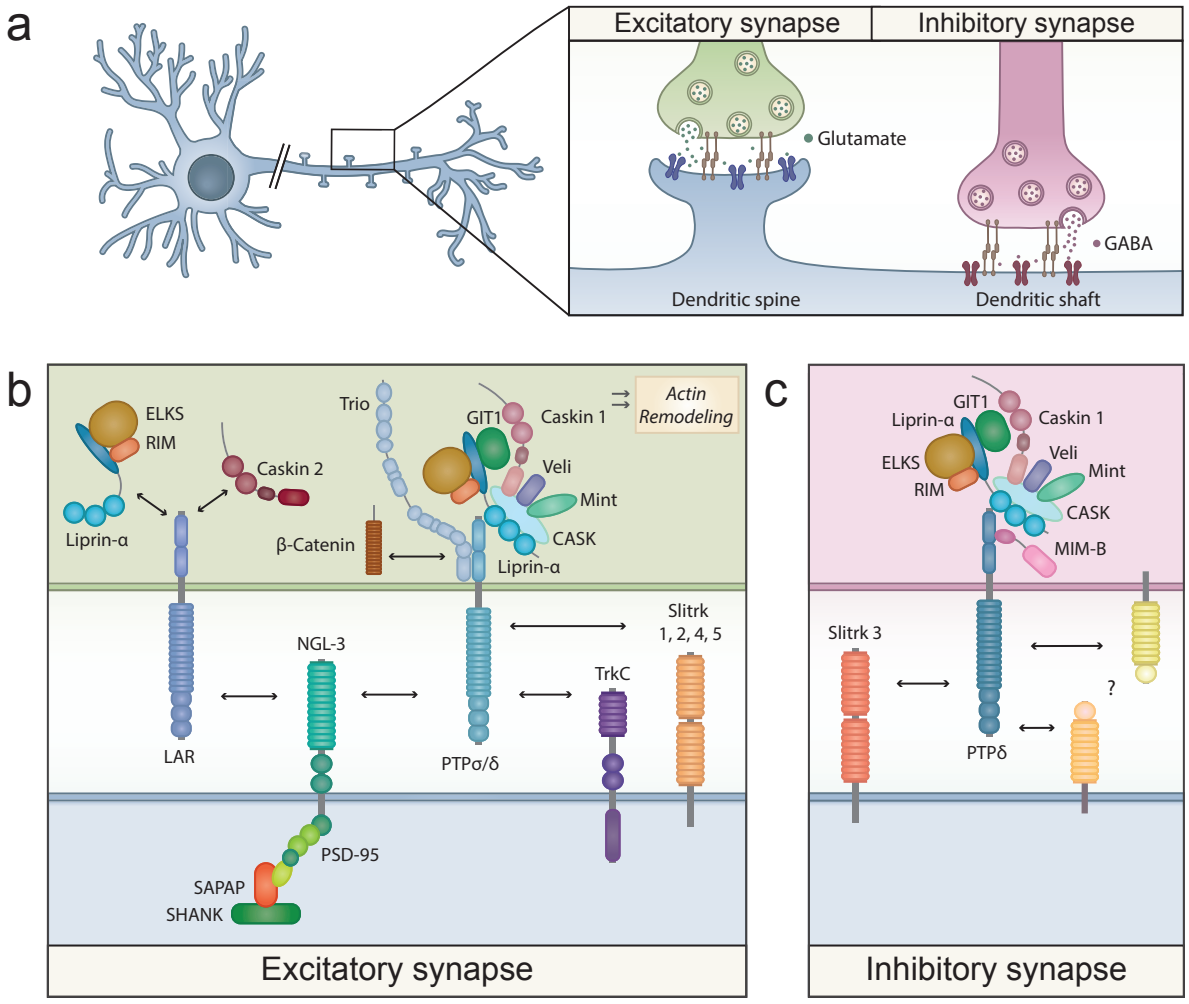


Figure 4. 8 – Leukocyte common antigen-related receptor protein tyrosine phosphatases (LAR-RPTPs) regulated synaptogenesis. (a) Schematic illustration of excitatory and inhibitory synapses. LAR-RPTPs associations with extracellular and intracellular synaptic protein complexes regulate excitatory synapse development. **(b)** Intracellular domains of LAR-RPTPs interact with diverse intracellular proteins, including liprin-α, Trio, β-catenin and Caskins, as part of synaptic retrograde neurotransmission. At the postsynaptic sites, LAR-RPTPs mediate anterograde neurotransmission through partnerships with postsynaptic proteins, such as protein PSD95/SAPAP/Shank complex. The intracellular interactions include proteins that bind receptor tyrosine kinase TrkC and Slitrks. **(c)** As an exception, PTPσ forms a complex with Slit- and Slitrk3 at inhibitory synapses. The possibility that specific coreceptors regulate LAR-RPTPs activity colocalizing with LAR-RPTPs at different subsets of synapses remains to be investigated (Han *et al*., 2016). Illustration revised from (Um & Ko, 2013) prepared by Lesia Szyca, (2016).

Figure Abbreviations: ELKS, glutamine, leucine, lysine, and serine-rich protein; Slitrk1/2/3/4/5, Slit- and Trk-like family protein (1-5); GIT1, G-protein-coupled receptor kinase-interacting protein 1; Shank, SH3 and multiple ankyrin repeat domains protein; PSD-95, postsynaptic density protein 95; SAPAP, PSD-95-associated protein; TrkC, neurotrophin receptor tyrosine kinase C; MALS/Veli, mammalian LIN-7/vertebrate homolog of LIN-7; Mint, Munc-18-interacting protein; NGL-3, netrin-G ligand-3; RIMs, Rab3-interacting molecules; MIM-B, missing in metastasis protein B.

I used the peptide SPOT blot method to identify the putative interaction surface on Caskin tandem SAM and LAR D2 domains. The first identified array-derived sequence C-F-Q-Y-W-P-H-E-R-S (*Drosophila*) is partially conserved across species (refer to **Figure 4.2c**), with a core sequence QYWP conserved in both D1 and D2 domains; the second peptide array sequence H-C-S-A-G-V-G-R corresponds to the PTPase active site which is fully conserved in both D1 and D2 domains. In the catalytically active LAR PTPase D1 domain the (p)Tyr-binding site is formed by three regions: a catalytic site (I/V)H-C-X-A-G-X-G-R(S/T)G with thiol residue is directly involved in nucleophilic attack on the pTyr and two additional loops contributing to acid-base catalysis and substrate recognition namely WPD-loop and pTyr-recognition loop. The tertiary structures of both intracellular domains of LAR are very similar and demonstrate remarkable sequence conservation, including preservation of the Cys-centered active site (Nam *et al.*, 1999). However, two amino acid differences in the D2 domain, Leu1644 (Tyr1644 in D1) and Glu-1779 (Asp1779 in D1), located in the pTyr recognition loop and WPD-loop respectively, induce a conformational difference and steric hindrance thereby preventing pTyr peptide from binding to D2 domain. Moreover, the reverse mutation of these residues (Leu1644Tyr; Glu1779) fully restored the robust PTPase catalytic activity of the D2 domain *in vitro* (Nam *et al.*, 1999). As mentioned earlier, in addition to Caskin,

several protein partners interact specifically with membrane distal domain (D2) although the precise recognition sites mostly remain unknown at the structural level. Taking together the spot blot observations and the fact that catalytic sites of both domains are ligand-accessible, it could be hypothesized that a "pseudo catalytic site" of the D2 domain may serve as a recognition site recruiting non-phosphorylated protein partners of LAR. Most likely, similarly to D1 domain recognition, D2 domain binding requires several non-consecutive surfaces. The SPOT blot method is applicable in a case of continuous surface-exposed binding epitopes, but could result in false positives or fail to identify the contact surface in which secondary structure defines the specificity, as well as if it is formed by non-consecutive residues. NMR titrations of Caskin with LAR-derived peptide (REKCFQYWPHERSVR) did not result in observation of substantial conformational changes in Caskin2 SAM1 SAM2 that would be considered indicative of binding. The other reason could be that to match the SPOT blot result the *Drosophila* sequence-based LAR peptide was tested against human Caskin SAM1-SAM2 protein. However, surveying the sequence alignment, it is apparent that this peptide is only partially conserved between human/mouse and fly orthologs (**Figure 4.2c**). An additional mutagenesis study would be required to confirm the interaction surface on LARD2. Alanine scanning mutagenesis or small truncations targeting the H-C-S-A-G-V-G-R pseudo-catalytic site (PTP-loop) and Q-Y-W-P-loop sequence could be tested. The human LAR sequence-derived peptides R-E-K-C-H-Q-Y-W-P-A-E-R and H-C-S-A-G-V-G-R could be tested for interaction with Caskin2 SAM1 SAM2 by NMR and/or ITC titrations. Another possibility is elucidating and characterizing the D2 binding site via structural

study of LARD2/D1D2 bound to Caskin or Liprin-α, or ligand-derived peptide (by X-ray crystallography).

Like Caskin2, Liprin-α is another SAM domain-containing protein known to interact with the LARD2 domain. Although the Liprin-LAR interaction has yet to be characterized at the structural level, it has been reported to take place at the triple-SAM-domain module of Liprin-α which is highly conserved in entire liprin family, and also called liprin homology domain (LHD) (Serra-Pagès *et al.*, 1995; 1998; Serra-Pagès, Streuli, & Medley, 2005). Furthermore, the intrinsic autophosphorylation ability had been suggested to regulate the liprin family protein partnerships (Serra-Pagès *et al.*, 2005). Another similarity between Caskins and Liprins is their predisposition to oligomerization, although Caskins oligomerize via SAM domains and liprins had been suggested to dimerize via their coiled-coil domains (Serra-Pagès *et al.*, 1998). In addition to homo-oligomerization, *Drosophila* liprins β and *γ* were shown to hetero-oligomerize with liprin α (Astigarraga *et al.*, 2010). As mentioned earlier, the triple SAM domain module is remarkably conserved among liprin family members. Surveying a sequence alignment of all Liprin-α human homologs with tandem SAM domains of human Caskins **(Figure 4.9b)** two sequence areas drive an attention as they demonstrate an exceptionally high degree of conservation, though overall conservation is rather low. For the purpose of this discussion, I designated these sequences as the "*conservation region 1* and *2*" (labeled accordingly in the figure **Figure 4.9b**). *Conservation region 2,* colored in green on Caskin structure **(Figure 4.9a),** corresponds to the sequence (100)L-C-A-L-G-L-P-Q-Y-H-K(110) within Caskin2 SAM2 domain as identified by SPOT blot analysis. It is a solvent exposed

continuous sequence region with the exception of a downstream leucine residue (L152) located in fifth α-helix, likewise fully conserved in all assessed protein sequences. Interestingly, glutamate (E98) and proline (P106) residues are conserved in all liprins and Caskin2, but not Caskin 1 (denoted with red boxes in the **Figure 4.9b**). From Caskin2 domain-swapped homodimer structure we learned that salt bridges between E98 and R151 reinforce the dimerization interface (Smirnova *et al.*, 2016); hence glutamate 98 is one of the residues defining the distinct mode of Caskin2 dimerization, not observed in Caskin1.

Another consensus sequence I-G-I/V-S/T-x-P-L/G-H-R denoted as *conservation region 1* is located within the first SAM domain of Caskin **(Figure 4.9a).** Surveying this region in the Caskin SAM1-SAM2 structure, it is evident that residues G70, H71, K73, K,74, plus two conserved residues upstream, L36 and Y39, contribute to the solvent-exposed conserved cluster. Therefore, in contrast to *conservation region 2*, the conserved surface 1 is non-continuous, hence methods such as peptide array would fail to identify it as an interaction site. However, in light that amino acids G70, K73 and K74 are located at the Head-to-Tail oligomerization interface of Caskin2 the high degree of conservation is not surprising since these surfaces are highly preserved in most SAM domains. According to our *in vitro* immunoprecipitation studies single/double mutants G70D, K73E, and [G70D/K73E] do not perturb interaction with LAR. Likewise, the $_{929}$LHR$_{931}$ (triple A mutant) Liprin-α1 assessed in (Serra-Pagès *et al.*, 2005) study remained LAR-active. At the same time, the D$_{990}$(A) mutation in SAM2 domain, assessed in the same study, resulted in reduced LAR binding as well as perturbed liprin auto-phosphorylation both *in vivo* and *in vitro* (Serra-Pagès *et al.*, 2005). Orange boxes denote all mutants examined by Serra

Pages *et al.* (2005) in **Figure 4.9b**. Furthermore, the most recent study by Astigarraga *et al.* (2011) confirmed that liprin-α and a new isoform that they named liprin-γ, but not liprin-β, interact with the D2 domain of Lar via LHD (triple-SAM domain). The LAR-liprin(s) partnership was once more shown to be necessary for normal synapse development in *Drosophila* in both R7 photoreceptors and motor neurons, with liprin-γ possibly serving as a functional antagonist counteracting and/or regulating the two other liprins functions (Astigarraga *et al.*, 2010).

Another mutagenesis coupled with yeast two-hybrid study by Weng *et al.* (2011) (in *Drosophila*), however, described Csk SAM1(R305Q) and Csk SAM1Δ324-331 mutant constructs as Lar-inactive. Initially, these regions were selected for mutagenesis as they were earlier connected to the loss-of-function phenotypes (LOFs) causing various degrees of motor axon pathfinding disruption. However, an examination of these mutations from a structural perspective suggests that the R305Q mutation and large Δ324-331 truncation would have the serious structural consequences: by disrupting the SAM1 domain structure they would consequently also perturb the SAM1-SAM2 intra-molecular interface, thus SAM1-SAM2 module integrity. Moreover, the perturbation at SAM1 domain would likely manifest in the loss of Caskin2 oligomerization, which is a known requirement for proper function of many scaffolding proteins. Therefore, the observed loss of function was rather anticipated in case of SAM1 domain mutants; at the same time, Weng *et al.* (2011) study inarguably confirmed, by a number of methods, the intact SAM1-SAM2 module of Csk as an interaction site for Lar PTP (Weng *et al*., 2011b).

