

AN ABSTRACT OF THE DISSERTATION OF

Amber Vogel for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on September 9, 2022.

Title: Strength in Numbers: How Multivalent WW-PPXY Interactions Regulate Cell Signaling.

Abstract approved:

Afua A. Nyarko

Cell signaling is often mediated by protein-protein interactions, which must be specific, tunable, and transient to allow agile responsiveness to cellular messages. Due to their unique properties, multivalent, intrinsically disordered proteins make ideal candidates to accomplish these vital tasks. A single protein with multiple binding sites may bind numerous partners, leading to diverse but integrated cellular outcomes. However, despite their high prevalence and importance, several challenges have hindered the detailed characterization of multivalent, disordered proteins and complexes. First, these proteins are often low yielding and have poor stability and solubility. Second, the inherent dynamic nature of multivalent, disordered proteins often translates to the formation of intricate heterogenous mixtures of protein complexes. Third, few biophysical techniques are amenable to studying these complicated and dynamic systems.

In this work, we overcome these challenges to elucidate the mechanisms of complex assembly for three sets of interactions. The defining feature of each interaction is multivalency, as each protein contains multiple recognition sites. At the heart of these interactions is the multivalent scaffolding protein, Angiomotin-Like 1 (AMOTL1), which plays essential roles in cell growth regulation, cell polarity, shape, and tight junction formation. Important for this thesis is a largely disordered segment of AMOTL1 with three short linear motifs of the sequence L/PPXY (L=leucine, P=proline, Y=tyrosine, and X=any amino acid, hereafter referred to as PPXY). The

partner proteins – YES-associated protein (YAP), Kidney and Brain-expressed protein (KIBRA), and neural precursor cell-expressed developmentally downregulated 4 (NEDD4) – contain multiple repeats of the globular WW domain, an arrangement that facilitates multivalent interactions with AMOTL1 that are functionally relevant. Thus, understanding how each protein complex assembles in the context of multivalency and how AMOTL1 discriminates between its partners could inform critical cellular processes.

Chapter 1 introduces unique features that make intrinsically disordered proteins ideal candidates in multivalent interactions. Emphasis is placed on specific multivalent interactions mediated by the largely disordered PPXY-motif region of AMOTL1 and its multivalent globular WW domain-containing partners. In the following chapters, we provide in-depth molecular biophysics studies of the solution properties of the disordered PPXY-rich segment of AMOTL1 and its interaction with YAP (Chapter 2), KIBRA (Chapter 3) and NEDD4 (chapter 5). Unique characteristics of each interaction are highlighted. Chapter 4 is a study of the solution properties of the multiple WW domains of NEDD4, with emphasis on structural dynamics and global communication between the domains. Chapter 6 summarizes key findings, impact, and future directions of this work.

©Copyright by Amber Vogel
September 9, 2022
All Rights Reserved

Strength in Numbers: How Multivalent WW-PPXY Interactions Regulate Cell Signaling.

by
Amber Vogel

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 9, 2022
Commencement June 2023

Doctor of Philosophy dissertation of Amber Vogel presented on September 9, 2022.

APPROVED:

Major Professor, representing Biochemistry and Biophysics

Head of the Department of Biochemistry and Biophysics

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Amber Vogel, Author

ACKNOWLEDGEMENTS

This body of work was made possible by Dr. Afua Nyarko and the Biochemistry and Biophysics Department at Oregon State University. When I applied to graduate school, I did not believe I could succeed as a scientist. However, five years of dedicated mentoring and support from this department have proven me wrong. The people I acknowledge here have transformed me into a scientist.

Afua, when I first rotated in your lab, I was not so sure biophysics was for me. However, after working in your lab for a single term, I was absolutely hooked on dissecting the subtle puzzles of multivalent protein-protein interactions. Furthermore, I thought you were the classiest person I had ever met, and I thought maybe I could learn some of that as well. While I don't think I've become classier, your guidance, investment, patience, and support have allowed me to succeed as a biophysicist. I cannot tell you how grateful I am to have had you as a mentor. You always made time for me, and I really felt like I could come to you with any question or concern. Also, I can't wait to pester you with more emails after I move on to the next stage of my career.

I want to acknowledge and thank all the Nyarko lab members. You have become some of my closest friends. Kasie, you are one of the most thoughtful people I have ever met! You are so supportive, open, kind, and helpful. You are my science sister, and I know that we will always be there for each other. Sanjay, you are such an incredibly fun person to do science with! I love talking to you about our projects, sharing ideas, and troubleshooting. You challenge me to be a better scientist, and I know I could trust you with my life. I would like to thank Kathy, Diego, and Lydia. Although we didn't work on projects together, sharing a space with you was delightful. I will be lucky to share future labs with such wonderful people.

During grad school, I had the pleasure of working with two wonderful undergraduates: Alexandra Crawford and Matthew McWhorter. Allie, although our time in lab was cut short because of Covid, you were the best first trainee I could have asked for. Not only are you bright, kind, and excited to learn, you became one of my closest friends. At times when I wanted to give up, I thought of you, and how I held the

responsibility of helping prove to you that women can succeed in science. You are one of the reasons I made it through. Matt, you impressed me with your perseverance and hard work. Even though you started in the lab before Covid, you never quit until you graduated, and your dedication transformed you into a smart and independent lab member. Also, thank you for always thinking of me and bringing me caffeine when you knew I was having a bad day.

Next, I would like to thank others in the department. I would like to acknowledge my cohort who provided endless comradery, support, and friendship. I am truly honored to have been part of our group, and I wish you all happiness and success. In addition to my cohort, I would like to thank all the graduate students for all the support and the fun experiences we shared. Whenever I needed help, you were there, and I am incredibly grateful for you.

I would like to thank my committee members: Dr. Elisar Barbar, Dr. Richard Cooley, Dr. Siva Kolluri, Dr. Glencora Borradaile, and my previous committee members, Dr. Wei Kong and Dr. Sandra Loesgen. Elisar, thank you for being my science grandmother. You were always willing to share your time and expertise to help me analyze data, review manuscripts, and plan experiments when I needed extra help. Rick, I truly can't count the number of times I came into your office asking for help or advice, and you were always willing to sit and talk to me, even though I wasn't your student. I really appreciate all your support, and it was a pleasure collaborating with you. Siva, thank you for supporting and believing in me, even though I didn't follow through on the collaboration we were planning (I'm sorry about that). I appreciate your kindness, and I hope we can do science together someday in the future.

Although I cannot thank everyone in the department by name, there are a few other individuals who I would like to highlight. Dr. Patrick Reardon, how can I thank you enough? If it wasn't for you, I would still be trying to process NMR data right now. Your knowledge and creativity sparked many of my experiments, and your support made me a better scientist. Dr. Andy Karplus, since the day I visited, you have always believed in me, and your confidence, guidance, honesty, and thoughtfulness built an important framework that helped me succeed. Whether it be writing a manuscript, conceptualizing project ideas, or helping me get a job, you were always there to support

me. Dr. Ryan Mehl, even though I didn't join your lab, you were always available to help me with science and planning my career. Thank you for going above and beyond to help me succeed.

In addition to the department, I must thank my family and friends. Dad, thank you for helping me move to Corvallis, giving me a new laptop to write this thesis with, supporting me, and believing in me. Granny and Gramps, thank you for helping me get to this point in my life and for supporting me endlessly. Thank you, Kameron, Mary, and August; your friendships mean the world to me.

Finally, I want to thank my partner: Ilana Gottfried-Lee. Ilana, you are the most loyal and honest person I have ever met, and your support has gotten me through grad school. You taught me the importance of balancing work with the rest of my life. Thank you for all the sacrifices you made to support me and be with me while I finished graduate school. Thank you for talking about science with me, and for making me stop talking about science. I can't wait to start the next adventure with you. I love you.

CONTRIBUTION OF AUTHORS

Dr. Afua Nyarko contributed to conceptual development, experimental design, data analysis, and writing all chapters presented herein. Alexandra Crawford contributed to cloning, protein production, and NMR experiments presented in Chapter 2, Appendix 1, and Appendix 6. Matthew McWhorter contributed to cloning, protein production, ITC, and NMR experiments presented in Chapters 3-5, and Appendices 1, 4, and 5. Hannah Stuwe contributed to protein production, NMR experiments, and data analysis in Chapter 4. Miranda Leek contributed to NMR experiments in Appendix 3.