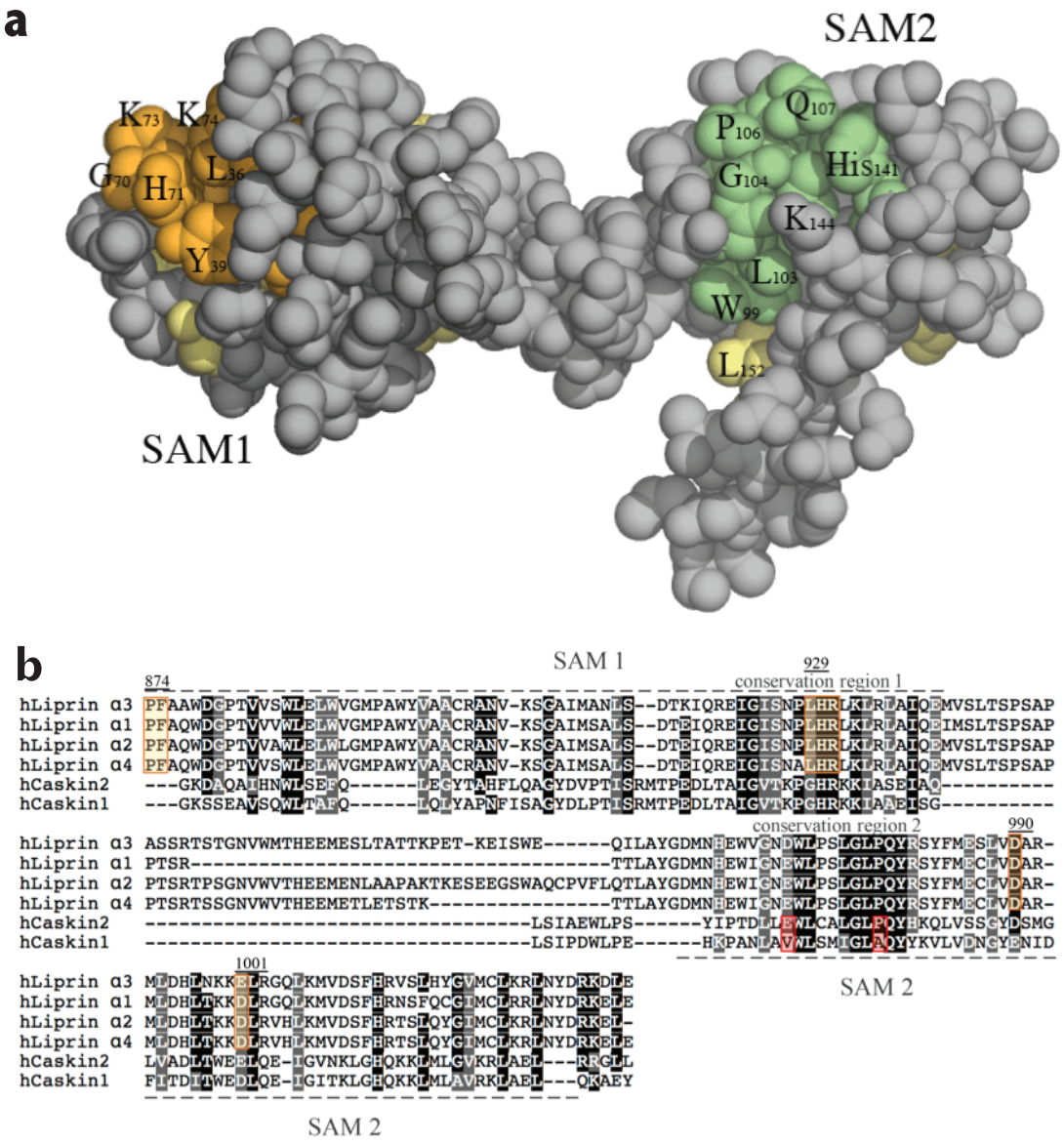


Figure 4. 9 – Sequence conservation of SAM1-SAM2 domains of human Liprin family members and human Caskins. (a) Caskin2 SAM tandem with highly conserved "region one" in SAM1 (orange) and "region two" (green) in SAM2 domains; other fully conserved residues colored in yellow. **(b)** Black regions denote the identity regions, grey – substitutable amino acids. Two red boxes denote differences between Caskin2 and Caskin1 in the "conservation region two". Orange boxes indicate liprin mutants $_{874}$PF$_{875}$(AA), D$_{990}$(A), D$_{1001}$(A), $_{929}$LHR$_{931}$(3A) evaluated in (Serra-Pagès, Streuli, & Medley, 2005) study.

An alignment was performed using Clustal Omega multiple protein sequence alignment program (*www.ebi.ac.uk/Tools/msa/clustalo*) and color-coded using Color Align Conservation program (*www.bioinformatics.org/sms2/color_align_cons.html*) UniProtKB sequence IDs: human liprin-α2 (AF034799), human liprin-α4 (AF034801), human liprin-α1 (U22815), human liprin-α3 (AF034800), Caskin2 (Q8WXE0), Caskin1(Q8WXE9). The SAM3 domain of liprin was excluded from the alignment.

Within the structural context of this discussion it is worth describing the Liprin-α2 interaction with C-lobe of CASK CaMK domain which is currently the only Liprin-based complex reported at atomic resolution. The liprin-α2/CASK complex crystal structure, solved by Wei at al. (2011), revealed that the intramolecular interface is formed by four salt bridges and an intense hydrogen bonding network from both SAM1 and SAM2 domains of Liprin, as well as short insertion helix between them 978-GNVWVTHE-985 (namely α_L helix). The Trp981 serves as a hydrophobic anchor intercalated deeply into CaMK-binding cleft (Wei *et al.*, 2011). Interestingly enough, α_L helix only exists in the α2, α3, and α4 isoforms of mammalian liprins and absent in *C-elegans* and *Drosophila* as well as in the mammalian liprin-α1 isoform, prompting liprin-α 2/CASK signaling complex as vertebrate-specific (Wei *et al.*, 2011). In the case of Caskin tandem SAM structure similar epitope is present three times: first in SAM1 domain (-NWL-), then in the conservation region SAM2 (-EWL-); in both cases, the tryptophans are buried in the molecular surface and incorporated in the α-helical motifs. However, the third sequence is located in the flexible linker between first and second SAM domain and conserved -E/D-W-L- in both Caskin2/1 homologs. In the crystal structure of the domain-swapped Caskin2 dimer the linker region provides the hydrophobic homo-dimerization interface. In contrast, in solution, the mix of monomeric/dimeric states is observed and the linker appears to be flexible (by NMR) (Smirnova *et al.*, 2016). Therefore, it could be speculated that linker W554 could potentially serve as a hydrophobic anchoring point similarly to (-VWV-) motif of Liprin-α2 insertion helix α_L. This hypothesis could be tested by the computational molecular docking methods and mutagenesis.

The extensive *in-vitro* immunoprecipitation series confirmed the interaction between mammalian LARD2 and Caskin SAM1-SAM2 homologs, as well as the requirement of SAM2 domain for the interaction. The single SAM1 domain expressed as GST-fusion did not co-elute with LARD1D2. However, from the structural studies, we have learned that SAM1 domain is highly unstable without SAM2 domain, therefore, the possibility of a SAM1-SAM2 interface or the linker region in between providing an additional specificity to the interaction could not be entirely eliminated. Additionally, the fluorescent anisotropy titrations were performed to characterize the interaction further. The low micromolar affinity was determined for LAR-D1D2 interaction with Caskin2 SAM1-SAM2, which is common for scaffolding/signaling assemblies at the synaptic sites where proteins must rapidly sample multiple ligands and relatively fast association/dissociation is required for efficient signal transduction (Burack & Shaw, 2000; Pan, Sudol, Sheetz, & Low, 2012; Pawson, Scott, & Scott, 1997). For comparison, liprin-a2_LH binds to CASK_CaMK with a Kd of ~0.6 μM (Wei *et al.*, 2011). Finally, co-transfections in Neuro-2a cells revealed a high degree of colocalization between LARD2/D1D2 and wild type and double mutant Caskin2 SAM1-SAM2 fluorescent constructs. The overexpression of LARD2/D1D2 (single transfections) demonstrated membrane accumulation in selected cells, although only cytosolic domains were expressed. That could be an indication of physical interactions with transmembrane proteins such as N-cadherin (Siu, Fladd, & Rotin, 2007) or other co-receptors (Han *et al.*, 2016) yet to be identified.

To summarize, the experimental evidence supports the interaction between mammalian orthologs of LAR and Caskin2. Although additional investigations are required to confirm the minimal binding epitopes, two regions were identified as plausible binding sites. Both are highly evolutionary conserved across evolution not only in human Caskin and Liprin homologs, but liprin-β1/2 and *Drosophila* liprin, *C. elegans* liprin-α and *Drosophila* Ckn. *Conservation region 2* appears to be the most plausible binding recognition epitope, since it remains ligand-accessible in Caskin2 monomeric or dimeric/oligomeric forms, whereas *conservation region 1* overlaps with the Caskin2 oligomerization interface (Smirnova *et al.*, 2016). Additional mutagenesis studies, peptide titrations, and SPOT blot array of Caskin SAM tandem sequence with consecutive residue substitutions can be conducted to further elucidate the binding requirements and provide a framework for a more detailed NMR/X-ray structural investigation of the LAR-Caskin complex.

CHAPTER 5:

CONCLUDING REMARKS AND FUTURE DIRECTIONS

5.1. Summary of Research

Neuronal protein AIDA-1 links synaptic signaling events with global changes in gene expression. Among other neuronal "shuffling" proteins AIDA-1 was suggested to serve in NMDAR-regulated neuronal signaling and protein trafficking, essentially linked to a long-term memory formation and synaptic plasticity (Dudek, 2007; Jordan & Kreutz, 2009). In addition, AIDA-1 has been implicated in APP processing; Cajal body regulation, including nucleolar formation and stability; PSD structural remodeling (Xu & Hebert, 2005; Dosemeci *et al.,* 2015, Jordan *et al.,* 2009, Tindi *et al.*, 2015). Loss of AIDA-1 functions have serious consequences for human health: a number of genetic studies have linked the AIDA-1 gene *ANKS1B* to schizophrenia (Purcell *et al.*, 2014; Snyder & Gao, 2013) and autism spectrum disorders (ASDs) (Pinto *et al.*, 2014; M. Uddin *et al.*, 2014).

In this study, we reported the NMR structure of the carboxy-terminally located PTB domain, preserved to all AIDA-1 splice variants, that could be viewed as the business end of the molecule capable of supporting multiple protein partnerships including amyloid protein precursor (APP), a Cajal body protein coilin, and Ephrin A8 receptor tyrosine kinase. Using a comprehensive survey of peptides, the consensus sequence YxNxΦYxΨFE around an NxxY motif of APP was identified. Employing a peptide docking simulation, we characterized the AIDA-1 PTB binding cleft specificity to AβPP-AICD derived minimal peptide. In addition, the binding affinity of ~10 μM for the PTB-APP interaction was determined using fluorescence anisotropy method. We have established that the AIDA-1 PTB domain recognizes unphosphorylated-tyrosine NxxY motif-containing ligands, a binding mode typical for the Dab-like class protein family, although the consensus sequence NxxY is less stringent than NPxY required for other Dab family proteins. A comparison of the binding clefts of the AIDA-1 PTB with X11 and Fe65 revealed a binding mode analogous to the X11/Mint PTB-APP interaction. The moderate affinity and slightly different minimal ligand recognition epitope could be advantageous for AIDA-1-regulated signaling, alternative to X11, as well as in non-APP signaling contexts.

Two mammalian homologs, Caskin 1 and Caskin 2, are neuronal scaffolding proteins highly enriched in neuronal synapses (Tabuchi *et al.*, 2002) as well as retinal synapses (Anjum *et al.*, 2014). Caskin1, but not Caskin2 contributes to CASK (calcium/calmodulin-dependent serine kinase) regulated signaling pathways. Despite an overall high level of homology, the apparent differences in their oligomerization modes

suggest the possibility of distinct functional outcomes in neuronal signaling circuits. The reported crystal structure of the Caskin2 SAM domain tandem revealed an oligomeric form different from Caskin1, with the minimal repeating unit being a domain-swapped dimer, rather than a monomer. Examination of the dimer interface suggested that residues critical for homo-dimerization in Caskin2 were not preserved in Caskin1, providing a possible explanation why Caskin2 and Caskin1 oligomerize differently *in vitro* and *in vivo*. Such structural diversity could be advantageous in a concentration-dependent mechanism regulating low-affinity protein interactions, common for signaling pathways. This study contributes to understanding how neuronal SAM domain-containing proteins facilitate the assembly of large macromolecular complexes in order to concentrate and amplify synaptic responses.

Growing evidence of the mammalian LAR receptor tyrosine phosphatase regulatory functions in complex processes such as axonogenesis, synaptic assembly, neuronal plasticity, and its implications for human health, including neuronal regeneration, autism spectrum and other neurodegenerative disorders, demands a detailed elucidation of mammalian LAR PTP signaling pathways. The experimental evidence supports the interaction between mammalian orthologs of LAR and Caskin2 as presented in this dissertation. A combination of biochemical and biophysical techniques was used to characterize the interaction between the membrane distal domain D2 of LAR and the tandem SAM domain of Caskin2. Although additional experiments would be required to characterize the interaction surface and binding specificity, accumulated experimental data suggests the minimal requirement of SAM2 domain of Caskin2.

5.2. Future Directions

5.2.1. AIDA-1

Considering that the abundance of AIDA-1 in the PSD is comparable to GKAP guanylate kinase-associated proteins and PSD-95 fractions (Lowenthal, Markey & Dosemeci, 2015), our current understanding of its physiological functions is very modest. Several outstanding questions remain to be addressed: What regulatory mechanisms are encoded by AIDA-1 via alternative splicing? Are the subcellular locations of AIDA-1 isoforms a reflection of their distinct functions? Moreover, the mechanism of AIDA1 NMDA receptor specific functions has yet to be fully elucidated. The AIDA-1 nuclear translocation mechanism and the gene expression consequences are also not entirely understood.

NMDAR activation triggers the proteolytic cleavage and subsequent translocation of the AIDA-1d fragment to the nucleus where it associates with Cajal bodies and stabilizes its interaction with nucleoli (Jordan *et al.*, 2007). Silencing AIDA-1 through siRNA knockdown resulted in disruption of Cajal bodies and increased cell death rate (Xu & Hebert, 2005). The basis of AIDA-1 interaction with Cajal bodies protein Coilin has not been elucidated at the structural level. Preliminary investigations, performed by our laboratory members, suggested the interaction is between the Coilin Tudor domain and PTB domain of AIDA-1. Additional mutagenesis and structural studies would be required to verify this hypothesis.

Bryen A. Jordan's group suggested that AIDA-1 acts as a subunit-specific NMDAR transport facilitator. AIDA1 mediation of NMDAR function by fluctuations of GluN2B subunit levels has been proposed (Tindi *et al.,* 2015). The structural basis for preferential binding of AIDA-1 to GluN2B versus GluN2A is currently unknown. This research avenue is necessary for understanding of AIDA-1 protein contributions to pathogenesis of neuropsychiatric disorders such as autism and schizophrenia, which are strongly associated with NMDA receptor synaptic dysfunctions.

The observations that the short AIDA-1a isoform demonstrates preferential binding to AβPP-AICD while isoform b is AβPP-inactive (Ghersi et al., 2004a) suggested the possibility of a self-inhibitory mechanism in which the sequence region encoded by exon14 is precluding AIDA-1 from association with AβPP by direct interaction with PTB domain. Similar self-inhibitory mechanisms have been described for PTB domains of X11 (Matos et al. 2012) and Talin (Goksoy et al. 2008) proteins. Although our investigations of the AIDA-1 self-inhibition, conducted up to date, delivered controversial results, if this segment of the study is pursued the possible contributions of exon14 flanking sequences and/or phosphorylation events should be considered. Phosphorylation-induced conformational change displaces Fe65, but does not prevent X11s from binding to the GYENPTY motif of the AβPP cytoplasmic domain (Suzuki & Nakaya, 2008).

The PTB domains of AIDA-1, X11 and Fe65 target the same overlapping binding site on AICD, and the competitive nature of their interaction with AICD has been previously reported. It would be instrumental to determine if a subset of AIDA-1-positive

synapses overlaps with X11 in order to address a key question of whether they have complementary or distinct roles in APP processing. Further structural and biochemical studies of AIDA-1 would lead towards better understanding of neuronal signaling pathways.

5.2.2. Caskin2 and Lar Tyrosine Phosphatase

Since the minimal binding requirement has not been determined for any of the currently known LAR intracellular ligands, pursuing the structural studies could provide insight into the ligand binding specificity of the D2 domain of LAR PTPase. An additional mutagenesis study would be required to confirm the interaction surface on LARD2. If the minimal binding epitope within LAR originates from consensus peptide sequence, further NMR-based structural investigations would be possible. Another option is elucidating and characterizing the D2 binding site via a structural study of LARD2/D1D2 bound to Caskin, Liprin-α, or ligand-derived peptide by X-ray crystallography. A competition between Liprin-α and Caskin2 for LAR could be tested by applying a fluorescent polarization spectroscopy method. Complementary ITC titrations could be used to confirm the binding affinity of LARD1D2 and Caskin SAM1-SAM2 determined by the fluorescence anisotropy method.

With respect to the synaptic functions of LAR PTP family proteins synaptic functions, many outstanding questions remain to be addressed: (1) whether or not LAR-PTPase activity regulates LAR interactions with intracellular ligands; (2) ligand selectivity and specificity that allows individual LAR family members (LAR, PTPσ and

PTPγ) to target a distinct set of protein partners; and (3) what functional consequences and regulatory mechanisms are regulated by specific synaptic localization and alternative splicing of mammalian LAR family members. Although a comprehensive functional investigation is beyond the scope of this dissertation and our laboratory specialization, additional behavioral, neurological and electrophysiological studies would be necessary to investigate if dysfunctions of LAR RPTPs postsynaptic ligands are directly associated with disruptions of LAR functions. The detailed characterization of LAR-regulated synaptic adhesion architecture would contribute to a fundamental knowledge of synaptic development and synaptic plasticity.

5.3. Concluding remarks

The modular nature of signal transduction proteins endorses them as the ideal modulators of cellular signaling circuits through the ability to interact with a diverse array of ligands. The structure-directed studies, described in this dissertation, expand our knowledge about domain organization and oligomerization properties of neuronal proteins AIDA-1 and Caskin2.

Structure-functional characterization of AIDA-1 PTB domain as the interaction hub, together with its SAM tandem structure, described earlier by our group, could provide a framework for future functional studies ultimately revealing the functional details of AIDA-1 functions at the synapse with important consequences for human health.

The reported distinct structural attributes of Caskin2 and its partnership with LAR RPTP is one step towards understanding their cellular functions. With much hope these studies would contribute to future investigations improving our understanding of LAR-RPTPs signaling outcomes in the development of various neurological conditions.

REFERENCES

Arons, M. H., Lee, K., Thynne, C. J., Kim, S. A., Schob, C., Kindler, S., et al. (2016). Shank3 Is Part of a Zinc-Sensitive Signaling System That Regulates Excitatory Synaptic Strength. *The Journal of Neuroscience*, 36(35), 9124–9134.

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, 66(2), 213–221.

Agrawal, P., & Hardin, P. E. (2016). The Drosophila Receptor Protein Tyrosine Phosphatase LAR Is Required for Development of Circadian Pacemaker Neuron Processes That Support Rhythmic Activity in Constant Darkness, But Not during Light/Dark Cycles. *The Journal of Neuroscience*, 36(13), 3860–3870.

Ala, U., Piro, R. M., Grassi, E., Damasco, C., Silengo, L., Oti, M., et al. (2008). Prediction of human disease genes by human-mouse conserved coexpression analysis. *PLoS Computational Biology*, 4(3), e1000043.

Almeida, F. C. L., Moraes, A. H., & Gomes-Neto, F. (2013). An Overview on Protein Structure Determination by NMR: Historical and Future Perspectives of the use of Distance Geometry Methods. In Distance Geometry (pp. 377–412). New York, NY: *Springer New York.*

Ando, K., Iijima, K. I., Elliott, J. I., Kirino, Y., & Suzuki, T. (2001). Phosphorylation-dependent regulation of the interaction of amyloid precursor protein with Fe65 affects the production of beta-amyloid. *Journal of Biological Chemistry*, 276(43), 40353–40361.

Anjum, R., Ayoubian, H., & Schmitz, F. (2014). Differential synaptic distribution of the scaffold proteins Cask and Caskin1 in the bovine retina. *Molecular and Cellular Neuroscience*, 62(C), 19–29.

Astigarraga, S., Hofmeyer, K., Farajian, R., & Treisman, J. E. (2010). Three Drosophila liprins interact to control synapse formation. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 30(46), 15358–15368.

Aviv, T., Amborski, A. N., Zhao, X. S., Kwan, J. J., Johnson, P. E., Sicheri, F., & Donaldson, L. W. (2006). The NMR and X-ray Structures of the Saccharomyces cerevisiae Vts1 SAM Domain Define a Surface for the Recognition of RNA Hairpins. *Journal of Molecular Biology*, 356(2), 274–279.

Aviv, T., Lin, Z., Lau, S., Rendl, L. M., Sicheri, F., & Smibert, C. A. (2003). The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nature Structural & Molecular Biology*, 10(8), 614–621.

Awasthi, M., Singh, S., Pandey, V. P., & Dwivedi, U. N. (2016). Alzheimer's disease: An overview of amyloid beta dependent pathogenesis and its therapeutic implications along with in silico approaches emphasizing the role of natural products. *Journal of the Neurological Sciences*, 361, 256–271.

Baker, M. W., & Macagno, E. R. (2010). Expression levels of a LAR-like receptor protein tyrosine phosphatase correlate with neuronal branching and arbor density in the medicinal leech. *Developmental Biology*, 344(1), 346–357.

Balázs, A., Csizmok, V., Buday, L., Rakács, M., Kiss, R., Bokor, M., et al. (2009). High levels of structural disorder in scaffold proteins as exemplified by a novel neuronal protein, CASK-interactive protein1. *FEBS Journal*, 276(14), 3744–3756.

Baron, M. K., Boeckers, T. M., Vaida, B., Faham, S., Gingery, M., Sawaya, M. R., et al. (2006). An architectural framework that may lie at the core of the postsynaptic density. *Science*, 311(5760), 531–535.

Beckett, C., Nalivaeva, N. N., Belyaev, N. D., & Turner, A. J. (2012). Nuclear signalling by membrane protein intracellular domains: the AICD enigma. *Cellular Signalling*, 24(2), 402–409.

Bhattacharya, A., Wunderlich, Z., Monleon, D., Tejero, R., & Montelione, G. T. (2007). Assessing model accuracy using the homology modeling automatically software. *Proteins: Structure, Function, and Bioinformatics*, 70(1), 105–118.

Borg, J. P., Lõpez-Figueroa, M. O., Lõpez-Figueroa, M. O., de Taddèo-Borg, M., de Taddèo-Borg, M., Kroon, D. E., et al. (1999). Molecular analysis of the X11-mLin-2/CASK complex in brain. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 19(4), 1307–1316.

Borg, J. P., Ooi, J., Levy, E., & Margolis, B. (1996). The phosphotyrosine interaction domains of X11 and FE65 bind to distinct sites on the YENPTY motif of amyloid precursor protein. *Molecular and Cellular Biology*, 16(11), 6229–6241.

Bórquez, D. A., González-Billault, C. (2012). The amyloid precursor protein intracellular domain-fe65 multiprotein complexes: a challenge to the amyloid hypothesis for Alzheimer's disease? *International Journal of Alzheimer's Disease*, 2012, 353145.

Borthakur, S., Lee, H., Kim, S., Wang, B.-C., & Buck, M. (2014). Binding and function of phosphotyrosines of the Ephrin A2 (EphA2) receptor using synthetic sterile α motif (SAM) domains. *The Journal of Biological Chemistry*, 289(28), 19694–19703.

Bourgeron, T. (2009). A synaptic trek to autism. *Current Opinion in Neurobiology*, 19(2), 231–234.

Brigidi, G. S., & Bamji, S. X. (2011). Cadherin-catenin adhesion complexes at the synapse. *Current Opinion in Neurobiology*, 21(2), 208–214.

Brookes, E., Cao W., and Demeler D. (2009). A Two-Dimensional Spectrum Analysis for Sedimentation Velocity Experiments of Mixtures with Heterogeneity in Molecular Weight and Shape. *European Biophysics Journal* 39 (3). Springer-Verlag: 405–14.

Burack, W. R., Shaw, A. S. (2000). Signal transduction: hanging on a scaffold. *Current Opinion in Cell Biology*, 12(2), 211–216.

Cavanagh, J., Fairbrother, W. J., Arthur G Palmer, I., Skelton, N. J., & Rance, M. (2010). *Protein NMR Spectroscopy*. New York: Academic Press.

Calderwood, D. A., Fujioka, Y., de Pereda, J. M., García-Alvarez, B., Nakamoto, T., Margolis, B., et al. (2003). Integrin beta cytoplasmic domain interactions with phosphotyrosine-binding domains: a structural prototype for diversity in integrin signaling. *Proceedings of the National Academy of Sciences*, 100(5), 2272–2277.

Cao, W., & Demeler, B. (2005). Modeling Analytical Ultracentrifugation Experiments with an Adaptive Space-Time Finite Element Solution of the Lamm Equation. *Biophysical Journal*, 89(3), 1589–1602.

Cao, X., Sudhof, T. C. (2001). A transcriptively active complex of APP with Fe65 and histone acetyltransferase Tip60. *Science*, 293(5527), 115–120.

Cao, X., Sudhof, T. C. (2004). Dissection of amyloid-beta precursor protein-dependent transcriptional transactivation. *Journal of Biological Chemistry*, 279(23), 24601–24611.

Caruana, G. (2002). Genetic studies define MAGUK proteins as regulators of epithelial cell polarity. *The International Journal of Developmental Biology*, 46(4), 511–518.

Cesareni, G., Gimona, M., Sudo, M., & Yaffe, M. (Eds.). (2005). Modular protein domains. *Weinheim*: Wiley-VCH Verlab GmbH & Co.

Chagnon, M. J., Uetani, N., & Tremblay, M. L. (2004). Functional significance of the LAR receptor protein tyrosine phosphatase family in development and diseases. *Biochemistry and Cell Biology = Biochimie Et Biologie Cellulaire*, 82(6), 664–675.

Chen, K., and Featherstone, D.E. (2011). Pre and Postsynaptic Roles for Drosophila CASK. *Molecular and Cellular Neuroscience* 48 (2): 171–82.

Chen, V. B., Arendall, W. B., III, Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., et al. (2009). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*, 66(1), 12–21.

Clandinin, T. R., Lee, C. H., Herman, T., Lee, R. C., Yang, A. Y., Ovasapyan, S., & Zipursky, S. L. (2001). Drosophila LAR regulates R1-R6 and R7 target specificity in the visual system. *Neuron*, 32(2), 237–248.

Cohen, G. B., Ren, R., & Baltimore, D. (1995). Modular binding domains in signal transduction proteins. *Cell*, 80(2), 237–248.

Coles, C. H., Jones, E. Y., & Aricescu, A. R. (2015). Extracellular regulation of type IIa receptor protein tyrosine phosphatases: mechanistic insights from structural analyses. *Seminars in Cell & Developmental Biology*, 37, 98–107.

Coles, C. H., Mitakidis, N., Zhang, P., Elegheert, J., Lu, W., Stoker, A. W., et al. (2014). Structural basis for extracellular cis and trans RPTPσ signal competition in synaptogenesis. *Nature Communications*, 5, 5209.

Corvin, A. P. (2010). Neuronal Cell Adhesion Genes: Key Players in Risk for Schizophrenia, Bipolar Disorder and Other Neurodevelopmental Brain Disorders? *Cell Adhesion & Migration* 4(4): 511–14.

Couturier, L., Mazouni, K., & Schweisguth, F. (2013). Inhibition of Notch recycling by Numb: relevance and mechanism(s). *Cell Cycle* (Georgetown, Tex.), 12(11), 1647–1648.

Dahanukar, A., Walker, J. A., & Wharton, R. P. (1999). Smaug, a novel RNA-binding protein that operates a translational switch in Drosophila. *Molecular Cell*, 4(2), 209–218.

Dai, Y., Taru, H., Deken, S. L., Grill, B., Ackley, B., Nonet, M. L., & Jin, Y. (2006). SYD-2 Liprin-alpha organizes presynaptic active zone formation through ELKS. *Nature Neuroscience*, 9(12), 1479–1487.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., et al. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, 6(3), 277–293.

Demeler, B., and Brookes, E. (2007). Monte Carlo Analysis of Sedimentation Experiments. *Colloid and Polymer Science* 286 (2). Springer-Verlag: 129–37.

Demeler, B., and van Holde, K. E. (2004). Sedimentation Velocity Analysis of Highly Heterogeneous Systems. *Analytical Biochemistry* 335 (2): 279–88.

Demeler, B., Brookes, E., Wang, R., Schirf, V., & Kim, C. A. (2010). Characterization of Reversible Associations by Sedimentation Velocity with UltraScan. *Macromolecular Bioscience*, 10(7), 775–782.

Demeler, B. (2001). Methods for the Design and Analysis of Sedimentation Velocity and Sedimentation Equilibrium Experiments with Proteins (pp. 7.13.1–7.13.24). *Hoboken*, NJ, USA: John Wiley & Sons, Inc.

Desai, C. J., Gindhart, J. G., Goldstein, L. S., & Zinn, K. (1996). Receptor tyrosine phosphatases are required for motor axon guidance in the Drosophila embryo. *Cell*, 84(4), 599–609.

Dhalluin, C., Yan, K. S., Plotnikova, O., Lee, K. W., Zeng, L., Kuti, M., et al. (2000). Structural basis of SNT PTB domain interactions with distinct neurotrophic receptors. *Molecular Cell*, 6(4), 921–929.

Dho, S. E., French, M. B., Woods, S. A., & McGlade, C. J. (1999). Characterization of four mammalian numb protein isoforms. Identification of cytoplasmic and membrane-associated variants of the phosphotyrosine binding domain. *Journal of Biological Chemistry*, 274(46), 33097–33104.

Donaldson, L. W., & Kwan, J. J. (2016). Structural investigation of an atypical binding cleft in the CASKIN2 SH3 domain. *BMC Structural Biology* (2016) 16:14

Dosemeci, A., Toy, D., Reese, T. S., & Tao-Cheng, J.-H. (2015). AIDA-1 Moves out of the Postsynaptic Density Core under Excitatory Conditions. *PLoS ONE*, 10(9), e0137216.

Dudek, S. (2007). Transcriptional Regulation by Neuronal Activity. New York ; London : *Springer Science & Business Media.*

Dumanis, S. B., Chamberlain, K. A., Jin Sohn, Y., Jin Lee, Y., Guénette, S. Y., Suzuki, T., et al. (2012). FE65 as a link between VLDLR and APP to regulate their trafficking and processing. *Molecular Neurodegeneration*, 7(1), 9.

Dunah, A. W., Hueske, E., Wyszynski, M., Hoogenraad, C. C., Jaworski, J., Pak, D. T., et al. (2005). LAR receptor protein tyrosine phosphatases in the development and maintenance of excitatory synapses. *Nature Neuroscience*, 8(4), 458–467.

Duyckaerts, C., Delatour, B., & Potier, M.-C. (2009). Classification and basic pathology of Alzheimer disease. *Acta Neuropathologica*, 118(1), 5–36.

Eck, M. J., Dhe-Paganon, S., Trüb, T., Nolte, R. T., & Shoelson, S. E. (1996). Structure of the IRS-1 PTB domain bound to the juxtamembrane region of the insulin receptor. *Cell*, 85(5), 695–705.

Emaduddin, M., Edelmann, M. J., Kessler, B. M., & Feller, S. M. (2008). Odin (ANKS1A) is a Src family kinase target in colorectal cancer cells. *Cell Communication and Signaling,* 6(1), 7.

Emsley P., and Cowtan K. (2004). Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica Section D Biological Crystallography* 60 (12): 2126–32.

Ernst, A., Sazinsky, S. L., Hui, S., Currell, B., Dharsee, M., Seshagiri, S., et al. (2009). Rapid Evolution of Functional Complexity in a Domain Family. *Science Signaling*, 2(87), ra50–ra50.

Farooq, A., Plotnikova, O., Zeng, L., & Zhou, M. M. (1999). Phosphotyrosine binding domains of Shc and insulin receptor substrate 1 recognize the NPXpY motif in a thermodynamically distinct manner. *Journal of Biological Chemistry*, 274(10), 6114–6121.

Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., et al. (1994). Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry*, 33(19), 5984–6003.

Ferrell, J. E. (2000). What Do Scaffold Proteins Really Do? *Science Signaling*, 2000(52), pe1–pe1.

Fiedler, M., Mendoza-Topaz, C., Rutherford, T. J., Mieszczanek, J., & Bienz, M. (2011). Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating β-catenin. *Proceedings of the National Academy of Sciences*, 108(5), 1937–1942.