TABLE OF CONTENTS

	<u>Page</u>
Chapter 1: Introduction – Assembly of multivalent, disordered signaling protein complexes: insights from WW domain-PPXY motif interactions.....	1
Intrinsically disordered proteins and their roles as molecular scaffolds.....	2
Overview.....	2
IDP scaffolds and binding properties.....	3
Multivalent IDP assemblies.....	5
Modes of multivalent binding.....	5
Regulation of multivalent binding.....	7
Significance of multivalency.....	8
Homovalent IDP assemblies: insights from WW-PPXY complexes.....	8
Overview.....	8
Classification and binding.....	9
Affinity and specificity of WW domain-PPXY motif interactions....	11
WW domains and PPXY motifs in cell signaling and disease.....	13
YAP.....	14
KIBRA.....	17
AMOTL1.....	18
NEDD4.....	20
Contents of dissertation.....	22
Chapter 2: Multivalent Angiotensin-like 1 and Yes-associated protein form a dynamic complex.....	24
Abstract.....	25

TABLE OF CONTENTS (Continued)

	<u>Page</u>
Abbreviations.....	25
Introduction.....	26
Results.....	28
Discussion.....	37
Materials and Methods.....	39
Acknowledgements.....	42
 Chapter 3: Domain and motif-specific roles in assembly of the multivalent WW-PPXY complex of Angiomotin-Like 1, Yes-associated protein and Kidney and Brain-expressed protein.....	 43
Abstract.....	44
Introduction.....	45
Results.....	47
Discussion.....	55
Materials and Methods.....	58
 Chapter 4: Domain-domain communication in multivalent WW proteins: insights from the tetravalent NEDD4-1 WW domains.....	 60
Abstract.....	61
Introduction.....	62
Results.....	64
Discussion.....	80
Materials and Methods.....	82
 Chapter 5: Biophysical characterization of the AMOTL1-NEDD4 complex.....	 85
Abstract.....	86

TABLE OF CONTENTS (Continued)

	<u>Page</u>
Introduction.....	87
Results.....	88
Discussion.....	97
Materials and Methods.....	99
Chapter 6: Conclusions and Future Directions.....	101
Highlights of reported work.....	102
Conclusions and Impact.....	103
Future directions.....	104
Appendices	107
References	167

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
Figure 1.1: Compositional features and functions of intrinsically disordered proteins and regions.....	4
Figure 1.2: Modes of multivalent binding.....	6
Figure 1.3: Representative WW domain sequences and structures.....	10
Figure 1.4: Domain architecture of YES-associated protein.....	15
Figure 1.5: Canonical and noncanonical regulation of YAP/TAZ through the Hippo signaling pathway.....	16
Figure 1.6: Domain architecture of Kidney and Brain-expressed protein.....	18
Figure 1.7: Domain architecture of Angiotensin-Like 1 protein.....	19
Figure 1.8: Domain architecture of NEDD4-1 protein.....	20
Figure 2.1: Domain architecture and solution properties of apo and bound AMOTL1 and YAP polypeptides.....	27
Figure 2.2: NMR backbone assignments and solution dynamics of AMOTL1 A ₁₂₃ polypeptide.....	30
Figure 2.3: A ₁₂₃ solvent accessibility.....	31
Figure 2.4: NMR mapping of the A ₁₂₃ -YWW _{TD} binding interface.....	33
Figure 2.5: ITC binding isotherms of the A ₁₂₃ -YWW _{TD} interaction.....	35
Figure 2.6: Model of the species formed in the AMOTL1-YAP complex.....	38
Figure 3.1: KWW _{TD} and A ₁₂₃ construct design and complex assembly.....	47
Figure 3.2: Binding of KWW _{TD} to A ₁₂₃ /mutants monitored by ITC.....	49
Figure 3.3: Binding of KWW _{TD} WW domain mutants to A ₁₂₃ and A ₃ monitored by ITC.....	50
Figure 3.4: Binding-induced chemical shift changes of the KWW _{TD} WW domains..	51

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 3.5: The binding interface of A ₁₂₃ mapped by NMR shows binding of all three PPXY motifs initiated by P3.....	52
Figure 3.6: Assembly of the YWW _{TD} -A ₁₂₃ -KWW _{TD} ternary complex monitored by analytical SEC, SEC-MALS, and ITC.....	53
Figure 3.7: ITC isotherm of KWW _{TD} titrated into YWW _{TD} at 30 °C.....	54
Figure 3.8: Proposed model of complex assembly.....	57
Figure 4.1: Domain organization, construct design, and sequences of the NEDD4-1 WW domains.....	63
Figure 4.2: Average secondary structure and oligomerization state of WW1-4.....	66
Figure 4.3: NMR comparison of isolated NEDD4-1 domains and WW1-4.....	67
Figure 4.4: ¹ H- ¹⁵ N-HSQC spectra and assignments for isolated NEDD4-1 WW domains.....	69
Figure 4.5: Residue-level secondary structure characterization of isolated NEDD4-1 WW domains.....	70
Figure 4.6: Residue-level dynamics of WW1, WW3, and WW4 determined by NMR.....	72
Figure 4.7: Interdomain differences in dynamics of WW1, WW3, and WW4.....	74
Figure 4.8: ¹ H- ¹⁵ N-HSQC comparison of NEDD4-1 WW1 and WW2 to tandem WW1-2.....	75
Figure 4.9: ¹ H- ¹⁵ N-HSQC comparison of WW3 and WW4 to WW3-4.....	76
Figure 4.10: Sortase A-mediated ligation and NMR validation of ¹⁵ N-WW3-4 to unlabeled WW1-2.....	78
Figure 4.11: ¹ H- ¹⁵ N-HSQC comparison of WW3-4 to WW1-2- ¹⁵ N(3-4).....	79
Figure 4.12: Proposed model of NEDD4-1 WW domain interconnectivity.....	81
Figure 5.1: Construct Design of NEDD4-1 and AMOTL1 with all putative binding domains.....	90

LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure 5.2: SEC-MALS and ITC show WW1-4 and A ₁₂₃ form a 1:1 complex.....	91
Figure 5.3: ITC of mono-, bi-, and trivalent NEDD4-1 WW domain constructs titrated into A ₁₂₃	93
Figure 5.4: The WW1-4-binding interface of A ₁₂₃ mapped by ITC and NMR.....	96
Figure 5.5: Proposed model of NEDD4-1-AMOTL1 assembly.....	98

LIST OF TABLES

<u>Table</u>	<u>Page</u>
Table 1.1: Valency of proteins discussed in this work.....	14
Table 2.1: Experimental and theoretical molar masses of apo and bound AMOTL1 A ₁₂₃ and YAP YWW _{TD}	29
Table 2.2: Thermodynamics parameters for the YAP-AMOTL1 interactions.....	36
Table 3.1: Average molar masses of apo and bound KWW _{TD}	48
Table 3.2: ITC of KWW _{TD} titrated into A ₁₂₃ /mutants collected at 25 °C.....	48
Table 3.3: ITC of KWW _{TD} mutants titrated into A ₁₂₃ or A ₃ collected at 25 °C.....	49
Table 3.4: ITC of YWW _{TD} or KWW _{TD} titrated into the alternate A ₁₂₃ -WW domain complex collected at 25 °C.....	54
Table 4.1: NEDD4-1 construct tags.....	65
Table 4.2: SEC-MALS of WW1-4 and WW1-2- ¹⁵ N(3-4).....	65
Table 5.1: Tags of NEDD4-1 constructs.....	89
Table 5.2: SEC-MALS of WW1-4 bound to A ₁₂₃	91
Table 5.3: Thermodynamic parameters for A ₁₂₃ bound to varying valency NEDD4-1 WW domains.....	92
Table 5.4: Thermodynamic parameters for the WW1-4 – A ₁₂₃ /mutant complexes....	95

LIST OF APPENDICES

<u>Appendix</u>	<u>Page</u>
Appendix 1: Cloning, expression, and purification of proteins.....	108
Appendix 2: Preliminary data on competition between YAP, KIBRA, and NEDD4-1 for AMOTL1.....	114
Appendix 3: NMR analysis of the tandem WW domains of Yes-associated protein.....	121
Appendix 4: Interactions of YAP, KIBRA, and NEDD4-1 WW domains with viral PPXY sequences.....	131
Appendix 5: Liquid-liquid phase separation studies of the NEDD4-1-AMOTL1 complex.....	153
Appendix 6: What factors govern AMOTL1 motif specificity?.....	162