Fogh, R. H., Fogh, R. H., Vranken, W. F., Vranken, W. F., Boucher, W., Boucher, W., et al. (2006). A nomenclature and data model to describe NMR experiments. *Journal of Biomolecular NMR*, 36(3), 147–155.

Forman-Kay, J. D., & Pawson, T. (1999). Diversity in protein recognition by PTB domains. *Current Opinion in Structural Biology*, 9(6), 690–695.

Fox, A. N., & Zinn, K. (2005). The heparan sulfate proteoglycan syndecan is an in vivo ligand for the Drosophila LAR receptor tyrosine phosphatase. *Current Biology*, 15(19), 1701–1711.

Frank, R. (2002). The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications. *Journal of Immunological Methods*, 267(1), 13–26.

Fromer, M., Pocklington, A. J., Kavanagh, D. H., Williams, H. J., Dwyer, S., Gormley, P., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature*, 506(7487), 179–184.

Fry, E. J., Chagnon, M. J., López-Vales, R., Tremblay, M. L., & David, S. (2010). Corticospinal tract regeneration after spinal cord injury in receptor protein tyrosine phosphatase sigma deficient mice. *Glia*, 58(4), 423–433.

Galvan, V., Gorostiza, O. F., Banwait, S., Ataie, M., Logvinova, A. V., Sitaraman, S., et al. (2006). Reversal of Alzheimer's-like pathology and behavior in human APP transgenic mice by mutation of Asp664. *Proceedings of the National Academy of Sciences*, 103(18), 7130–7135.

García-Alvarez, B., de Pereda, J. M., Calderwood, D. A., Ulmer, T. S., Critchley, D., Campbell, I. D., et al. (2003). Structural determinants of integrin recognition by talin. *Molecular Cell*, 11(1), 49–58.

Garrington, T. P., & Johnson, G. L. (1999). Organization and regulation of mitogen-activated protein kinase signaling pathways. *Current Opinion in Cell Biology*, 11(2), 211–218.

Gauthier, J., Champagne, N., Lafrenière, R. G., Xiong, L., Spiegelman, D., Brustein, E., et al. (2010). De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America,* 107(17), 7863–7868.

Ghersi, E., Noviello, C., & D'Adamio, L. (2004a). Amyloid-beta protein precursor (AbetaPP) intracellular domain-associated protein-1 proteins bind to AbetaPP and modulate its processing in an isoform-specific manner. *Journal of Biological Chemistry*, 279(47), 49105–49112.

Ghersi, E., Vito, P., Lopez, P., Abdallah, M., & D'Adamio, L. (2004b). The intracellular localization of amyloid beta protein precursor (AbetaPP) intracellular domain associated protein-1 (AIDA-1) is regulated by AbetaPP and alternative splicing. *Journal of Alzheimer's Disease: JAD*, 6(1), 67–78.

Goksoy, E., Ma, Y.-Q., Wang, X., Kong, X., Perera, D., Plow, E. F., & Qin, J. (2008). Structural basis for the autoinhibition of talin in regulating integrin activation. *Molecular Cell*, 31(1), 124–133.

Gorbet, G. E., Pearson, J. Z., Demeler, A. K., Cölfen, H., & Demeler, B. (2015). Next-Generation AUC. *Analytical Ultracentrifugation* (Vol. 562, pp. 27–47). Elsevier.

Green, J. B., Gardner, C. D., Wharton, R. P., & Aggarwal, A. K. (2003). RNA recognition via the SAM domain of Smaug. *Molecular Cell*, 11(6), 1537–1548.

Grimshaw, S. J., Mott, H. R., Stott, K. M., Nielsen, P. R., Evetts, K. A., Hopkins, L. J. (2004). Structure of the sterile alpha motif (SAM) domain of the Saccharomyces cerevisiae mitogen-activated protein kinase pathway-modulating protein STE50 and analysis of its interaction with the STE11 SAM. *Journal of Biological Chemistry*, 279(3), 2192–2201.

Grochulski, P., Fodje, M., Labiuk, S., Gorin, J., Janzen, K., & Berg, R. (2012). Canadian macromolecular crystallography facility: a suite of fully automated beamlines. *Journal of Structural and Functional Genomics*, 13(2), 49–55.

Guerois, R., Nielsen, J. E., Serrano, L. (2002). Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations. *Journal of Molecular Biology*, 320(2), 369–387.

Gundelfinger, E. D., Boeckers, T. M., Baron, M. K., & Bowie, J. U. (2006). A role for zinc in postsynaptic density asSAMbly and plasticity? *Trends in Biochemical Sciences*, 31(7), 366–373.

Hamid, R., Kilger, E., Willem, M., Vassallo, N., Kostka, M., Bornhövd, C., et al. (2007). Amyloid precursor protein intracellular domain modulates cellular calcium homeostasis and ATP content. *Journal of Neurochemistry*, 102(4), 1264–1275.

Han, D. C., Shen, T.-L., Miao, H., Wang, B., & Guan, J.-L. (2002). EphB1 associates with Grb7 and regulates cell migration. *Journal of Biological Chemistry*, 277(47), 45655–45661.

Han, K. A., Jeon, S., Um, J. W., & Ko, J. (2016a). Emergent Synapse Organizers: LAR-RPTPs and Their Companions. *International Review of Cell and Molecular Biology*, 324, 39–65.

Han, S.-H., Park, J.-C., & Mook-Jung, I. (2016b). Amyloid β-interacting partners in Alzheimer's disease: From accomplices to possible therapeutic targets. *Progress in Neurobiology*, 137, 17–38.

Harada, B. T., Knight, M. J., Imai, S.-I., Qiao, F., Ramachander, R., Sawaya, M. R., et al. (2008). Regulation of Enzyme Localization by Polymerization: Polymer Formation by the SAM Domain of Diacylglycerol Kinase δ1. *Structure* (London, England : 1993), 16(3), 380–387.

Hayashi, M. K., Tang, C., Verpelli, C., Narayanan, R., Stearns, M. H., Xu, R.-M., et al. (2009). The postsynaptic density proteins Homer and Shank form a polymeric network structure. *Cell*, 137(1), 159–171.

Hoch, Jeffrey C, and Alan S Stern. 1996. NMR Data Processing. *Wiley*.

Hofmeyer, K., & Treisman, J. E. (2009). The receptor protein tyrosine phosphatase LAR promotes R7 photoreceptor axon targeting by a phosphatase-independent signaling mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46), 19399–19404.

Howell, B. W., Lanier, L. M., Frank, R., Gertler, F. B., & Cooper, J. A. (1999). The disabled 1 phosphotyrosine-binding domain binds to the internalization signals of transmembrane glycoproteins and to phospholipids. *Molecular and Cellular Biology*, 19(7), 5179–5188.

Hsueh, Y.-P. (2006). The Role of the MAGUK Protein CASK in Neural Development and Synaptic Function. *Current Medicinal Chemistry*, 13(16), 1915–1927.

Hu, T., Shi, G., Larose, L., Rivera, G. M., Mayer, B. J., & Zhou, R. (2009). Regulation of process retraction and cell migration by EphA3 is mediated by the adaptor protein Nck1. *Biochemistry*, 48(27), 6369–6378.

Hyberts, S. G., Milbradt, A. G., Wagner, A. B., Arthanari, H., & Wagner, G. (2012). Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *Journal of Biomolecular NMR*, 52(4), 315–327.

Iasevoli, F., Tomasetti, C., & de Bartolomeis, A. (2013). Scaffolding proteins of the post-synaptic density contribute to synaptic plasticity by regulating receptor localization and distribution: relevance for neuropsychiatric diseases. *Neurochemical Research*, 38(1), 1–22.

Jacob, A. L., Jordan, B. A., & Weinberg, R. J. (2010). Organization of amyloid-beta protein precursor intracellular domain-associated protein-1 in the rat brain. *The Journal of Comparative Neurology*, 518(16), 3221–3236.

Jády, B. E., Richard, P., Bertrand, E., & Kiss, T. (2006). Cell cycle-dependent recruitment of telomerase RNA and Cajal bodies to human telomeres. *Molecular Biology of the Cell*, 17(2), 944–954.

Jeyifous, O., Waites, C. L., Specht, C. G., Fujisawa, S., Schubert, M., Lin, E. I., et al. (2009). SAP97 and CASK mediate sorting of NMDA receptors through a previously unknown secretory pathway. *Nature Neuroscience*, 12(8), 1011–1019.

Jiang, Y.-H., & Ehlers, M. D. (2013). Modeling Autism by SHANK Gene Mutations in Mice. *Neuron*, 78(1), 8–27.

Jin, J., & Pawson, T. (2012). Modular evolution of phosphorylation-based signalling systems. Philosophical Transactions of the Royal Society of London. Series B, *Biological Sciences*, 367(1602), 2540–2555.

Jin, J., Xie, X., Chen, C., Park, J. G., Stark, C., James, D. A., et al. (2009). Eukaryotic protein domains as functional units of cellular evolution. *Science Signaling*, 2(98), ra76–ra76.

Johnson, K. G., Tenney, A. P., Ghose, A., Duckworth, A. M., Higashi, M. E., Parfitt, K., et al. (2006). The HSPGs Syndecan and Dallylike bind the receptor phosphatase LAR and exert distinct effects on synaptic development. *Neuron*, 49(4), 517–531.

Johnson, P. E., & Donaldson, L. W. (2006). RNA recognition by the Vts1p SAM domain. *Nature Structural & Molecular Biology*, 13(2), 177–178.

Jordan, B. A. (2004). Identification and Verification of Novel Rodent Postsynaptic Density Proteins. *Molecular & Cellular Proteomics*, 3(9), 857–871.

Jordan, B. A., & Kreutz, M. R. (2009). Nucleocytoplasmic protein shuttling: the direct route in synapse-to-nucleus signaling. *Trends in Neurosciences*, 32(7), 392–401.

Jordan, B. A., Fernholz, B. D., Khatri, L., & Ziff, E. B. (2007). Activity-dependent AIDA-1 nuclear signaling regulates nucleolar numbers and protein synthesis in neurons. *Nature Neuroscience*, 10(4), 427–435.

Jordan, B. A., Fernholz, B. D., Boussac, M., Xu, C., Grigorean, G., Ziff, E. B., & Neubert, T. A. (2004). Identification and verification of novel rodent postsynaptic density proteins. *Molecular & Cellular Proteomics*, 3(9), 857–871.

Kabsch, W., Kabsch, W., & IUCr. (2010). XDS. *Acta Crystallographica Section D Biological Crystallography*, 66(2), 125–132.

Kantardjieff, K. A., & Rupp, B. (2003). Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Science : a Publication of the Protein Society*, 12(9), 1865–1871.

Kay, L. E., Torchia D. A., and Bax A. (1989). Backbone Dynamics of Proteins as Studied by 15N Inverse Detected Heteronuclear NMR Spectroscopy: Application to Staphylococcal Nuclease. *Biochemistry* 28 (23): 8972–79.

Kay, L. E., Ikura, M., Tschudin, R., & Bax, A. (2011). Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. 1990. *Journal of Magnetic Resonance* (San Diego, Calif. : 1997), 213(2), 423–441.

Kiel, C., Beltrao, P., & Serrano, L. (2008). Analyzing protein interaction networks using structural information. *Annu Rev Biochem*, 77(1), 415–441.

Kim, C. (2003). SAM domains: uniform structure, diversity of function. *Trends in Biochemical Sciences,* 28(12), 625–628.

Kim, C. A., & Kim, C. A. (2005). Structural Organization of a Sex-comb-on-midleg/Polyhomeotic Copolymer. *Journal of Biological Chemistry*, 280(30), 27769–27775.

Kim, C. A., Gingery, M., Pilpa, R. M., & Bowie, J. U. (2002). The SAM domain of polyhomeotic forms a helical polymer. *Nature Structural & Molecular Biology*, 9(6), 453–457.

Kim, C. A., Phillips, M. L., Kim, W., Gingery, M., Tran, H. H., Robinson, M. A., et al. (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *The EMBO Journal*, 20(15), 4173–4182.

Kim, E., & Sheng, M. (2004). PDZ domain proteins of synapses. *Nature Reviews Neuroscience*, 5(10), 771–781.

Kim, J., Lee, H., Kim, Y., Yoo, S., Park, E., & Park, S. (2010). The SAM domains of Anks family proteins are critically involved in modulating the degradation of EphA receptors. *Molecular and Cellular Biology*, 30(7), 1582–1592.

Kim, S.-E., Huang, H., Zhao, M., Zhang, X., Zhang, A., Semonov, M. V., et al. (2013). Wnt stabilization of β-catenin reveals principles for morphogen receptor-scaffold assemblies. *Science*, 340(6134), 867–870.

King, G. D., & Scott Turner, R. (2004). Adaptor protein interactions: modulators of amyloid precursor protein metabolism and Alzheimer's disease risk? *Experimental Neurology*, 185(2), 208–219.

Knight, M. J., Leettola, C., Gingery, M., Li, H., & Bowie, J. U. (2011). A human sterile alpha motif domain polymerizome. *Protein Science*, 20(10), 1697–1706.

Knight, M. J., Joubert, M. K., Plotkowski, M. L., Kropat, J., Gingery, M., Sakane, F., et al. (2010). Zinc Binding Drives Sheet Formation by the SAM Domain of Diacylglycerol Kinase δ. *Biochemistry*, 49(44), 9667–9676.

Ko, J., Kim, S., Valtschanoff, J. G., Shin, H., Lee, J.-R., Sheng, M., et al. (2003). Interaction between liprin-alpha and GIT1 is required for AMPA receptor targeting. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 23(5), 1667–1677.

Koradi, R., Billeter, M., & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. Journal of Molecular Graphics, 14(1), 51–5–29–32.

Krauss, G. (2008). Biochemistry of Signal Transduction and Regulation. (4th edition). *Weinheim*: Wiley-VCH Verlab GmbH & Co.

Krissinel, E, and K Henrick. 2004. "Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions.." *Acta Crystallographica Section D Biological Crystallography* 60 (Pt 12 Pt 1). International Union of Crystallography: 2256–68.

Krishna, N. R., & Berliner, L. (1999). Modern Techniques in Protein NMR. *Springer Science & Business Media.*

Krueger, N. X., Van Vactor, D., Wan, H. I., Gelbart, W. M., Goodman, C. S., & Saito, H. (1996). The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila. *Cell*, 84(4), 611–622.

Kurabi, A., Brener, S., Mobli, M., Kwan, J. J., & Donaldson, L. W. (2009). A Nuclear Localization Signal at the SAM–SAM Domain Interface of AIDA-1 Suggests a Requirement for Domain Uncoupling Prior to Nuclear Import. *Journal of Molecular Biology*, 392(5), 1168–1177.

Kwan, J. J., Warner, N., Pawson, T., & Donaldson, L. W. (2004). The Solution Structure of the S.cerevisiae Ste11 MAPKKK SAM Domain and its Partnership with Ste50. *Journal of Molecular Biology*, 342(2), 681–693.

Kyba, M., & Brock, H. W. (1998). The Drosophila polycomb group protein Psc contacts ph and Pc through specific conserved domains. *Molecular and Cellular Biology*, 18(5), 2712–2720.

LaConte, L., & Mukherjee, K. (2013). Structural constraints and functional divergences in CASK evolution. *Biochemical Society Transactions*, 41, 1017–1022.

LaConte, L. E. W., Chavan, V., Liang, C., Willis, J., Schönhense, E.-M., Schoch, S., & Mukherjee, K. (2016). CASK stabilizes neurexin and links it to liprin-α in a neuronal activity-dependent manner. *Cellular and Molecular Life Sciences*, 73(18), 3599–3621.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M., & IUCr. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291.

Lee, H. J., Hota, P. K., Chugha, P., Guo, H., Miao, H., Zhang, L., et al. (2012). NMR Structure of a Heterodimeric SAM:SAM Complex: Characterization and Manipulation of EphA2 Binding Reveal New Cellular Functions of SHIP2. *Structure* (London, England : 1993), 20(1), 41–55.