LIST OF APPENDIX FIGURES

<u>Figure</u>	<u>Page</u>
Figure A2.1: Competition SEC-MALS.....	116
Figure A2.2: Competition ITC.....	118
Figure A3.1: Initial NMR characterization of YWW _{TD}	123
Figure A3.2: HSQCs of P202A _{TD}	124
Figure A3.3: Optimization of YWW _{TD} NMR conditions.....	126
Figure A3.4: YWW _{TD} assignments and peak intensities at pH 5.....	127
Figure A3.5: Amide solvent exchange measured by CLEANEX NMR.....	128
Figure A4.1: Viral protein domain architecture.....	133
Figure A4.2: ITC of viral peptides titrated into YWW _{TD} (YAP), KWW _{TD} (KIBRA), WW1-4 and NWW3 (NEDD4-1).....	136
Figure A4.3: NMR titration of HBC PPAY peptide into ¹⁵ N-labeled KWW _{TD}	138
Figure A4.4: ITC titration of bdSUMO-nCoV-S into WW1-4.....	140
Figure A4.5: Oligomerization state of HBC149.....	141
Figure A4.6: SEC-MALS and ITC studies show no binding of HBC149 and WW1- 4.....	142
Figure A4.7: SEC-MALS and native gel electrophoresis show no binding of HBC149 and KWW _{TD}	143
Figure A4.8: ITC and NMR studies of HBC149 and KWW _{TD} show no binding.....	145
Figure A4.9: Structure of HBC149 dimer.....	147
Figure A5.1: Liquid-liquid phase separation diagram.....	155
Figure A5.2: SEC-MALS of WW1-4.....	156
Figure A5.3: Turbidity of WW1-4.....	157

LIST OF APPENDIX FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
Figure A5.4: Confocal microscopy shows droplet formation of WW1-4 mixed with A ₁₂₃	159
Figure A6.1: PPXY motif preferences of YAP, KIBRA, and NEDD4.....	163
Figure A6.2: Motif and domain linker length.....	165
Figure A6.3: Binding of A _{RYΔAA} and KWW _{TD}	166

LIST OF APPENDIX TABLES

<u>Table</u>	<u>Page</u>
Table A1.1: Construct information for genes used in Chapters 2-5.....	110
Table A1.2: Protein expression conditions.....	111
Table A2.1: SEC-MALS of A ₁₂₃ bound to WW1-4, YWW _{TD} , and KWW _{TD}	116
Table A2.2: Thermodynamic parameters for WW1-4/KWW _{TD} /YWW _{TD} titrated into preformed A ₁₂₃ -WW domain complexes.....	118
Table A4.1: Sequences of viral PPXY peptides.....	133
Table A4.2: Binding parameters for viral peptides titrated into WW domain proteins monitored by ITC.....	135
Table A4.3: HBC149 oligomerization state determined by SEC-MALS.....	141
Table A5.1: Phase separation prediction scores.....	155
Table A5.2: SEC-MALS of WW1-4 at increasing concentrations.....	156

DEDICATION

This thesis is dedicated to all women scientists and scientists assigned female at birth who faced doubt because of their gender but who persevered, nonetheless.

Chapter 1

Assembly of multivalent, disordered signaling protein complexes: insights from WW domain-PPXY motif interactions

Introduction

Signal transduction is an essential component of cellular function and is universal in living systems[6]. Signaling pathways control a myriad of processes including cell proliferation, apoptosis, response to cellular stress, hormone regulation, glucose metabolism, immune response, and development. These pathways are mediated by many coordinated protein-protein interactions, which must be specific, easily reversible, and sensitive to changes in the cellular environment[7]. Intrinsically disordered proteins and regions are highly prevalent in signaling pathways, and their unique binding properties make them ideal candidates to achieve the specificity, tunability, and integration required for cell signaling. Understanding how these versatile protein complexes assemble will illuminate features of selectivity which allow precise cellular control and reveal how these interactions may be manipulated to treat diseases. However, traction in characterizing transient, disordered protein complexes has proven challenging, due to difficulties in producing the large samples needed for biophysical studies and the instability of these proteins and complexes. In this work, we overcome these challenges and elucidate the binding mechanisms of three multivalent protein assemblies involved in cell signaling. Our findings reveal that differential binding mechanisms may underly the diverse functions of disordered proteins in cell signaling.

Intrinsically disordered proteins and their roles as molecular scaffolds

Overview. Intrinsically disordered proteins (IDPs) and intrinsically disordered regions of proteins (IDRs) are polypeptide chains of ≥ 30 residues that lack well-defined 3D structure[8]. In the last three decades, there has been an explosion in the field of IDPs/IDRs. While the first article on IDPs was published in 1989[9], a PubMed.gov search for “intrinsically disordered proteins” now produces 6,122 results. In contrast to the traditional paradigm that protein structure confers function, we now know that disordered proteins play a vast number of important roles in the cell, particularly in signaling pathways. This paradigm shift has opened investigation into a huge number of proteins, as IDPs/IDRs are predicted to compose roughly 50% of the human proteome[10].

IDPs/IDRs are characterized by several sequence-based and physical features (Fig. 1.1A). First, since a primary force driving protein folding is the hydrophobic effect[11, 12], it is unsurprising that unfolded proteins are depleted in hydrophobic and aromatic residues and enriched in polar and charged residues, having a much higher net charge to mean hydrophobicity ratio relative to folded proteins[9, 13, 14]. In a 2000 study comparing 91 experimentally validated IDPs and 275 small globular proteins, plotting the mean hydrophobicity versus net charge separated folded and unfolded proteins into two distinct groupings[9]. Another compositional feature of IDPs is their high glycine and proline content. Glycine is highly flexible, making structure more entropically unfavorable[15]. Proline composes 7.3% of IDP sequences vs. 4.4% in folded proteins[14]. The effect of proline on IDP conformation seems to be sequence-context dependent. Generally, proline is thought to confer flexibility to IDPs by disrupting structure[14].

Overall amino acid composition doesn't tell the full story. Although IDPs/IDRs have a higher mutation rate than folded proteins[16], sequence-specific features are important for preserving IDP function, evidenced by the fact that changes in sequence and/or environment can lead to disease, often as a result of protein aggregation[17-19]. In some cases, sequence features of IDPs can be used to predict and rescue their biological functions[19]. For example, raised isoelectric point and hydrophobicity in IDPs/IDRs are associated with mitochondrial localization and targeting[20]. The distribution of charge also has an impact on IDP stability and conformation. Charge clustering causes IDPs to take on more compact conformations with lower hydrodynamic radii due to intramolecular charge-charge interactions and can increase secondary structure content[21-25]. The degree of compaction or expansion of disordered polypeptide chains affects their binding mechanisms and range of binding partners[26].

IDP scaffolds and binding properties. Binding of IDPs/IDRs can occur through versatile mechanisms, which provide the tunability and sensitivity required for cell signaling. Assembly is usually enthalpically driven, having specific hydrogen bonds and polar contacts which provide specificity, while the entropic cost of binding reduces the overall affinity, allowing complexes to easily disassociate[27, 28]. A common

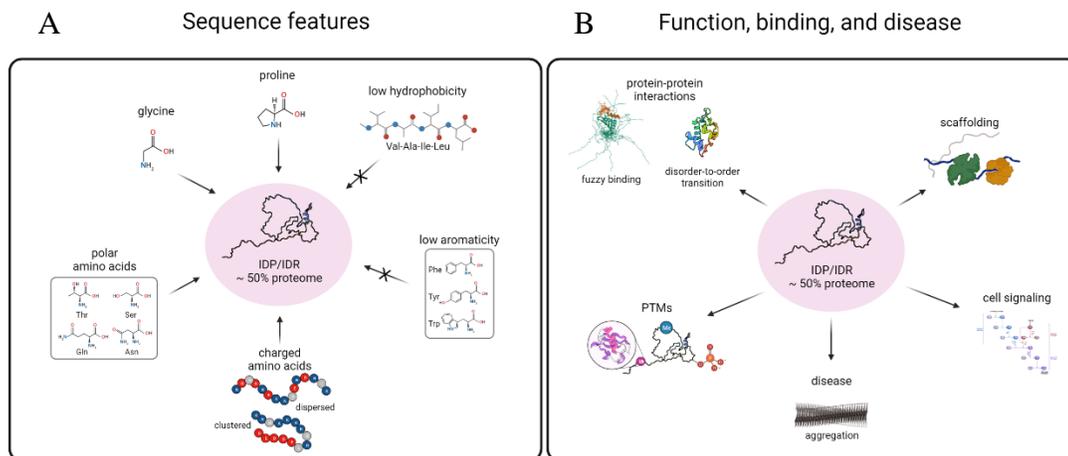


Figure 1.1: Compositional features and functions of intrinsically disordered proteins and regions. IDPs/IDRs compose approximately 50% of the human proteome. The representative IDP depicted here was generated using the structure of bacterial polar organizing protein Z (PopZ, PDB ID: 6XRY). (A) Sequence features such as high charge, high polar content, high proline and glycine content, and low hydrophobicity and aromaticity confer disorder. (B) IDPs/IDRs facilitate protein-protein interactions, which can occur through fuzzy binding (PDB: 2YKA) or disorder-to-order transitions (PDB: 7AD0); are prone to PTMs; provide scaffolds for protein assembly; participate in cell signaling pathways; and regulate disease states. Created with BioRender.com.

method by which IDPs/IDRs participate in cell signaling is through protein scaffolding (Fig. 1.1B)[29]. Scaffolding interactions are highly tunable and help integrate signals between pathways. Scaffolding proteins act as molecular hubs by binding numerous targets simultaneously. These interactions organize the binding partners in space, bringing them together in close proximity to enhance their intermolecular functions. In some cases, IDP/IDR binding interactions retain disorder and are highly dynamic, exhibiting a continuum of conformations or an ensemble of complexes described as “fuzzy binding”[30]. In other cases, IDPs undergo binding-induced folding, which helps provide the defined enthalpic interactions needed to overcome entropic losses[31]. Disordered scaffold-mediated complexes participate in pathways including p53 signaling, the Wnt pathway, tyrosine kinase pathways, TGF- β signaling, JNK signaling, Akt signaling, GPCR signaling, and the Hippo signaling pathway[29, 32, 33].