Lee, H., & Bennett, A. M. (2013). Receptor protein tyrosine phosphatase-receptor tyrosine kinase substrate screen identifies EphA2 as a target for LAR in cell migration. *Molecular and Cellular Biology*, 33(7), 1430–1441.

Leettola, C. N., Knight, M. J., Cascio, D., Hoffman, S., & Bowie, J. U. (2014). Characterization of the SAM domain of the PKD-related protein ANKS6 and its interaction with ANKS3. *BMC Structural Biology*, 14(1), 17.

Leone, M., Cellitti, J., & 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(48), 12721–12728.

Li, L., Su, Y., Zhao, C., Zhao, H., Liu, G., Wang, J., & Xu, Q. (2006). The role of Ret receptor tyrosine kinase in dopaminergic neuron development. *Neuroscience*, 142(2), 391–400.

Li, S.-C., Zwahlen, C., Vincent, S. J. F., McGlade, C. J., Kay, L. E., Pawson, T., & Forman-Kay, J. D. (1998). Structure of a Numb PTB domain|[ndash]|peptide complex suggests a basis for diverse binding specificity. *Nature Structural &*

Molecular Biology, 5(12)

Long, J., Wei, Z., Feng, W., Yu, C., Zhao, Y.-X., & Zhang, M. (2008). Supramodular nature of GRIP1 revealed by the structure of its PDZ12 tandem in complex with the carboxyl tail of Fras1. *Journal of Molecular Biology*, 375(5), 1457–1468.

Lorenzo, A., Yuan, M., Zhang, Z., Paganetti, P. A., Sturchler-Pierrat, C., Staufenbiel, M., et al. (2000). Amyloid beta interacts with the amyloid precursor protein: a potential toxic mechanism in Alzheimer's disease. *Nature Neuroscience*, 3(5), 460–464.

Lowenthal, M. S., Markey, S. P., & Dosemeci, A. (2015). Quantitative mass spectrometry measurements reveal stoichiometry of principal postsynaptic density proteins. *Journal of Proteome Research*, 14(6), 2528–2538.

Madrzak, J., Fiedler, M., Johnson, C. M., Ewan, R., Knebel, A., Bienz, M., & Chin, J. W. (2015). Ubiquitination of the Dishevelled DIX domain blocks its head-to-tail polymerization. *Nature Communications*, 6, 6718.

Manders, E., & Tyberghein, J. (1993). The effects of ventilation tube placement on hearing, speech, language, cognition and behaviour. *Acta Oto-Rhino-Laryngologica Belgica*, 47(1), 27–32.

Mao, X., Kikani, C. K., Riojas, R. A., Langlais, P., Wang, L., Ramos, F. J., et al. (2006). APPL1 binds to adiponectin receptors and mediates adiponectin signalling and function. *Nature Cell Biology*, 8(5), 516–523.

Margolis, B., Margolis, B., Borg, J. P., Borg, J. P., Straight, S., Straight, S., et al. (1999). The function of PTB domain proteins. *Kidney International*, 56(4), 1230–1237.

Marks, N., & Berg, M. J. (2010). BACE and gamma-secretase characterization and their sorting as therapeutic targets to reduce amyloidogenesis. *Neurochemical Research*, 35(2), 181–210.

Martel, V., Racaud-Sultan, C., Dupe, S., Marie, C., Paulhe, F., Galmiche, A., et al. (2001). Conformation, localization, and integrin binding of talin depend on its interaction with phosphoinositides. *Journal of Biological Chemistry*, 276(24), 21217–21227.

Matos, M. F., Xu, Y., Dulubova, I., Otwinowski, Z., Richardson, J. M., Tomchick, D. R., et al. (2012). Autoinhibition of Mint1 adaptor protein regulates amyloid precursor protein binding and processing. *Proceedings of the National Academy of Sciences of the United States of America*, 109(10), 3802–3807.

Maurel-Zaffran, C., Suzuki, T., Gahmon, G., Treisman, J. E., & Dickson, B. J. (2001). Cell-Autonomous and -Nonautonomous Functions of LAR in R7 Photoreceptor Axon Targeting. *Neuron*, 32(2), 225–235.

McIntosh, L. P., & Dahlquist, F. W. (1990). Biosynthetic incorporation of 15N and 13C for assignment and interpretation of nuclear magnetic resonance spectra of proteins. *Quarterly Reviews of Biophysics*, 23(1), 1–38.

McLoughlin, D. M., McLoughlin, D. M., Miller, C. C. J., & Miller, C. C. J. (2008). The FE65 proteins and Alzheimer's disease. *Journal of Neuroscience Research*, 86(4), 744–754.

Meister, M., Tomasovic, A., Banning, A., & Tikkanen, R. (2013). Mitogen-Activated Protein (MAP) Kinase Scaffolding Proteins: A Recount. *International Journal of Molecular Sciences*, 14(3), 4854–4884.

Mercurio, F. A., Marasco, D., Pirone, L., Pedone, E. M., Pellecchia, M., & Leone, M. (2012). Solution Structure of the First Sam Domain of Odin and Binding Studies with the EphA2 Receptor. *Biochemistry*, 51(10), 2136–2145.

Mertens, A. E. E., Pegtel, D. M., & Collard, J. G. (2006). Tiam1 takes PARt in cell polarity. *Trends in Cell Biology*, 16(6), 308–316.

Middleton, F. A., Carrierfenster, K., Mooney, S. M., & Youngentob, S. L. (2009). Gestational ethanol exposure alters the behavioral response to ethanol odor and the expression of neurotransmission genes in the olfactory bulb of adolescent rats. *Brain Research*, 1252, 105–116.

Minard, M. E., Kim, L.-S., Price, J. E., & Gallick, G. E. (2004). The role of the guanine nucleotide exchange factor Tiam1 in cellular migration, invasion, adhesion and tumor progression. *Breast Cancer Research and Treatment*, 84(1), 21–32.

Minopoli, G., De Candia, P., Bonetti, A., Faraonio, R., Zambrano, N., & Russo, T. (2001). The beta-amyloid precursor protein functions as a cytosolic anchoring site that prevents Fe65 nuclear translocation. *Journal of Biological Chemistry*, 276(9), 6545–6550.

Mishra, S. K., Watkins, S. C., & Traub, L. M. (2002). The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery. *Proceedings of the National Academy of Sciences*, 99(25), 16099–16104.

Mota, S. I., Ferreira, I. L., & Rego, A. C. (2014). Dysfunctional synapse in Alzheimer's disease - A focus on NMDA receptors. *Neuropharmacology*, 76 Pt A, 16–26.

Multhaup, G., Huber, O., Buée, L., & Galas, M.-C. (2015). Amyloid Precursor Protein (APP) Metabolites APP Intracellular Fragment (AICD), Aβ42, and Tau in Nuclear Roles. *The Journal of Biological Chemistry*, 290(39), 23515–23522.

Mulvihill, M. M., Guttman, M., & Komives, E. A. (2011). Protein interactions among Fe65, the low-density lipoprotein receptor-related protein, and the amyloid precursor protein. *Biochemistry*, 50(28), 6208–6216.

Naisbitt, S., Kim, E., Tu, J. C., Xiao, B., Sala, C., Valtschanoff, J., et al. (1999). Shank, a Novel Family of Postsynaptic Density Proteins that Binds to the NMDA Receptor/PSD-95/GKAP Complex and Cortactin. *Neuron*, 23(3), 569–582.

Nam, H. J., Nam, H. J., Poy, F., Krueger, N. X., Krueger, N. X., Saito, H., et al. (1999). Crystal structure of the tandem phosphatase domains of RPTP LAR. *Cell*, 97(4), 449–457.

Oberstrass, F. C., Lee, A., Stefl, R., Janis, M., Chanfreau, G., & Allain, F. H. T. (2006). Shape-specific recognition in the structure of the Vts1p SAM domain with RNA. *Nature Structural & Molecular Biology,* 13(2), 160–167.

Ojeh, N., Pekovic, V., Jahoda, C., & Määttä, A. (2008). The MAGUK-family protein CASK is targeted to nuclei of the basal epidermis and controls keratinocyte proliferation. *Journal of Cell Science*, 121(16), 2705–2717.

Okamoto, M., & Sudhof, T. C. (1997). Mints, Munc18-interacting proteins in synaptic vesicle exocytosis. *Journal of Biological Chemistry*, 272(50), 31459–31464.

Olmsted, J. B., Carlson, K., Klebe, R., Ruddle, F., & Rosenbaum, J. (1970). Isolation of microtubule protein from cultured mouse neuroblastoma cells. *Proceedings of the National Academy of Sciences,* 65(1), 129–136.

Otten, R., Chu, B., Krewulak, K. D., Vogel, H. J., & Mulder, F. A. A. (2010). Comprehensive and cost-effective NMR spectroscopy of methyl groups in large proteins. *Journal of the American Chemical Society*, 132(9), 2952–2960.

Olsen, O., Moore, K. A., Fukata, M., Kazuta, T., Trinidad, J. C., Kauer, F. W., et al. (2005). Neurotransmitter release regulated by a MALS-liprin-alpha presynaptic complex. *The Journal of Cell Biology*, 170(7), 1127–1134.

Pan, C. Q., Sudol, M., Sheetz, M., & Low, B. C. (2012). Modularity and functional plasticity of scaffold proteins as p(l)acemakers in cell signaling. *Cellular Signalling*, 24(11), 2143–2165.

Park, S., Lee, H., & Park, S. (2015). In Vivo Expression of the PTB-deleted Odin Mutant Results in Hydrocephalus. *Molecules and Cells*, 38(5), 426–431.

Passer, B., Pellegrini, L., Russo, C., Siegel, R. M., Lenardo, M. J., Schettini, G., et al. (2000). Generation of an apoptotic intracellular peptide by gamma-secretase cleavage of Alzheimer's amyloid beta protein precursor. *Journal of Alzheimer's Disease : JAD*, 2(3-4), 289–301.

Patel, M. R., Lehrman, E. K., Poon, V. Y., Crump, J. G., Zhen, M., Bargmann, C. I., & Shen, K. (2006). Hierarchical assembly of presynaptic components in defined C. elegans synapses. *Nature Neuroscience*, 9(12), 1488–1498.

Pawson, T., & Nash, P. (2000). Protein-protein interactions define specificity in signal transduction. *Genes & Development*, 14(9), 1027–1047.

Pawson, T., & Scott, J. D. (1997). Signaling Through Scaffold, Anchoring, and Adaptor Proteins. *Science*, 278(5346), 2075–2080.

Pesti, S., Balázs, A., Udupa, R., Szabó, B., Fekete, A., Bőgel, G., & Buday, L. (2012). Complex formation of EphB1/Nck/Caskin1 leads to tyrosine phosphorylation and structural changes of the Caskin1 SH3 domain. *Cell Communication and Signaling*, 10(1), 36.

Peterson, A. J., Kyba, M., Bornemann, D., Morgan, K., Brock, H. W., & Simon, J. (1997). A domain shared by the Polycomb group proteins Scm and ph mediates heterotypic and homotypic interactions. *Molecular and Cellular Biology*, 17(11), 6683–6692.

Peterson, A. J., Mallin, D. R., Francis, N. J., Ketel, C. S., Stamm, J., Voeller, R. K., et al. (2004). Requirement for sex comb on midleg protein interactions in Drosophila polycomb group repression. *Genetics*, 167(3), 1225–1239.

Petsko, G. A., & Ringe, D. (2004). Protein Structure and Function. *New Science Press*.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., Vorstman, J. A. S., Thompson, A., Regan, R., Pilorge, M., Pellecchia, G., Pagnamenta, A. T., Oliveira, B., Marshall, C. R., Magalhaes, T. R., Lowe, J. K., Howe, J. L., Griswold, A. J., Gilbert, J., Duketis, E., Dombroski, B. A., De Jonge, M. V., Cuccaro, M., Crawford, E. L., Correia, C. T., Conroy, J., Conceição, I. C., Chiocchetti, A. G., Casey, J. P., Cai, G., Cabrol, C., Bolshakova, N., Bacchelli, E., Anney, R., Gallinger, S., Cotterchio, M., Casey, G., Zwaigenbaum, L., Wittemeyer, K., Wing, K., Wallace, S., van Engeland, H., Tryfon, A., Thomson, S., Soorya, L., Rogé, B., Roberts, W., Poustka, F., Mouga, S., Minshew, N., McInnes, L. A., McGrew, S. G., Lord, C.,

Leboyer, M., Le Couteur, A. S., Kolevzon, A., Jiménez González, P., Jacob, S., Holt, R., Guter, S., Green, J., Green, A., Gillberg, C., Fernandez, B. A., Duque, F., Delorme, R., Dawson, G., Chaste, P., Café, C., Brennan, S., Bourgeron, T., Bolton, P. F., Bölte, S., Bernier, R., Baird, G., Bailey, A. J., Anagnostou, E., Almeida, J., Wijsman, E. M., Vieland, V. J., Vicente, A. M., Schellenberg, G. D., Pericak-Vance, M., Paterson, A. D., Parr, J. R., Oliveira, G., Nurnberger, J. I., Monaco, A. P., Maestrini, E., Klauck, S. M., Hakonarson, H., Haines, J. L., Geschwind, D. H., Freitag, C. M., Folstein, S. E., Ennis, S., Coon, H., Battaglia, A., Szatmari, P., Sutcliffe, J. S., Hallmayer, J., Gill, M., Cook, E. H., Buxbaum, J. D., Devlin, B., Gallagher, L., Betancur, C., & Scherer, S. W. (2014). Convergence of Genes and Cellular Pathways Dysregulated in Autism Spectrum Disorders. *The American Journal of Human Genetics*, 94(5), 677–694.

Pontes, O., Pontes, O., Pikaard, C. S., & Pikaard, C. S. (2008). siRNA and miRNA processing: new functions for Cajal bodies. *Current Opinion in Genetics & Development*, 18(2), 197–203.

Ponting, C. P. (1995). SAM: a novel motif in yeast sterile and Drosophila polyhomeotic proteins. *Protein Science : a Publication of the Protein Society*, 4(9), 1928–1930.

Pot, D. A., & Dixon, J. E. (1992). A thousand and two protein tyrosine phosphatases. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*, 1136(1), 35–43.

Prochnow, N., Hoffmann, S., Dermietzel, R., & Zoidl, G. (2009). Replacement of a single cysteine in the fourth transmembrane region of zebrafish pannexin1 alters hemichannel gating behavior. *Experimental Brain Research*, 199(3-4), 255–264.

Pulido, R., Serra-Page, C., Tang, M., & Streuli, M. (1995). The LAR/PTPδ/PTPσ Subfamily of Transmembrane Protein- Tyrosine-Phosphatases: Multiple Human LAR, PTPδ, and PTPσ Isoforms are Expressed in a Tissue-Specific Manner and Associate with the LAR-Interacting Protein LIP.1. *PNAS*, 92(25),11686-11690.

Purcell, S. M., Moran, J. L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S. E., Kähler, A., Duncan, L., Stahl, E., Genovese, G., Fernández, E., Collins, M. O., Komiyama, N. H., Choudhary, J. S., Magnusson, P. K. E., Banks, E., Shakir, K., Garimella, K., Fennell, T., DePristo, M., Grant, S. G. N., Haggarty, S. J., Gabriel, S., Scolnick, E. M., Lander, E. S., Hultman, C. M., Sullivan, P. F., McCarroll, S. A., & Sklar, P. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*, 506(7487), 185–190.

Qiao, F., & Bowie, J. U. (2005). The Many Faces of SAM. *Sci STKE*, 2005(286), re7–re7.

Qiao, F., Harada, B., Song, H., Whitelegge, J., Courey, A. J., & Bowie, J. U. (2006). Mae inhibits Pointed- P2 transcriptional activity by blocking its MAPK docking site. *The EMBO Journal*, 25(1), 70–79.