Multivalent IDP assemblies

The scaffolding ability of IDPs/IDRs is conferred by multivalency. A multivalent molecule is defined by the presence of more than one identical (homovalent) or non-identical (heterovalent) site capable of making interactions with other molecules (Fig. 1.2A). Within disordered proteins, multivalent binding sites are often composed of stretches of three to ten amino acids known as short linear motifs (SLiMs). SLiM sequences are often composed of a combination of invariant and variable residues. While invariant residues define the SLiM identity and class of binding partner, variable residues can provide protein/domain-level specificity, as SLiM recognition sequences are often found in multiple proteins with different cellular roles[34].

Modes of multivalent binding. With the expansive diversity and prevalence of multivalent binding sites, it is unsurprising that multivalent binding can result in many different mechanistic outcomes. In the simplest multivalent system, molecule A has multiple binding sites and its partner, molecule B has only one (Fig. 1.2B). Even in this simple case, multiple binding modes can occur. One possible interaction mode is independent binding, where each multivalent binding site on molecule A acts as an autonomous entity, and binding of molecule B at one site has no impact on binding of a second molecule of B at a second site. Another option is positive cooperativity (otherwise termed synergy or avidity), where binding at the first site increases the affinity of molecule B for the second site. A final possible outcome is negative cooperativity (or attenuation of affinity), where binding at the first site decreases the affinity of molecule B for the second site.

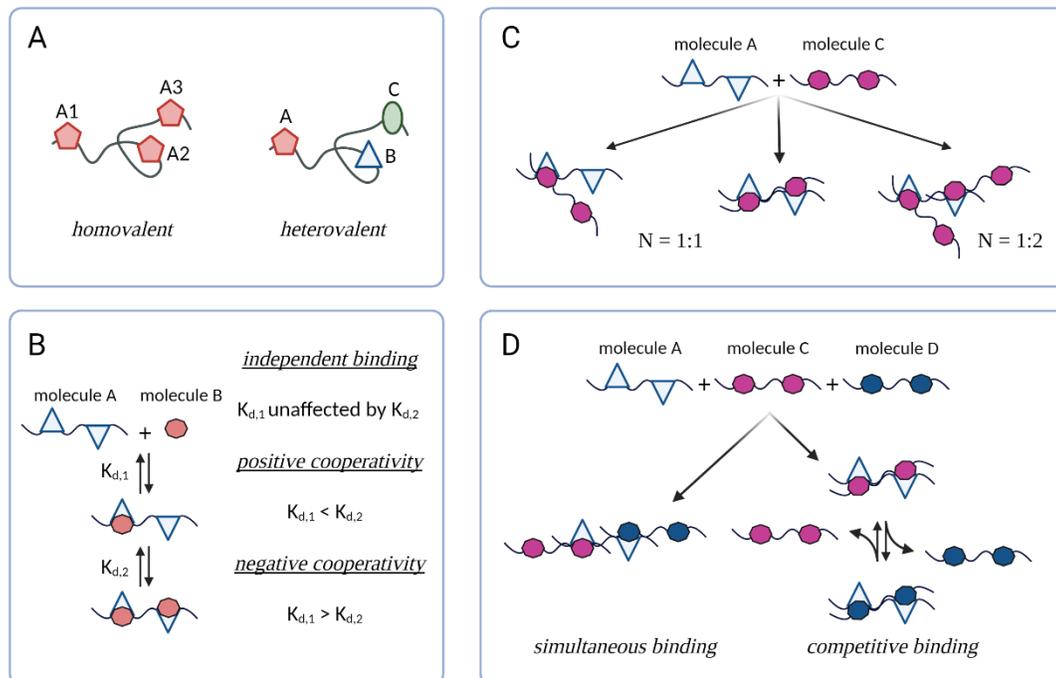


Figure 1.2: Modes of multivalent binding. (A) Cartoon representation of two different trivalent molecules: one with three identical, binding sites (A1, A2, and A3), the other with three non-identical binding sites (A, B, and C). (B) Cartoon representation of a bivalent molecule A and its monovalent partner, molecule B. Molecule B can interact with both sites of A independently ($K_{d,1}$ unaffected by $K_{d,2}$), with positive cooperativity ($K_{d,1} < K_{d,2}$), or with negative cooperativity ($K_{d,1} > K_{d,2}$). (C) Cartoon representation of two bivalent molecules, A and C, with can form complexes of varying stoichiometry (N) and configuration. (D) Cartoon representation of three bivalent molecules, A, C, and D, where molecules C and D are partners of molecule A. Molecules C and D may bind to A simultaneously or compete for binding at the same sites. Created with BioRender.com

Potential binding modes quickly gain complexity as valency increases. When both binding partners have two sites, in addition to factors such as independent and cooperative binding which impact the affinity of the overall interaction, binding configuration also becomes a variable (Fig. 1.2C). Binding may occur with a stoichiometry of one where one site of molecule A is bound to one site of molecule C. However, a stoichiometry of one is also observed when both sites of molecule A bind both sites of molecule C. Alternatively, two different molecules of C could bind each site of molecule A, resulting in a stoichiometry of 1:2. This mode of binding can lead to complex oligomerization states where many molecules of A and C are strung

together forming large polymer-like structures. These scenarios gain further complexity when one or more partners has more than two binding sites.

Multivalent molecules are often degenerate, having multiple copies of the same site which recognize cognate binding sites of multiple different partner molecules. This opens the door to competition and the formation multimeric protein assemblies, making selectivity an important question. For example, take molecules A and C and introduce a third bivalent molecule D, which also binds molecule A (Fig. 1.2D). Molecules C and D may bind alternate sites of A, in which case both molecules could bind A simultaneously, forming a ternary complex. However, molecules C and D may bind the same sites of A, in which case they would compete for molecule A. In the latter scenario, typically the complex with the higher affinity predominates, although a mixture of complexes may also form. Many biomolecules have numerous binding partners which recognize the same binding sites, and understanding which complexes form under different circumstances is important for understanding and tuning the cellular processes these complexes control.

Regulation of multivalent binding. The binding mechanisms, affinities, and specificities of multivalent interactions are modulated by a host of molecular features including sequence identity/composition, valency, alternative splicing, flanking and sometimes distal regions, linker length and structure, dynamics, and post-translational modifications (PTMs)[35-40]. Given the short length of SLiMs, it is unsurprising that flanking residues impact affinity and specificity, contributing an average of 20% to interactions[34, 41]. Linker length and structure play important roles in controlling the effective concentration of multivalent SLiMs. For inert, disordered linkers, shorter length is often favorable due to increased effective concentration of the binding sites and reduced loss in conformational entropy[39, 42]. Increasing the linker length beyond the distance between cognate binding sites can remove the cooperative component of multivalent binding by effectively eliminating increased apparent SLiM concentration[38, 39]. However, if the linker interacts favorably with components of the system, increased length has a less negative impact on binding affinity[36]. In some systems, linkers have propensity to form structure, which can increase binding affinity[43]. However, it is important to consider the possibility of steric clash and

conformational restriction which can result in decreased binding ability[38]. The extended nature of IDPs/IDRs exposes more residues to PTMs. PTMs such as phosphorylation change the charge, binding, and sometimes modulate structure in IDPs/IDRs, especially as PTM sites are often within or adjacent to binding sites[35, 37, 40, 44]. In some cases, PTMs act as molecular switches by completely inducing or abolishing binding[35, 45].

Significance of multivalency. The eukaryotic linear motif resource (ELM, available at <http://elm.eu.org>) curates experimentally validated SLiMs and currently contains 3,934 individual protein motifs composed of 317 classes of SLiMs[46]. Not only is multivalency highly prevalent, it participates in vital processes including intracellular signaling[39, 47], cell-cell communication[39, 48], immune function[49], and morphogenesis[50]. Furthermore, multivalency is now being incorporated into therapeutics including vaccines[51, 52], cancer treatments[53, 54], antiviral drugs[55, 56], antibiotics[57], and neurodegenerative treatments[58, 59]. Dissecting the mechanisms and regulation of multivalent binding interactions is foundational to understanding and manipulating biological function and disease.

Homovalent IDP assemblies: insights from WW-PPXY complexes

Overview. In addition to binding other disordered regions, multivalent IDPs/IDRs can also interact with folded domains. In 1994, Marius Sudol and colleagues discovered the smallest, autonomously-folding interaction domain in the proteome in the cytoskeletal protein dystrophin – the WW domain[60-63]. This domain is 35-40 residues in length and named for the presence of two highly conserved tryptophan (W) residues separated by 20-22 amino acids and folds into a triple-stranded, antiparallel β -sheet. Following the second tryptophan is a highly conserved proline residue, hence the less common name, the WWP domain. WW domains are composed of a high percentage of hydrophobic, aromatic, and proline residues, are typically flanked by IDRs, and their internal structure is composed of a hydrophobic core. WW domains are widespread, with at least 52 proteins containing one or more WW domain in the human proteome and over 10,000 reported in proteins of all species[64]. These proteins are involved in many biological processes including cell

signaling, growth, differentiation, cytoskeletal control, transcription, and protein stability and trafficking[64, 65]. Their pinnacle role in proper cell function can also lead to disease states when WW domains or their binding partners are mutated or dysregulated[66, 67].

Classification and binding. WW domains are interaction modules that bind proline-rich motifs (PRMs). They are organized into five classes based on their cognate SLiM sequence. Class I, which represents the largest group of WW domains, bind (L/P)PXY sequences (hereafter referred to as PPXY motifs), where X is any amino acid[68, 69]. Binding of class I WW domains to PPXY motifs is abolished when the Tyr residue is phosphorylated[70]. Class II WW domains bind PPLP sequences[71]. Class III WW domains bind PGM sequences[72]. Class IV WW domains bind (phospho-S/T)P sequences[73]. Class V WW domains bind (P/R)PR sequences[74]. Unless otherwise stated, further discussion will focus only on class I WW domains and their binding partners. The sequences of the eight WW domains described in this dissertation are shown in Figure 1.3A.

Within the human proteome, there are close to 2,000 identified PPXY motifs[75]. These SLiMs are usually disordered but adopt polyproline II helical structure upon binding WW domains[76]. The constrained flexibility of proline within PPXY motifs is important for reducing the entropic cost of binding to WW domain partners[70]. The first WW domain structure was determined for YES-associated protein bound to a 10-residue PPXY peptide with the sequence: GTPPPPYTVG[77]. Later structures showed that other WW domain-PPXY complexes adopt similar canonical structure with shared features. The three antiparallel β -strand lengths are slightly variable, but strands 2 and 3 tend to be longest and shortest, respectively. Each strand contains or is flanked by aromatic residues, with strand 1 and 3 containing the conserved tryptophan residues and strand 2 containing tyrosine residues. The strand-separating loops tend to be rigid and well-defined. The β -strands are slightly twisted and bent, forming concave and convex sides of the WW domain (Fig. 1.3B).

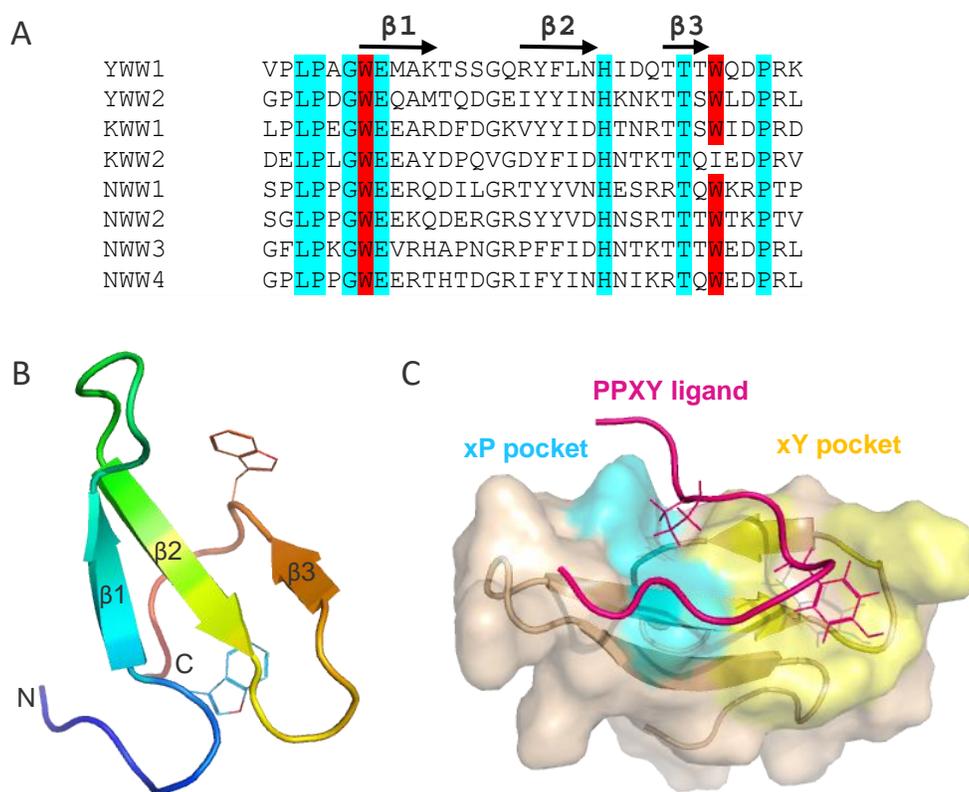


Figure 1.3: Representative WW domain sequences and structures. (A) Sequence alignment of the WW domains discussed in this dissertation: YAP (YWW1 and YWW2), KIBRA (KWW1 and KWW2), and NEDD4-1 (NWW1, NWW2, NWW3, and NWW4). Residues highlighted in cyan are conserved in all eight domains. The two highly conserved tryptophan residues are highlighted in red. The second tryptophan position in the KWW2 is replaced by an isoleucine. The approximate β -strand positions are indicated with arrows above the alignment. (B) Apo structure of the second WW domain of YAP (PDB: 2L4J), with the N terminus shown in blue and the C terminus shown in red. The highly conserved Trp sidechains are shown. Structural view is from the concave face. (C) Structure of the second WW domain of YAP (wheat) bound to a PPPY peptide of SMAD7 (hot pink), (PDB: 2LTV). The xP and xY pockets are shown in cyan and yellow, respectively. The side chains of the second Pro and Tyr of the PPPY motif are shown to indicate their positions in the xP and xY pockets, respectively.

The WW domain termini, sidechains within loops 1 and 2, and the first tryptophan form an exposed hydrophobic buckle on the convex side important for domain stability[78]. Loop 2 and the aromatic side chains within strands 2 and 3 form a shallow, hydrophobic binding surface on the concave side important for substrate recognition. This binding surface contains two pockets, the xP pocket containing the

second Trp, which packs against the second Pro of the PPXY motif, and the xY or specificity pocket, which packs against and hydrogen bonds with the Tyr of the PPXY motif and is critical for achieving motif specificity (Fig. 1.3C) [79].

Affinity and specificity of WW domain-PPXY motif interactions. Given the abundance of WW domains and PPXY motifs in the proteome and their vital roles in cell signaling, specificity of these interactions is critical for proper cell function. However, WW domains are degenerate, have small, relatively featureless binding interfaces and share similar structural features[80]. Thus, how WW domain and PPXY motif proteins achieve specificity is a major question in the field. In initial binding screens, isolated WW domains and PPXY motif peptides were shown to bind promiscuously[81]. Of the 65 WW domains and 1056 functional PPXY motifs evaluated, 632 motifs interacted moderately with 4-30 WW domains, indicating low specificity. However, this study was conducted with small protein fragments *in vitro*, which does not account for the context provided by the rest of the protein or the cell. A later study failed to recapitulate the binding observed between the WW2 domain of WW domain-containing oxidoreductase (WWOX) to the 16 identified partners in this high throughput screen when conducted with biological protein concentrations[82]. This suggests that factors beyond the isolated interaction sites are important for achieving specificity.

For some WW domain-PPXY motif interactions, features which confer specificity have been identified. First, WW domains and PPXY motifs often occur as homovalent sets within a single protein. The presence of multiple domains/motifs can lead to varying effects on binding as described previously in this chapter. In some cases, WW domain multivalency results in negative cooperativity. For example, in a study characterizing binding of the isolated and tandem WW domains of YES-associated protein (YAP, discussed below) to monovalent PPXY motif peptides of 10 proteins, affinities of the isolated domain interactions were roughly two-fold tighter than the tandem domains[83]. This negative cooperativity was attributed to intramolecular domain interactions in the apo form. However, most multivalent WW domain-PPXY motif interactions exhibit varying degrees of positive cooperativity. Examples include the tandem WW domains of Kidney and Brain-expressed protein (KIBRA, discussed

below) binding multivalent PPXY motif-containing partners PTPN14, Angiomin (discussed below), Expanded, Synaptopodin, and Dendrin[3, 84, 85], where both WW domains and two PPXY motifs are needed to form tight complexes.