Qiao, F., Song, H., Kim, C. A., Sawaya, M. R., Hunter, J. B., Gingery, M., et al. (2004). Derepression by Depolymerization. *Cell*, 118(2), 163–173.

Rademacher, N., Kunde, S.-A., Kalscheuer, V. M., & Shoichet, S. A. (2013). Synaptic MAGUK multimer formation is mediated by PDZ domains and promoted by ligand binding. *Chemistry & Biology*, 20(8), 1044–1054.

Radzimanowski, J., Simon, B., Sattler, M., Beyreuther, K., Sinning, I., & Wild, K. (2008). Structure of the intracellular domain of the amyloid precursor protein in complex with Fe65-PTB2. *EMBO Reports*, 9(11), 1134–1140.

Ramachander, R., & Bowie, J. U. (2004). SAM domains can utilize similar surfaces for the formation of polymers and closed oligomers. *Journal of Molecular Biology*, 342(5), 1353–1358.

Ramachander, R., Kim, C. A., Phillips, M. L., Mackereth, C. D., Thanos, C. D., McIntosh, L. P., & Bowie, J. U. (2002). Oligomerization-dependent association of the SAM domains from Schizosaccharomyces pombe Byr2 and Ste4. *Journal of Biological Chemistry*, 277(42), 39585–39593.

Raveh, Barak, Raveh, Barak, Nir London, Nir London, Lior Zimmerman, Lior Zimmerman, Ora Schueler-Furman, and Ora Schueler-Furman. 2011. Rosetta FlexPepDock Ab-Initio: Simultaneous Folding, Docking and Refinement of Peptides Onto Their Receptors. *PLoS ONE* 6 (4): e18934.

Ravichandran, K. S. (2001). Signaling via Shc family adapter proteins. *Oncogene*, 20(44), 6322–6330.

Ravichandran, K. S., Zhou, M. M., Pratt, J. C., Harlan, J. E., Walk, S. F., Fesik, S. W., & Burakoff, S. J. (1997). Evidence for a requirement for both phospholipid and phosphotyrosine binding via the Shc phosphotyrosine-binding domain in vivo. *Molecular and Cellular Biology*, 17(9), 5540–5549.

Raychaudhuri, M., & Mukhopadhyay, D. (2007). AICD and its adaptors - in search of new players. *Journal of Alzheimer's Disease : JAD*, 11(3), 343–358.

Roussignol, G., Ango, F., Romorini, S., Tu, J. C., Sala, C., Worley, P. F., et al. (2005). Shank expression is sufficient to induce functional dendritic spine synapses in aspiny neurons. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 25(14), 3560–3570.

Saito, Y., Akiyama, M., Araki, Y., Sumioka, A., Shiono, M., Taru, H., et al. (2011). Intracellular trafficking of the amyloid β-protein precursor (APP) regulated by novel function of X11-like. *PLoS ONE*, 6(7), e22108.

Saito, Y., Sano, Y., Vassar, R., Gandy, S., Nakaya, T., Yamamoto, T., & Suzuki, T. (2008). X11 proteins regulate the translocation of amyloid beta-protein precursor (APP) into detergent-resistant membrane and suppress the amyloidogenic cleavage of APP by beta-site-cleaving enzyme in brain. *Journal of Biological Chemistry*, 283(51), 35763–35771.

Sakuma, M., Sakuma, M., Tanaka, E., Taru, H., Taru, H., Tomita, S., et al. (2009). Phosphorylation of the amino-terminal region of X11L regulates its interaction with APP. *Journal of Neurochemistry*, 109(2), 465–475.

Sakurai, T., Kaneko, K., Okuno, M., Wada, K., Kashiyama, T., Shimizu, H., et al. (2008). Membrane microdomain switching: a regulatory mechanism of amyloid precursor protein processing. *The Journal of Cell Biology*, 183(2), 339–352.

Sayou, C., Nanao, M. H., Jamin, M., Posé, D., Thévenon, E., Grégoire, L., et al. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nature Communications*, 7, 11222.

Schaapveld, R. Q. J., Schepens, J. T. G., Robinson, G. W., Attema, J., Oerlemans, F. T. J. J., Fransen, J. A. M., et al. (1997). Impaired Mammary Gland Development and Function in Mice Lacking LAR Receptor-like Tyrosine Phosphatase Activity. *Developmental Biology*, 188(1), 134–146.

Schaeffer, H. J., & Weber, M. J. (1999). Mitogen-activated protein kinases: specific messages from ubiquitous messengers. *Molecular and Cellular Biology*, 19(4), 2435–2444.

Schoch, S., Castillo, P. E., Jo, T., Mukherjee, K., Geppert, M., Wang, Y., et al. (2002). RIM1alpha forms a protein scaffold for regulating neurotransmitter release at the active zone. *Nature*, 415(6869), 321–326.

Schultz, J., Bork, P., Ponting, C. P., & Hofmann, K. (1997). SAM as a protein interaction domain involved in developmental regulation. *Protein Science*, 6(1), 249–253.

Schultz, J., Milpetz, F., Bork, P., & Ponting, C. P. (1998). SMART, a simple modular architecture research tool: identification of signaling domains. *Proceedings of the National Academy of Sciences*, 95(11), 5857–5864.

Schwarz-Romond, T., Fiedler, M., Shibata, N., Butler, P. J. G., Kikuchi, A., Higuchi, Y., & Bienz, M. (2007). The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. *Nature Structural & Molecular Biology*, 14(6), 484–492.

Selkoe, D. J. (1991). The molecular pathology of Alzheimer's disease. *Neuron*, 6(4), 487–498.

Selkoe, D. J., & Hardy, J. (2016). The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Molecular Medicine*, 8(6), 595–608.

Serra-Pagès, C., Kedersha, N. L., Fazikas, L., Medley, Q., Debant, A., & Streuli, M. (1995). The LAR transmembrane protein tyrosine phosphatase and a coiled-coil LAR-interacting protein co-localize at focal adhesions. *The EMBO Journal*, 14(12), 2827–2838.

Serra-Pagès, C., Medley, Q. G., Tang, M., Hart, A., & Streuli, M. (1998). Liprins, a family of LAR transmembrane protein-tyrosine phosphatase-interacting proteins. *Journal of Biological Chemistry*, 273(25), 15611–15620.

Serra-Pagès, C., Streuli, M., & Medley, Q. G. (2005). Liprin phosphorylation regulates binding to LAR: evidence for liprin autophosphorylation. *Biochemistry*, 44(48), 15715–15724.

Sheng, M., & Hoogenraad, C. C. (2007). The postsynaptic architecture of excitatory synapses: a more quantitative view. *Annu Rev Biochem*, 76(1), 823–847.

Sheng, M., & Kim, E. (2000). The Shank family of scaffold proteins. *Journal of Cell Science* 113, 1851-1856

Shi, N., Zhou, W., Tang, K., Gao, Y., Jin, J., Gao, F., et al. (2002). Expression, crystallization and preliminary X-ray studies of the recombinant PTB domain of human dok-5 protein. *Acta Crystallographica Section D Biological Crystallography*, 58(Pt 12), 2170–2172.

Shin, J., Gu, C., Gu, C., Park, E., & Park, S. (2007). Identification of Phosphotyrosine Binding Domain-Containing Proteins as Novel Downstream Targets of the EphA8 Signaling Function. *Molecular and Cellular Biology*, 27(23), 8113–8126.

Siu, R., Fladd, C., & Rotin, D. (2007). N-cadherin is an in vivo substrate for protein tyrosine phosphatase sigma (PTPsigma) and participates in PTPsigma-mediated inhibition of axon growth. *Molecular and Cellular Biology*, 27(1), 208–219.

Skinner, S. P., Goult, B. T., Fogh, R. H., Boucher, W., Stevens, T. J., Laue, E. D., & Vuister, G. W. (2015). Structure calculation, refinement and validation using CcpNmr Analysis. *Acta Crystallographica Section D Biological Crystallography*,

71(1), 154–161.

Smalla, M., Schmieder, P., Kelly, M., Laak, Ter, A., Krause, G., Ball, L., et al. (1999). Solution structure of the receptor tyrosine kinase EphB2 SAM domain and identification of two distinct homotypic interaction sites. *Protein Science : a Publication of the Protein Society*, 8(10), 1954–1961.

Smirnova, E., Kwan, J. J., Siu, R., Gao, X., Zoidl, G., Demeler, B., et al. (2016). A new mode of SAM domain mediated oligomerization observed in the CASKIN2 neuronal scaffolding protein. *Cell Communication and Signaling*, 1–14.

Smirnova, E., Shanbhag, R., Kurabi, A., Mobli, M., Kwan, J. J., Sheng, Y., & Donaldson, L. W. (2013). Solution structure and amyloid precursor protein (APP) binding of the PTB domain from the AIDA1 postsynaptic signaling scaffolding protein. *PLoS ONE*. 8(6): e65605.

Smith, F. D., & Scott, J. D. (2013). Scaffolding proteins: not such innocent bystanders. *Current Biology : CB*, 23(12), R515–7.

Snyder, M. A., & Gao, W.-J. (2013). NMDA hypofunction as a convergence point for progression and symptoms of schizophrenia. *Frontiers in Cellular Neuroscience*, 7, 31.

Song, H., Nie, M., Qiao, F., Bowie, J. U., & Courey, A. J. (2005). Antagonistic regulation of Yan nuclear export by Mae and Crm1 may increase the stringency of the Ras response. *Genes & Development*, 19(15), 1767–1772.

Spangler, S. A., & Hoogenraad, C. C. (2007). Liprin-alpha proteins: scaffold molecules for synapse maturation. *Biochemical Society Transactions*, 35(Pt 5), 1278–1282.

Stafford, R. L., Ear, J., Knight, M. J., & Bowie, J. U. (2011a). The Molecular Basis of the Caskin1 and Mint1 Interaction with CASK. *Journal of Molecular Biology*, 412(1), 3–13.

Stafford, R. L., Hinde, E., Knight, M. J., Pennella, M. A., Ear, J., Digman, M. A., et al. (2011b). Tandem SAM Domain Structure of Human Caskin1: A Presynaptic, Self-Assembling Scaffold for CASK, *Structure* (London, England : 1993), 19(12), 1826–1836.

Stapleton, D., Balan, I., Pawson, T., & Sicheri, F. (1999). The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nature Structural Biology*, 6(1), 44–49.

Stolt, P. C., Jeon, H., Song, H. K., Herz, J., Eck, M. J., & Blacklow, S. C. (2003). Origins of peptide selectivity and phosphoinositide binding revealed by structures of disabled-1 PTB domain complexes. *Structure/Folding and Design*, 11(5), 569–579.

Stryker, E., & Johnson, K. G. (2007). LAR, liprin α and the regulation of active zone morphogenesis. *Journal of Cell Science*, 120(21), 3723–3728.

Suzuki, T., & Nakaya, T. (2008). Regulation of amyloid beta-protein precursor by phosphorylation and protein interactions. *Journal of Biological Chemistry*, 283(44), 29633–29637.

Tabuchi, K., Biederer, T., Butz, S., & Sudhof, T. C. (2002). CASK participates in alternative tripartite complexes in which Mint 1 competes for binding with caskin 1, a novel CASK-binding protein. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 22(11), 4264–4273.

Takeuchi, H., Matsuda, M., Yamamoto, T., Kanematsu, T., Kikkawa, U., Yagisawa, H., et al. (1998). PTB domain of insulin receptor substrate-1 binds inositol compounds. *Biochemical Journal*, 334 (Pt 1), 211–218.

Taru, H., & Jin, Y. (2011). The Liprin homology domain is essential for the homomeric interaction of SYD-2/Liprin-α protein in presynaptic assembly. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 31(45), 16261–16268.

Thanos, C. D., & Thanos, C. D. (1999). Oligomeric Structure of the Human EphB2 Receptor SAM Domain. *Science*, 283(5403), 833–836.

Tindi, J. O., Chávez, A. E., Cvejic, S., Calvo-Ochoa, E., Castillo, P. E., & Jordan, B. A. (2015). ANKS1B Gene Product AIDA-1 Controls Hippocampal Synaptic Transmission by Regulating GluN2B Subunit Localization. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 35(24), 8986–8996.

Tonks, N. K. (2003). Overview of Protein Tyrosine Phosphatases. In Handbook of Cell Signaling, Volume 1, (641–651). *Elsevier*.

Tootle, T. L., Lee, P. S., & Rebay, I. (2003). CRM1-mediated nuclear export and regulated activity of the Receptor Tyrosine Kinase antagonist YAN require specific interactions with MAE. *Development*, 130(5), 845–857.

Tran, H. H., Kim, C. A., Faham, S., Siddall, M.-C., & Bowie, J. U. (2002). Native interface of the SAM domain polymer of TEL. *BMC Structural Biology*, 2(1), 5.

Tsujikawa, K., Kawakami, N., Uchino, Y., Ichijo, T., Furukawa, T., Saito, H., & Yamamoto, H. (2001). Distinct functions of the two protein tyrosine phosphatase domains of LAR (leukocyte common antigen-related) on tyrosine dephosphorylation of insulin receptor. *Molecular Endocrinology* (Baltimore, Md.), 15(2), 271–280.

Uddin, L. Q., Supekar, K., Lynch, C. J., Cheng, K. M., Odriozola, P., Barth, M. E., et al. (2015). Brain State Differentiation and Behavioral Inflexibility in Autism. *Cerebral Cortex* (New York, N.Y. : 1991), 25(12), 4740–4747.

Uddin, M., Tammimies, K., Pellecchia, G., Alipanahi, B., Hu, P., Wang, Z., et al. (2014). Brain-expressed exons under purifying selection are enriched for de novo mutations in autism spectrum disorder. *Nature Genetics*, 46(7), 742–747.

Uhlik, M. T., Temple, B., Bencharit, S., Kimple, A. J., Siderovski, D. P., & Johnson, G. L. (2005a). Structural and Evolutionary Division of Phosphotyrosine Binding (PTB) Domains. *Journal of Molecular Biology*, 345(1), 1–20.

Um, J. W., & Ko, J. (2013). LAR-RPTPs: synaptic adhesion molecules that shape synapse development. *Trends in Cell Biology*, 23(10), 465–475.

Van den Berk, L. C. J., Landi, E., Walma, T., Vuister, G. W., Dente, L., & Hendriks, W. J. A. J. (2007). An allosteric intramolecular PDZ-PDZ interaction modulates PTP-BL PDZ2 binding specificity. *Biochemistry*, 46(47), 13629–13637.

Van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., et al. (2014). Classification of intrinsically disordered regions and proteins. *Chemical Reviews,* 114(13), 6589–6631.

Van der Zee, C. E. E. M., Man, T. Y., Van Lieshout, E. M. M., Van der Heijden, I., Van Bree, M., & Hendriks, W. J. A. J. (2003). Delayed peripheral nerve regeneration and central nervous system collateral sprouting in leucocyte common antigen-related protein tyrosine phosphatase-deficient mice. *The European Journal of Neuroscience*, 17(5), 991–1005.

Wagner, S. A., Beli, P., Weinert, B. T., Schölz, C., Kelstrup, C. D., Young, C., et al. (2012). Proteomic Analyses Reveal Divergent Ubiquitylation Site Patterns in Murine Tissues. *Molecular & Cellular Proteomics : MCP*, 11(12), 1578–1585.

Wallace, M. J., Fladd, C., Batt, J., & Rotin, D. (1998). The second catalytic domain of protein tyrosine phosphatase delta (PTP delta) binds to and inhibits the first catalytic domain of PTP sigma. *Molecular and Cellular Biology*, 18(5), 2608–2616.

Wang, F., Wolfson, S. N., Gharib, A., & Sagasti, A. (2012). LAR receptor tyrosine phosphatases and HSPGs guide peripheral sensory axons to the skin. *Current Biology*, 22(5), 373–382.