In systems containing multiple WW domain/PPXY motif copies, linker length and structure are also factors which can impact binding. The tandem WW domains of yeast splicing factor, Prp40 are separated by a 10-residue α -helical linker, causing the domains to act as a single, modular unit with fixed binding orientations[86]. This rigid domain orientation likely favors binding of some PPXY partners over others, in a linker length and conformation-dependent manner. Structural characterization of the tandem WW domain of KIBRA in complex with various bivalent PPXY motif partners revealed that the domains behave as a supramodule, dependent on linker and C-terminal helical extension interactions[3]. This bivalent but interconnected behavior is key to forming high affinity complexes, as disruption of linker interactions decreased binding affinity. In this case, a PPXY motif linker of two residues was shown to be critical in forming a tight complex with KIBRA. In contrast, some systems require linker flexibility to form stable complexes. For example, human formin-binding protein 21 (FBP21) involved in pre-mRNA splicing contains two tandem class III WW domains which bind the PRMs of splicing factor SIPP1[87]. The FBP21 WW domains are separated by a 12-residue, highly flexible linker important for multivalent substrate recognition, likely due to orientation-independent binding at the two domains.

The residues flanking WW domains and PPXY motifs are also important in achieving binding affinity and specificity. A study which clearly exemplifies this examined binding affinities of NEDD4 WW3 (discussed below) to eight PPXY peptides. Peptides of p53bp2 with the sequence EYPPYPPPYPSG, two peptides of HTLV-1 Gag viral protein with sequences SDPQIPPPYVEP and PPPYVEP, and the core PPPY motif peptide in isolation bound NEDD4 WW3 with affinities of 5.3, 61, 178, and 210 μ M, respectively[88]. This study illustrates that while the PPXY motif in isolation is recognized by WW domains, the presence and identity of residues on either side of the motif are important in forming a stable complex. While there are fewer examples in the literature, WW domain-flanking residues are also important. For example, an isoleucine residue adjacent to the WW domain of YAP1 interacts with the

hydrophobic surface of the domain, stabilizing its structure[77]. Another example is the α -helical extension C-terminal to the second WW domain of KIBRA, deletion of which decreases KIBRA's affinity for its partners[3].

Allosteric changes and/or interactions with adjacent regions are present in some WW domain-PPXY motif interactions. For example, the Dystrophin WW domain interaction with the β -Dystroglycan PPXY motif, important in signal transduction from the cytoskeleton to the extracellular matrix, requires additional interactions from the adjacent Dystrophin EF-hand motif[89]. Some systems utilize multiple methods to finetune affinity and selectivity. The *Drosophila* transcriptional coactivator, Yorkie and transcription factor Tondu domain-containing growth inhibitor (Tgi), which contain two WW domains and three PPXY motifs, respectively, bind with positive cooperativity[90]. Here, one PPXY motif serves as the binding initiation site but requires contributions from the others to achieve full affinity. In addition to cooperativity, a stretch of residues outside the cognate binding sites within Tgi was seen to participate in binding, suggesting a form of allosteric regulation.

WW domains and PPXY motifs in cell signaling and disease

WW domains and PPXY motifs are degenerate, both within specific cellular contexts, and within individual protein molecules, where some mammalian proteins contain up to four WW domain or PPXY motif copies[91]. However, one signaling pathway exemplifies this degeneracy above all others, as it contains at least five proteins with WW domains and 14 proteins with PPXY motifs: the Hippo signaling pathway[91, 92]. This pathway is a key regulator of cell growth, proliferation, and apoptosis and relies on a phosphorylation cascade mediated by multivalent WW domain-PPXY motif interactions. Protein-protein interactions in cell signaling such as the Hippo signaling pathway must be specific, tunable, and responsive for proper function, indicating that WW domain-PPXY motif interactions have evolved mechanisms to achieve partner selectivity. Dysregulation of the Hippo signaling pathway is implicated in a host of cancers, making it a desirable candidate for mechanistic investigation. The following sections describe the functions, regulation, and disease implications of four homovalent WW domain and PPXY motif proteins

from the Hippo signaling pathway investigated in this body of work: YES-associated protein, Kidney and Brain-expressed protein, Angiomotin-Like 1, and neural precursor cell-expressed developmentally downregulated 4-1 (Table 1.1).

Table 1.1: Valency of proteins discussed in this work.

Protein	WW domains (#)	PPXY motifs (#)
YES-associated protein (YAP)	2	-
Kidney and Brain-expressed protein (KIBRA)	2	-
Angiomotin-Like 1 (AMOTL1)	-	3
neural precursor cell-expressed developmentally downregulated 4-1 (NEDD4-1)	4	-

YAP. YES-associated protein (YAP) was first identified in 1994 by Marius Sudol and named for its ability to bind the Src homology domain 3 (SH3) domain of the oncogenic tyrosine kinase, YES protein, through a PVKQPPPLAP motif[93]. There are eight isoforms of YAP, which fall into two primary groupings: YAP1-1 (450-470 residues, YAP1) and YAP1-2 (488-508 residues, YAP2), which contain 1 and 2 WW domains, respectively[94, 95]. YAP is composed of a conserved, modular domain architecture consisting of a disordered, proline-rich N-terminal region followed by a transcription enhancer factor domain (TEAD)-binding domain, 1-2 WW domains, an SH3-binding motif, a coiled-coil (CC) domain, and a C-terminal transactivation domain (TAD), followed by a post synaptic density protein domain (PDZ)-binding motif (Fig. 1.4)[96, 97]. YAP is the prototype for studying WW domains. A closely related, 400-residue paralog of YAP is transcriptional coactivator with PDZ-binding motif (TAZ) which arose from a gene duplication event in vertebrates[96, 98]. TAZ shares a similar domain architecture with YAP but is missing the N-terminal pro-rich region, the SH3-binding motif, and contains only one WW domain like YAP1-1.

YAP and TAZ are transcriptional co-activators which primarily bind transcription factors TEAD1-4, but can also interact with SMADs, EGR-1, ERBB4, RUNXs, and p73[99]. YAP/TAZ upregulate transcription of genes directly and indirectly involved in cell growth and proliferation, DNA replication and repair, apoptosis, mitosis, transcriptional regulation, and response to reactive oxygen

species[100]. These genes include AXL, CTGF, CYR61, IAP, ANKRD1, and BCL2. YAP/TAZ function is critical for cell differentiation and migration, proper organ size, embryonic STEM cell development, inflammatory response, neurological function, and wound healing[101-104]. The vital nature of these cellular processes requires tight regulation of YAP and TAZ to achieve healthy cellular function and development.

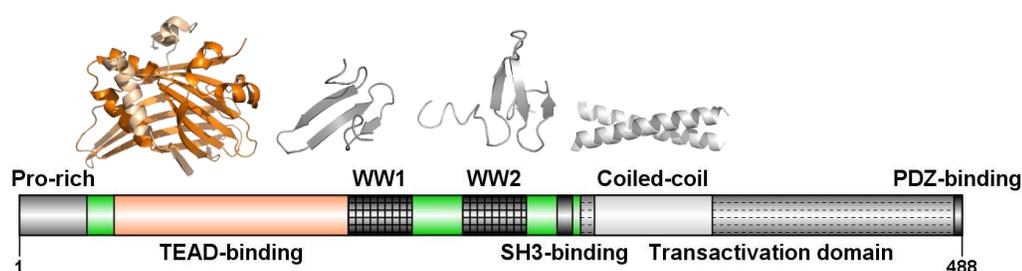


Figure 1.4: Domain architecture of YES-associated protein. Domain organization of human YAP showing (from left to right) the proline-rich region, the TEAD-binding domains, WW1, WW2, SH3-binding motif, the transactivation domain, and the coiled-coil domain[1]. Known structures are shown for the YAP TEAD-binding domain (wheat) bound to TEAD (orange), (PDB: 3KYS); the WW1 domain (PDB: 2LTW); the WW2 domain (PDB: 2L4J); and a generic coiled-coil domain (PDB: 1P9I).

The primary mode of YAP/TAZ regulation occurs through control of their subcellular localization. This is achieved through phosphorylation of critical serine residues of YAP/TAZ mediated by a cassette of kinases which compose the Hippo signaling pathway (Fig. 1.5). This evolutionarily conserved phosphorylation cascade was first identified in *Drosophila* in the 2000s and named for the hippopotamus-like overgrowth phenotype of fruit flies with mutations in the Hippo kinase gene[105]. The Hippo signaling pathway contains a set of core kinases, upstream regulators, downstream effectors, and crosstalk with related signaling pathways including the Wnt pathway[106], autophagy[107], G-protein-coupled receptor (GPCR) signaling[108], and bone morphogenetic protein signaling.