Wang, R., & Brattain, M. G. (2007). The maximal size of protein to diffuse through the nuclear pore is larger than 60kDa. *FEBS Letters*, 581(17), 3164–3170.

Wei, Z., Zheng, S., Spangler, S. A., Yu, C., Hoogenraad, C. C., & Zhang, M. (2011). Liprin-mediated large signaling complex organization revealed by the liprin-α/CASK and liprin-α/liprin-β complex structures. *Molecular Cell*, 43(4), 586–598.

Weng, Y. L., Liu, N., DiAntonio, A., & Broihier, H. T. (2011). The Cytoplasmic Adaptor Protein Caskin Mediates Lar Signal Transduction during Drosophila Motor Axon Guidance. T*he Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 31(12), 4421–4433.

Wider, G. (2000). Structure determination of biological macromolecules in solution using nuclear magnetic resonance spectroscopy. *BioTechniques*, 29(6), 1278–82–1284–90– 1292 passim.

Wolf, G., Trüb, T., Ottinger, E., Groninga, L., Lynch, A., White, M. F., et al. (1995). PTB domains of IRS-1 and Shc have distinct but overlapping binding specificities. *Journal of Biological Chemistry*, 270(46), 27407–27410.

Wood, L. D., Irvin, B. J., Nucifora, G., Luce, K. S., & Hiebert, S. W. (2003). Small ubiquitin-like modifier conjugation regulates nuclear export of TEL, a putative tumor suppressor. *Proceedings of the National Academy of Sciences*, 100(6), 3257–3262.

Xie, X., Yan, X., Wang, Z., Zhou, H., Diao, W., Zhou, W., et al. (2013). Open-closed motion of Mint2 regulates APP metabolism. *Journal of Molecular Cell Biology*, 5(1), 48–56.

Xie, Y., Yeo, T. T., Zhang, C., Yang, T., Tisi, M. A., Massa, S. M., & Longo, F. M. (2001). The Leukocyte Common Antigen-Related Protein Tyrosine Phosphatase Receptor Regulates Regenerative Neurite Outgrowth In Vivo. *Journal of Neuroscience*, 21(14), 5130–5138.

Xu, H., & Hebert, M. D. (2005). A novel EB-1/AIDA-1 isoform, AIDA-1c, interacts with the Cajal body protein coilin. *BMC Cell Biology*, 6(1), 23.

Xu, Y., & Fisher, G. J. (2012). Receptor type protein tyrosine phosphatases (RPTPs) - roles in signal transduction and human disease. Journal of Cell Communication and Signaling, 6(3), 125–138.

Xu, Y., & Matthews, S. (2013). TROSY NMR spectroscopy of large soluble proteins. *Topics in Current Chemistry*, 335(Chapter 228), 97–119.

Yaffe, M. B. (2002). Phosphotyrosine-Binding Domains In Signal Transduction. *Nature Reviews Molecular Cell Biology*, 3(3), 177–186.

Yan, K. S., Kuti, M., & Zhou, M.-M. (2002a). PTB or not PTB - that is the question. *FEBS Letters*, 513(1), 67–70.

Yan, K. S., Kuti, M., Yan, S., Mujtaba, S., Farooq, A., Goldfarb, M. P., & Zhou, M.-M. (2002b). FRS2 PTB domain conformation regulates interactions with divergent neurotrophic receptors. *Journal of Biological Chemistry*, 277(19), 17088–17094.

Yang, T., Bernabeu, R., Xie, Y., Zhang, J. S., Massa, S. M., Rempel, H. C., & Longo, F. M. (2003). Leukocyte antigen-related protein tyrosine phosphatase receptor: a small ectodomain isoform functions as a homophilic ligand and promotes neurite outgrowth. *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience*, 23(8), 3353–3363.

Yeo, T. T., Yang, T., Massa, S. M., Zhang, J. S., Honkaniemi, J., Butcher, L. L., & Longo, F. M. (1997). Deficient LAR expression decreases basal forebrain cholinergic neuronal size and hippocampal cholinergic innervation. *Journal of Neuroscience Research*, 47(3), 348–360.

Yun, M., Keshvara, L., Keshvara, L., Park, C.-G., Park, C.-G., Zhang, Y.-M., et al. (2003). Crystal structures of the Dab homology domains of mouse disabled 1 and 2. The *Journal of Biological Chemistry*, 278(38), 36572–36581.

Zambrano, N., Buxbaum, J. D., Minopoli, G., Fiore, F., De Candia, P., De Renzis, S., Raffaella Faraonio, R., Sabo S., Cheetham J., Sudol M., and Russo T. (1997). Interaction of the phosphotyrosine interaction/phosphotyrosine binding-related domains of Fe65 with wild-type and mutant Alzheimer's beta-amyloid precursor proteins. *Journal of Biological Chemistry*, 272(10), 6399–6405.

Zhang, J. S., & Longo, F. M. (1995). LAR tyrosine phosphatase receptor: alternative splicing is preferential to the nervous system, coordinated with cell growth and generates novel isoforms containing extensive CAG repeats. *The Journal of Cell Biology*, 128(3), 415–431.

Zhang, Z. (1997). Sequence-specific recognition of the internalization motif of the Alzheimer's amyloid precursor protein by the X11 PTB domain. *The EMBO Journal*, 16(20), 6141–6150.

Zhang, Z. Y., Wang, Y., & Dixon, J. E. (1994). Dissecting the catalytic mechanism of protein-tyrosine phosphatases. *Proceedings of the National Academy of Sciences,* 91(5), 1624–1627.

Zhong, J., Chaerkady, R., Chaerkady, R., Kandasamy, K., Gucek, M., Cole, R. N., & Pandey, A. (2011). The interactome of a PTB domain-containing adapter protein, Odin, revealed by SILAC. *Journal of Proteomics*, 74(3), 294–303.

Zhou, M. M., Huang, B., Olejniczak, E. T., Meadows, R. P., Shuker, S. B., Miyazaki, M., et al. (1996). Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nature Structural Biology*, 3(4), 388–393.

Zhou, M. M., Ravichandran, K. S., Olejniczak, E. F., Petros, A. M., Meadows, R. P., Sattler, M., et al. (1995). Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature*, 378(6557), 584–592.

Zhu, J., Shang, Y., Xia, C., Wang, W., Wen, W., & Zhang, M. (2011). Guanylate kinase domains of the MAGUK family scaffold proteins as specific phospho-protein-binding modules. *The EMBO Journal*, 30(24), 4986–4997.

Zhuang, G., Hunter, S., Hwang, Y., & Chen, J. (2007). Regulation of EphA2 receptor endocytosis by SHIP2 lipid phosphatase via phosphatidylinositol 3-Kinase-dependent Rac1 activation. *Journal of Biological Chemistry*, 282(4), 2683–2694.

Zwahlen, C., Li, S. C., Kay, L. E., Pawson, T., & Forman-Kay, J. D. (2000). Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. *The EMBO Journal*, 19(7), 1505–1515.

Appendix A

Protein NMR Spectroscopy

In the post-genomic era, protein structure has become an essential component for understanding complex biological processes. Structure focused studies also benefit therapeutic discovery, drug development, and biotechnology. For decades, NMR spectroscopy and X-ray crystallography were the only techniques permitting protein structure determination at atomic resolution; however, cryo-EM techniques are now gaining favour in the structural biology community. In contrast to X-ray diffraction methodology that requires protein crystallization, NMR allows studying proteins in their natural solution state. In addition to three-dimensional structure, NMR spectroscopy can provide information about molecular dynamics, conformational equilibria, and ligand-binding kinetics. In the context of this thesis, I will briefly discuss the fundamental principles of NMR to set the foundation for the following section which will focus on the aspects of NMR protein structure determination.

NMR basic theory

Nuclear magnetic resonance (NMR) spectroscopy technique is based on observation of the quantum-mechanical property of atomic nuclei, termed nuclear spin. The nuclei with an even mass and odd charge have an unpaired electron and have spin angular momentum of ½. This means that nuclei exist as an equilibrium of two distinct

energy levels - one of which (ground state) is lower than the other. This electromagnetic property of the nuclei is exploited in NMR experiments. In the presence of an external magnetic field, nuclei absorb, resonate and emit radio frequency energy as electromagnetic radiation. In the absence of an external magnetic field, the nuclear spins orientations are random. Once the magnetic field is applied the spins reorient in such way that one fraction aligns with the field and the other against the field. When nuclei are irradiated with electromagnetic waves of certain radio frequency (RF) the lower energy nuclei spin flips to the higher state, in other words, resonate with the RF. The range of the frequencies used in NMR range from 50 to 1000 MHz.

The perturbation of equilibrium states by short RF pulse or a series of the pulses followed by the return of the system to the equilibrium state is accompanied by a release of energy. The process of a system returning to its thermodynamic equilibrium is called relaxation. The resulting oscillating magnetic field producing an NMR signal is termed Free Induction Decay (FID). FID represents the series of electromagnetic waves measured as a function of time. The Fourier transformed FID signal gives the typical NMR spectrum as a collection of absorption peaks at various resonance frequencies. Different nuclei have distinct resonance frequencies which match the energy difference between the spin populations of this particular nucleus: protons (^{1}H) resonate at ten times higher frequencies than ^{15}N and four times higher than ^{13}C nuclei. Furthermore, when coupled with different atoms into chemical groups the nuclei of the same type will have different resonance frequencies. For instance, amide protons resonate at ≈ 8 ppm, Hα protons at ≈ 4 ppm and methyl protons at ≈ 1 ppm (Wider, 2000). This phenomenon, termed chemical

shift, primarily arises from the influence of elections from surrounding atoms (atom electronegativity) on the local magnetic field around the resonating atom of interest (Cavanagh *et al.*, 2010). The electrons create their own small magnetic field which slightly shields the nuclei from the external field. The chemical shift (δ) value is expressed in ppm (parts per million) and defined as:

$$\delta = ((\omega_{signal} - \omega_{reference}) / \omega_{reference}) * 10^6$$

Instead of Hz the universal ppm units are used as independent of the magnetic field strength B_0. The reference signal of the methyl groups of tetramethylsilane (TMS) or 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) is used as a reference frequency for the ppm scale and defined as the chemical shift of 0 ppm).

Nuclear spins influence each other in two ways: through chemical bonds (scalar coupling) and through space (dipolar coupling). The latter is especially important for protein structure determination. Dipolar interactions manifest themselves as a change of NMR signal intensity of nuclear spin which thermodynamic equilibrium is being perturbed as well as the neighboring spin systems that are situated close enough to have dipolar interactions with it. This phenomenon is called nuclear the Overhauser effect (NOE). NOE arises as a result of cross relaxation and magnetization transfer between the magnetic dipoles of neighboring atoms. It provides valuable spatial information about inter-atomic distances in the macromolecule. J-coupling experiments give correlation

information through one, two or three bonds (^{1}J, ^{2}J, ^{3}J). Amide group heteronuclear coupling (^{1}J ^{15}N-^{1}H) is an example of one bond coupling, whereas CαH-CβH proton coupling is a two bond coupling. Three-bond coupling, such as amide hydrogen to alpha hydrogen, gives information about dihedral angles essential for structure determinations (Almeida, Moraes, & Gomes-Neto, 2013).

The 1-dimensional (1D) NMR spectra of large macromolecules such as proteins are very complex with many overlapping peaks. Therefore, protein NMR spectroscopy exploits the information from a number of NMR experiments in order to assign the collection of resonance signals to the particular amino acid. The same amino acid type in the folded protein structure will have a different chemical shift profile depending on its spatial position in the globular structure. NMR dimensionality, correlation experiments and their interpretation resulting in 3D NMR structures will be discussed in subsequent sections.

Sample requirements

The success of high resolution NMR data acquisition suitable for structure determination relies on the protein sample purity, stability and requires high concentration. Since the natural abundance of NMR-detectable nuclei ^{15}N and ^{13}C is only 0.37% and 1.1% respectively (Wider, 2000) biochemical isotopic enrichment is required (McIntosh & Dahlquist, 1990). Bacterial protein expression systems are the most commonly used for protein labeling. The so-called minimal medium contains salts and trace mineral required for bacterial growth and proliferation and supplemented with

supplemented with $^{15}NH_4Cl$ and unlabeled or ^{13}C-glucose as a sole source of nitrogen and carbon. An addition of 10% D_2O is required for required for control and stabilization of magnetic field strength. For proteins bigger than 30 kDa, a ^{15}N,^{13}C,^{2}H triple labelling technique is used, achieved by substituting H_2O with D_2O in addition to ^{15}N/^{13}C isotopic labeling. The fact that larger molecules tumble slower in solution affects their relaxation rate and causes NMR signal broadening. By eliminating the bulk of proton signal the spectral resolution is greatly improved. Other methodologies such as methyl-specific labeling (Otten *et al.*, 2010) or amino acid specific labeling (Krishna & Berliner, 1999; McIntosh & Dahlquist, 1990) is also used for large proteins.

Signal correlations and NMR dimensionality

As noted earlier, NMR structural determinations require incorporation of both through-bond and through-space correlation experiments (Cavanagh *et al.*, 2010; Wider, 2000). 2D experiments employing J-coupling (scalar coupling) for direct through-bond correlation of different nuclei are called correlation spectroscopy or COSY. For instance, HSQC (Heteronuclear Single Quantum Correlation/Coherence) reflects all amide group N-H correlations, most of which are backbone originated. In the HSQC experiment, the magnetization is transferred from ^{1}H to ^{15}N then after chemical shift evolution it is transferred back to proton and recorded at the ^{1}H channel. The 3D experiments are practically based on 2D experiments propagated to the third dimension, where *x* and *y* are ^{1}H and ^{15}N and *z* is ^{13}C dimension. The correlation of all three signals as a result of a combination of HSQC and HNCO data will gain a single cross-referenced peak for each

amide group and adjacent carbonyl (C'O-N-NH) (Kay *et al.*, 2011) (refer to **Figure A1**).

Through-space correlation experiments that employ nuclear Overhauser effect are observed when two or more spins are not more than a few angstroms apart and become coupled with each other through the dipolar coupling (Almeida *et al.*, 2013). The experiments that derive the distance constraints between protons called NOE and also referred as NOESY (Almeida *et al.*, 2013; Wider, 2000). They allow the correlations between residues that could be far apart in the primary sequence, but nearby in space in folded protein structure. The NOE intensity depends on cross-relaxation and reduced with the power of six proportionally to the distance between two interacting spins (r^{-6}). Therefore it allows the measurement of dipolar coupling within 6 Å only (Almeida *et al.*, 2013; Wider, 2000).

Backbone assignment strategies and generation of 3D structure

High-resolution data acquisition necessary for successful structural calculations is achieved by use of the higher field magnets (750 MHz and up). Several multidimensional NMR strategies are used for tailoring an experimental information into the final globular structure. The most straightforward backbone assignment method requires triple resonance experiments such as CBCANNH and CBCA(CO)NNH that are based on the observation of all three nuclei: ^{1}H, ^{15}N and ^{13}C (illustrated in **Figure A.1**). These experiments allow the sequential assignment of molecular backbone, in other words, the protein primary sequence. The TROSY (transverse relaxation optimized spectroscopy) techniques could simplify the backbone assignments for larger proteins where signal overlap is an issue (Xu & Matthews, 2013).

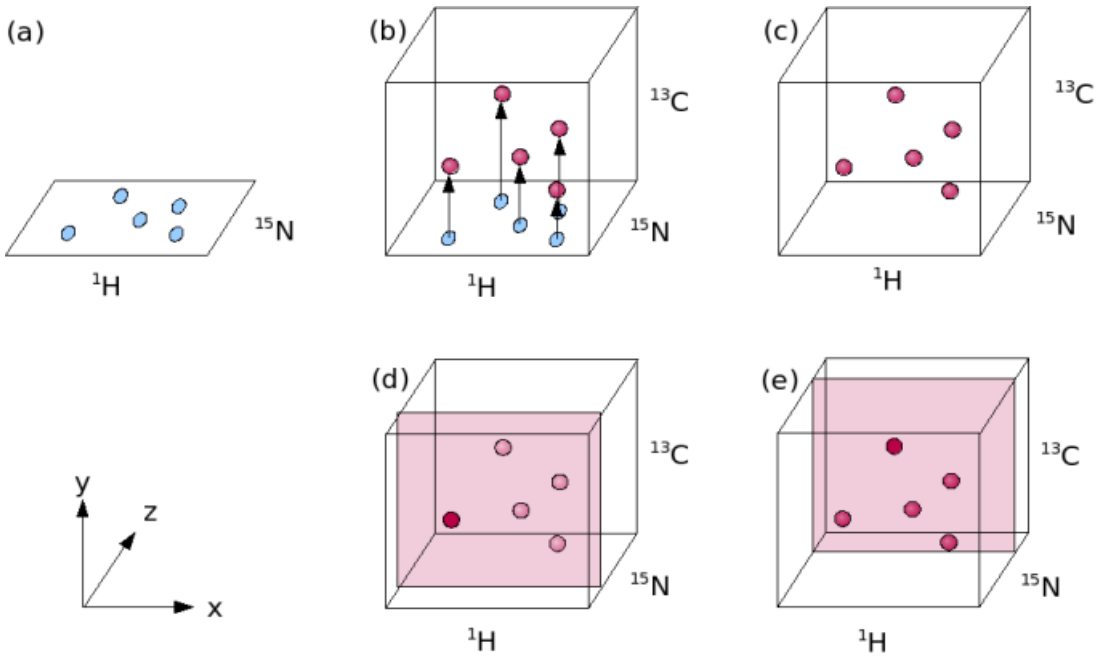


Figure A1 – The schematic presentation of 2D and 3D NMR spectrum. (a) The crosspeaks correlating [^{1}H, ^{15}N] resonances in 2D (HSQC) spectrum are translated into the third ^{13}C-dimension **(b, c).** The ^{1}H dimension is generally left in the x-dimension and in most cases the ^{13}C dimension is viewed along the y-axis, leaving ^{15}N to form the z-axis. Essentially the 3D spectrum can be visualized as a stack of a ^{1}H-^{13}C 2D spectra lined up along the ^{15}N dimension **(d** and **e).** The figure is reprinted from the online resource: *Protein NMR. A practical Guide:* (*www.protein-nmr.org.uk*) by Dr. Victoria A. Higman with author's permission.

The side chain assignments follow the back-bone assignments. For example, the most commonly used combination is HCCH-TOCSY together with HCCH-COSY. Complex analysis of the two links the side chain resonances with previously assigned backbone resonances. NOE experiments contribute the missing geometric and distance information based on dihedral angles and dipolar coupling. NMR structure calculations generate the solutions that optimally satisfy the atomic distance and angular restraints. Therefore, instead of a single structure the final NMR structure represents an average of the population of best structures. The degree of consistent distance constraints used for structural calculations directly affects the quality of 3D structure (Cavanagh *et al.*, 2010; Kay *et al.*, 2011).

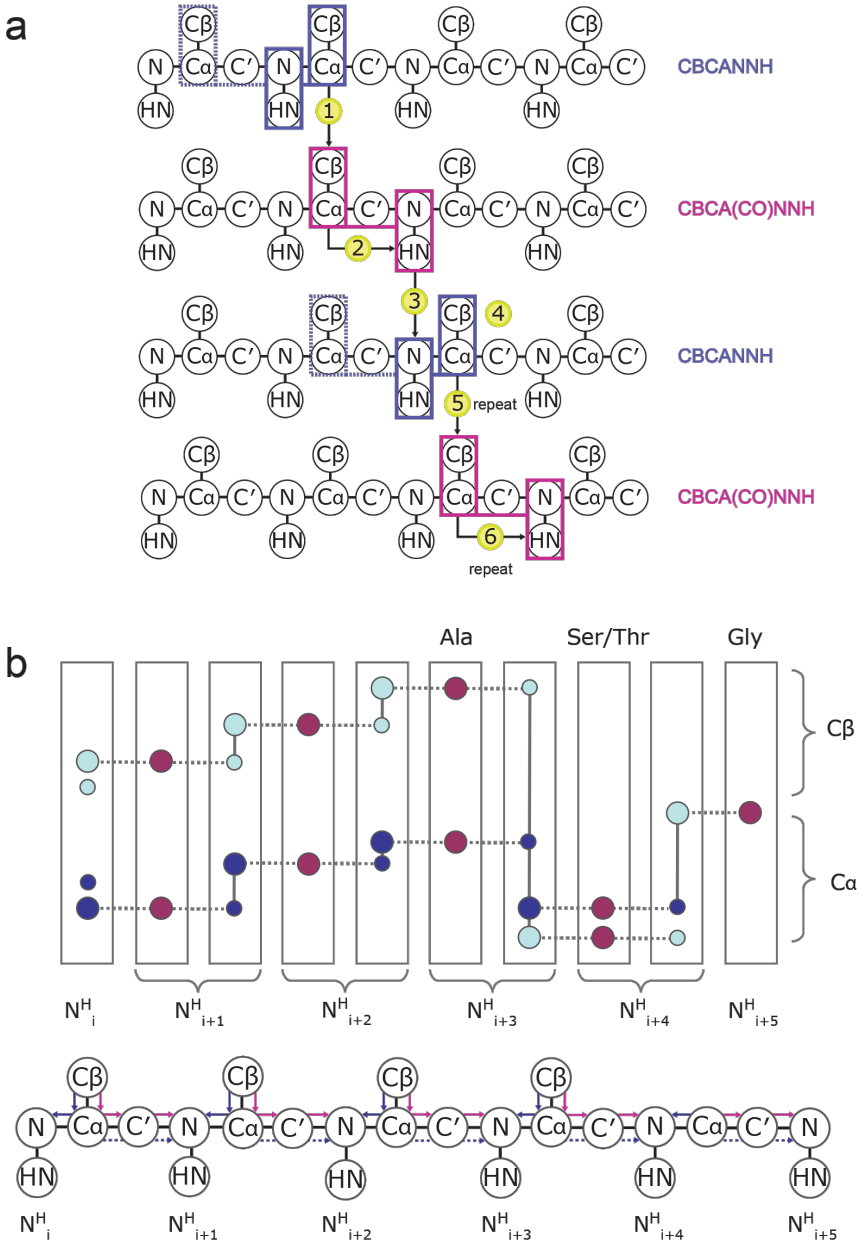


Figure A2 – A schematic diagram of sequential backbone assignment using triple resonance correlation experiments: CBCANNH and CBCA(CO)NNH. (a) Logistics of sequential correlation of Cα/Cβs and NH resonances in the peptide chain. CBCANNH strongly correlates each NH group with the Cα and Cβ chemical shifts of its own residue and more weakly with Cα-Cβ of the preceding residue. The CBCA(CO)NNH correlates the NH group to the preceding Cα and Cβ chemical shifts. **(b)** Two superimposed experiments converted into a strip view, with Cαs are shown in dark blue, Cβs in light blue, and dashed lines showing the correlations between the experiments. The figure is reproduced from the online resource: *Protein NMR. A Practical Guide:* (*www.protein-nmr.org.uk*) by Dr. Victoria A. Higman with author's permission.

Appendix B

Additional Research Accomplishments

A combined ITC and NMR spectroscopy study, to characterize a gap junction protein Connexin 36 (Cx36) and its interaction with calmodulin (CaM) has been included as an additional research accomplishment. This work was completed as part of a collaboration with the laboratory of Dr. G. Zoidl and published in:

> Siu R.C.F., Smirnova E., Brown C.A., Zoidl C., Spray D.C., Donaldson L.W., Zoidl G. (2016) "Structural and functional consequences of Connexin 36 (Cx36) interaction with Calmodulin". *Frontiers in Molecular Neuroscience. 9:120.*

I include here the manuscript abstract and outline the experiments performed by myself.

Summary

Functional plasticity of neuronal gap junctions involves the interaction of the neuronal connexin36 with calcium/calmodulin-dependent kinase II (CaMKII). The important relationship between Cx36 and CaMKII must also be considered in the context of another protein partner, Ca2+ loaded calmodulin, binding an overlapping site in the carboxy-terminus of Cx36. We demonstrate that CaM and CaMKII binding to Cx36 is calcium-dependent, with Cx36 able to engage with CaM outside of the gap junction plaque. Furthermore, Ca2+ loaded calmodulin activates Cx36 channels, which is different to other connexins. The NMR solution structure demonstrates that CaM binds Cx36 in its characteristic compact state with major hydrophobic contributions arising from W277 at

anchor position 1 and V284 at position 8 of Cx36. Our results establish Cx36 as a hub binding Ca2+ loaded CaM and they identify this interaction as a critical step with implications for functions preceding the initiation of CaMKII mediated plasticity at electrical synapses.

Isothermal Titration Calorimetry (ITC)

Isothermal titration calorimetry (ITC) measurements were performed on MicroCal VP-ITC calorimeter (MicroCal Inc., Northampton, MA, USA). The recombinant H6-CaM (affinity purified as described above) was titrated with Cx36 derived peptides reflecting the wild type (GSGWRKIKLAVRGAQAKRKSVYEIR; CanPeptide Inc., Montréal, QC, Canada) or the W227A mutant (GSGARKIKLAVRGAQAKRKSVY; BioBasic, Markham, ON, Canada). An optimal titration was achieved with 25 μM CaM in the reaction cell and 250 μM Cx36 peptide in the syringe, each in a buffer containing 50 mM NaCl, 5 mM BisTris pH 7.0, 5 mM CaCl2. After an initial injection of 2 μL, the bulk of the titration consisted of 34 successive 8 μL injections with an equilibration delay of 300 s. Heats of dilution were determined by titrating the same peptide solutions into buffer alone. Titration profiles corrected for the heat of dilution were fitted into one-binding-site model using MicroCal Origin vv5.0. All values reported were calculated based on three individual experiments.

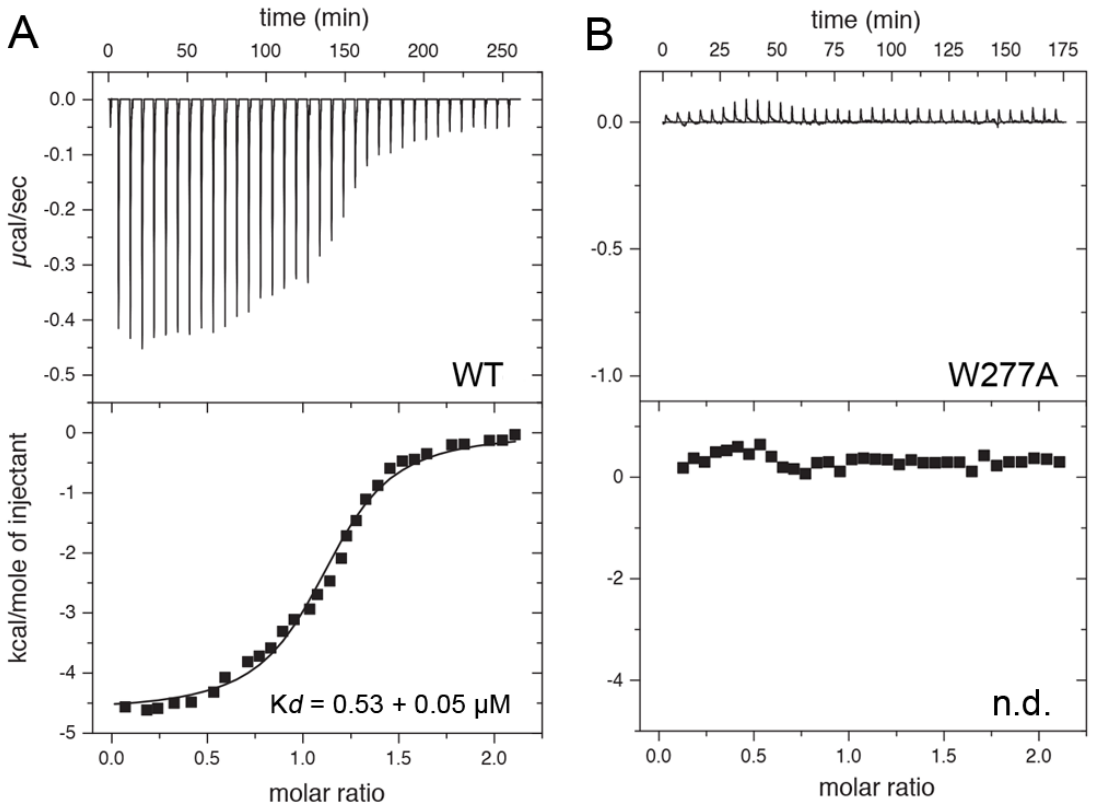


Figure B1 – Isothermal titration calorimetry (ITC) of Cx36 peptides and CaM. Calorimetric traces and integrated isotherms acquired at 30 ºC for (**A**) Calcium saturated CaM titrated with a Cx36 derived peptide, and (**B**) Calcium saturated CaM titrated with the same peptide bearing a W277A substitution. No binding was detected between CaM and W277A mutant peptide. Abbreviations: K_d, dissociation constant; n.d., not detected.

The ITC experiments were performed to determine the affinity of CaM for a 25 amino acid long peptide derived from the Cx36 cytoplasmic tail (**Figure B1**). As predicted from the NMR structure, the Cx36 peptide bound Ca^{2+}-CaM stoichiometrically with a K_d of 0.53 ± 0.05 μM (n = 3). Highlighting the role of the amino acid W277 at the position 1 of the 1–8–(14) motif, no binding was observed for the peptide bearing a W277A substitution (n = 3). Consistent with the structure and biological assays described throughout the study, peptide binding was only observed with Ca2+ saturated CaM. In addition, two other Cx36 flanking residue mutants G276A and G276A were tested, as well as Ca2+-free CaM was titrated with the wild type Cx36 peptide. Results are summarized in the **Table B1**.

Table B1: The Dissociation Constants and Thermodynamic Characteristics of CaM Interactions with WT and Mutant Cx36 Peptides Used in This Study.

Data	**Ka (M-1) remove later**	**Fitting mode**	**Kd (μM)**	**ΔH (kcal mol-1)**	**-TΔS (kcal mol-1)**
Ca2+CaM titrated with Cx36 WT peptide	1.9075x106	Single site binding	0.528 ± 46	– 5.00 ± 0.93	– 3.71 ± 0.97
CaM (Ca2+ free) titrated with Cx 36 WT peptide	nbd	na	nbd	nbd	nbd
Ca2+CaM titrated with Cx36 G276A peptide	3.42x105	Single site binding	2.92 ± 1.7	-5.09 ± 0.72	-2.57 ± 0.79
Ca2+CaM titrated with Cx36 R278A peptide	6.28x104	Double site binding	15.9 ± 2.9	-4.45 ± 1.3	-3.33 ± 0.19
	5.77x107		0.017 ± 0.0011	-5.33 ± 0.17	-5.41 ± 0.27
	1.64x105	Single site binding*	6.1 ± 0.02	-6.10 ± 0.02	-1.602 ± 0.02
Ca2+CaM titrated with Cx36 W277A peptide	nbd	na	nbd	nbd	nbd

Data acquired in 50mM NaCl, 5mM BisTris pH7.0, 5mM $CaCl_2$ (no $CaCl_2$ for apoCaM experiment).
The kinetic and thermodynamic parameters reported were calculated as averages of 2-4 individual experiments for each peptide, with error range expressed as a standard deviation. All experiments used for analysis were corrected for the heat of dilution. Ndb = no binding detected. of the titrant.
*A double binding site fitting algorithm produced the best results for R278A peptide, parameters produced by the single binding site fitting provided for the comparison.