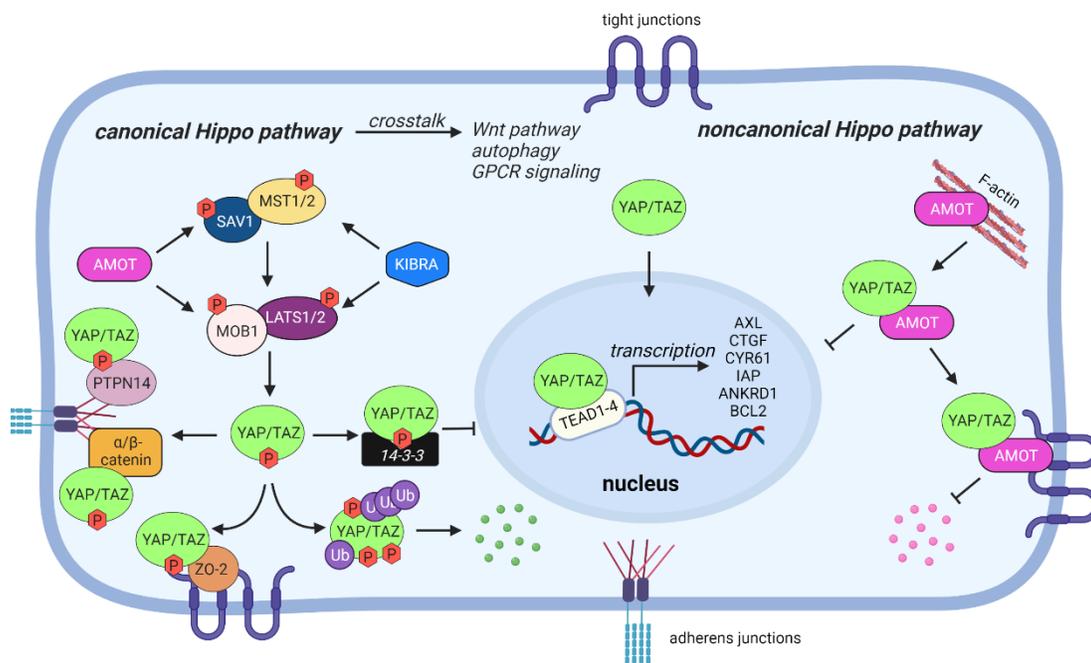


Figure 1.5: Canonical and noncanonical regulation of YAP/TAZ through the Hippo signaling pathway. Pictured on the left is the canonical Hippo signaling pathway, which includes the phosphorylation cascade proteins SAV1, MST1/2, MOB1, LATS1/2, and YAP/TAZ. This cascade is also regulated by scaffolding proteins AMOT and KIBRA. Phosphorylated YAP/TAZ are localized to the cytoplasm through sequestration or degraded. Pictured on the right is the phosphorylation-independent or noncanonical Hippo signaling pathway. YAP/TAZ can bind AMOT independent of their phosphorylation state and sequester these transcriptional coactivators to TJs. YAP/TAZ and F-actin compete for binding to AMOT, and YAP/TAZ-binding prevents degradation of AMOT.

The canonical Hippo signaling pathway consists of a core phosphorylation cascade which starts with mammalian STE20-like protein kinase 1 and 2 (MST1/2), the mammalian ortholog of the *Drosophila* kinase, Hippo. MST1/2 forms a heterodimer with Salvador 1 (SAV1). The MST1/2-SAV1 complex then binds and phosphorylates large tumor suppressor 1 and 2 (LATS1/2) kinase with the assistance of the adaptor protein, MOB kinase activator 1A and B (MOB1A/B). LATS1/2 then binds and phosphorylates YAP and TAZ at S127 and S89, respectively, along with numerous other sites less critical to the regulation of their subcellular localization. Phosphorylated YAP/TAZ are clients of proteins including 14-3-3, which binds and sequesters these coactivators in the cytoplasm; PTPN14 and α/β -catenin at adherens junctions (AJs); and ZO-2, which localized YAP/TAZ to tight junctions (TJs),

preventing the upregulation of growth activator genes in the nucleus[109]. Furthermore, phosphorylated YAP/TAZ are recognized and further phosphorylated by Casein kinase 1 δ/ϵ and ubiquitinated by SCF E3 ubiquitin ligase resulting in degradation[110].

Due to their vital roles in controlling cell growth and related processes, dysregulation of YAP/TAZ, either through mutations in these proteins or upstream regulators, is implicated in countless diseases, particularly numerous types of cancer[111]. YAP/TAZ inhibit apoptosis in cancers including gastric, renal, colon, liver, pancreatic, breast, gliomas, leukemia, and carcinomas[96]. However, in some contexts, YAP and TAZ can function as tumor suppressors in breast, lung, and liver cancer by eliciting an apoptotic response. In addition to cancer, YAP is involved in several other diseases. For example, overactivation of Hippo signaling and reduced YAP expression are correlated with arrhythmogenic cardiomyopathy and aortic dissections, respectively[111].

KIBRA. Kidney and Brain-expressed protein (KIBRA), otherwise known as WW and C2 domain containing 1 (WWC1), was first identified in 2003 and is the most conserved member of the WWC family which also contains paralogs WWC2 and WWC3[112]. As its name suggests, KIBRA expression is enriched in the kidneys and brain tissue, but it is also expressed in testis, breast tissue, gastric tissue, heart, colon, and pancreas. The domain architecture of this 125 kDa protein includes two N-terminal WW domains, a phospholipid-binding C2 domain, coiled-coil domains, a glutamate-rich region, and a C-terminal PDZ-binding motif (Fig. 1.6)[113]. KIBRA is primarily cytosolic but can also be found in the nucleus. KIBRA is an upstream activator of the Hippo signaling pathway, whereby it can synergize with NF2 and FRMD6 to trigger phosphorylation and activation of LATS1/2 independent of MST1/2 or bind directly to the PPXY motifs of LATS1/2 to protect it from degradation (Fig. 1.5)[114, 115].

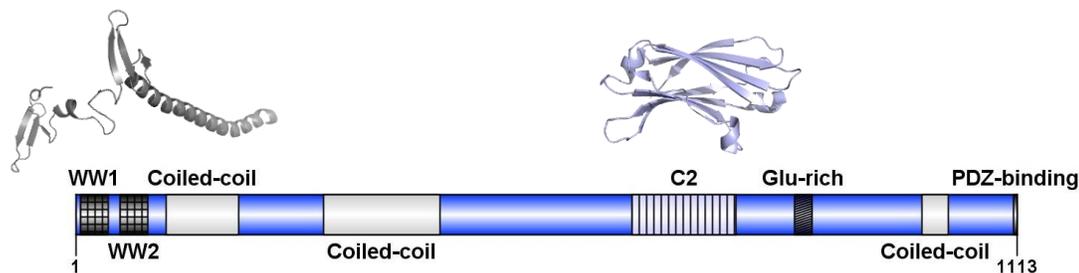


Figure 1.6: Domain architecture of Kidney and Brain-expressed protein. Domain organization of human KIBRA showing (from left to right) the tandem WW domains, coiled-coil domains, C2 domain, glutamate-rich region, and PDZ-binding motif[1]. Known structures are shown for the tandem WW domains and adjacent coiled-coil domain (PDB: 6JJX); and the C2 domain (PDB: 2Z0U).

AMOTL1. Angiomotin (AMOT) family proteins were first discovered in 2001 as regulators of endothelial cell migration through mediation of angiostatin – an inhibitor of angiogenesis and cell growth – in a placenta yeast two-hybrid screen for angiostatin-binding proteins[116]. The angiomotin family contains four members with high sequence homology: AMOT-p80 (675 residues), AMOT-p130 (1084 residues), Angiomotin-Like 1 (AMOTL1, 956 residues, otherwise known as junction-enriched and associated protein – JEAP), and Angiomotin-Like 2 (AMOTL2, 780 residues, otherwise known as MAGI-1-associated coiled-coil tight junction protein – MASCOT)[117-119]. Although it was the last to be identified, AMOT-p130 is considered the parent angiomotin isoform and is the most studied[120]. AMOT-p130 is an alternatively spliced isoform of AMOT-p80 and contains an additional 409 N-terminal residues which are also present in AMOTL1 and AMOTL2 and contains 2-3 N-terminal PPXY motifs, which mediate protein-protein binding interactions[117]. All four isoforms contain middle coiled-coil (CC) and Bin/Amphiphysin/Rvs (BAR) domains and C-terminal PDZ-binding motifs. AMOT-p130 and AMOT-p80 contain angiostatin-binding domains which are not present in AMOTL1 or AMOTL2. Although the other angiomotin proteins are discussed briefly, the focus of this work is on AMOTL1, which contains three PPXY motifs with sequences LPTY, PPEY, and PPEY (Fig. 1.7).

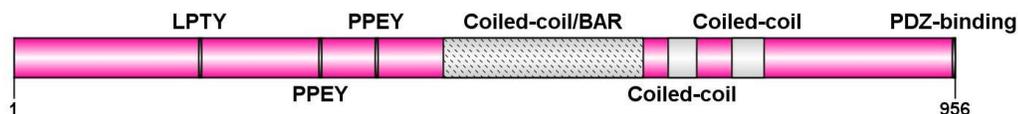


Figure 1.7: Domain architecture of Angiomotin-Like 1 protein. Domain organization of human AMOTL1 showing (from left to right) the three PPXY motifs (LPTY, PPEY, and PPEY), and coiled-coil/BAR domain, two additional coiled-coil domains, and a C-terminal PDZ-binding motif[1].

Despite their high sequence and domain homology, the angiomotin family proteins show differential spatiotemporal expression, function, and binding[117]. Human mRNA expression patterns show the highest expression of AMOT in testis, brain, and thyroid tissue and lowest expression in adrenal glands and liver tissue[117]. AMOTL1 expression was highest in skeletal muscle and lowest in the blood. AMOTL2 expression was highest in breast tissues and lowest in the liver. These expression patterns are also shown to vary at different stages of development[121-123].

Angiomotin proteins play important roles in cell growth regulation through several mechanisms. As a scaffold, AMOT(L1/2) help facilitate phosphorylation of YAP through Hippo signaling by simultaneously binding YAP, LATS, MOB1, and SAV1, which then binds MST1/2 (Fig. 1.5)[33]. This complex enhances phosphorylation of Hippo signaling components through proximity, which results in more efficient YAP phosphorylation and downregulation of cell growth. AMOTs can also regulate YAP subcellular localization independent of YAP phosphorylation through direct WW domain-PPXY motif interactions which sequester YAP in the cytosol or to TJs, a phenomenon known as noncanonical Hippo signaling (discussed above)[124].

AMOT(L1/2) cellular concentrations are mediated by ubiquitination and proteasomal degradation. This process is finely regulated by a variety of proteins. USP9x is a deubiquitylating enzyme that removes ubiquitin from K496 of AMOT, resulting in increased protein stability[125]. AMOTL1 degradation is also mediated by direct interactions with YAP, which blocks the interaction with ubiquitin ligases through competitive binding[126]. Intriguingly, not all AMOT-ubiquitin ligase interactions are negative. AMOTL1 interacts with HECW2 E3 ubiquitin ligase, but this

interaction results in increased AMOTL1 stability[127]. Similarly, AMOTs interact through WW domain-PPXY motif interactions with KIBRA, which results in decreased AMOT degradation in hepatic cells[2].

NEDD4. Neural precursor cell-expressed developmentally downregulated 4-1 (NEDD4-1) belongs to the family of HECT E3 ubiquitin ligases and was first discovered in 1992 and named for its downregulation in developing mouse brain tissue[128]. The HECT E3 ligase family contains nine members in mammals: NEDD4-1 (NEDD4), NEDD4-2 (NEDD4L), NEDL1, NEDL2, SMURF1, SMURF2, WWP1, WWP2, and ITCH[129]. All nine members contain a conserved modular domain architecture composed of an N-terminal Ca^{2+} /lipid-binding C2 domain, 2-4 WW domains responsible for substrate recognition, and a C-terminal catalytic HECT domain responsible for substrate mono/polyubiquitination[130]. NEDD4 family proteins are highly conserved from yeast to humans[129]. While the need for nine NEDD4 E3 ligases is not fully understood, it is thought that differing expression, localization, and substrate specificity underly their different functions[130]. The founding member, NEDD4, is 120 kDa, contains four WW domains, and will be the focus of further discussion (Fig. 1.8).

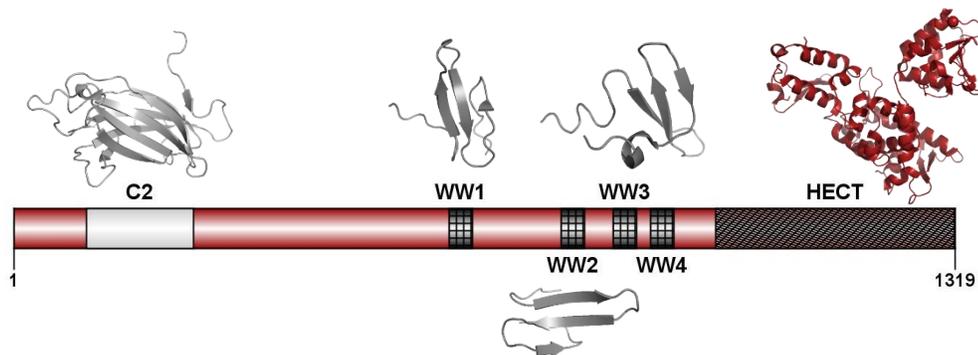


Figure 1.8: Domain architecture of NEDD4-1 protein. Domain organization of human NEDD4-1 showing (from left to right) the C2 domain, four WW domains, and the catalytic HECT domain. Known structures are shown for the C2 domain (PDB: 3B7Y), WW1 from rat NEDD4 (PDB: 2N8S), WW2 from NEDD4L (PDB: 2LYT), WW3 (PDB: 5AHT), and the HECT domain (PDB: 2XBF).

NEDD4 is a ubiquitously expressed, cytosolic protein involved in a plethora of vital cellular functions including regulation of ion channels, endocytosis and membrane

transport, immunity, autophagy, cell growth, proliferation, motility, and signaling[129, 131]. Unlike YAP and KIBRA, NEDD4 is not directly involved in the Hippo signaling pathway. However, NEDD4 indirectly regulates Hippo signaling (among other pathways) by regulating the homeostasis of proteins in the pathway through ubiquitination[129]. Ubiquitination is an important mechanism for regulation of protein degradation, discovery of which was awarded the 2004 Nobel Prize in Chemistry to Ciechanover, Hershko, and Rose. Now, there are > 600 known human E3 ubiquitin ligases[132]. Ubiquitination is a PTM that involves the ATP-dependent transfer of the 8 kDa protein ubiquitin (Ub) to lysine residues of substrates[133]. This reaction involves three steps: (1) Ub activation through the ATP-dependent covalent attachment of Ub to a cysteine residue of an E1 ligase through a thioester bond; (2) transfer of activated Ub to an E2 ligase through a second thioester bond; and (3) transfer of Ub from E2 to the substrate mediated by an E3 ligase, resulting in covalent attachment of Ub through the C-terminal G76 to the Lys sidechain of the substrate.

NEDD4 plays a role in a plethora of disease states including various cardiovascular diseases[134], neurodegenerative diseases[135, 136], cancers[137], and viral infection[138]. The role of NEDD4 in tumorigenesis is complex and context-dependent[131]. In the Hippo signaling pathway, NEDD4 acts as a cell growth activator by ubiquitinating upstream regulators of YAP/TAZ. For example, the WW domains of NEDD4 interact with the PPXY motifs of LATS1 kinase, leading to its ubiquitination, degradation, and the subsequent enhanced nuclear localization of YAP[139]. NEDD4 also binds and degrades SAV1 in a WW domain-PPXY motif-independent manner[140]. However, this reaction is inhibited by interactions between SAV1 and MST1. Additionally, NEDD4, NEDD4L, and ITCH E3 ubiquitin ligases bind AMOT(L1/2) PPXY motifs through their WW domains, which results in AMOT(L1/2) dissociation from tight junctions (TJs), poly-ubiquitination, and proteasomal degradation[126, 141]. Decreased cellular concentrations of AMOT(L1/2) results in the concomitant increase in YAP nuclear localization.

In viral infection, NEDD4 can be hijacked by viral proteins and assist in important roles in the viral life cycle of many types of viruses including retroviruses (HIV, RSV, HTLV-1, and MLV), filoviruses (Ebola and Marburg), adenoviruses, and

SARS-CoV-2 [138, 142]. Many viral proteins contain regions called Late (L)-domains, which contain SLiMs with sequences PPXY, PT/SAP, or YXXL[138] used to bind host proteins. However, not all NEDD4-mediated viruses contain a L-domain. Instead, some require adaptor proteins to partition NEDD4 and viral particles. Viruses use the ubiquitin system in cell entry, to evade immune response, enhance viral replication, and for viral particle budding and escape from the cell by usurping the ESCRT pathway[143]. Understanding the mechanisms behind substrate binding and specificity of NEDD4 and its partners is critical for developing therapeutics to combat this broad range of diseases.

Contents of dissertation

Interactions of AMOTL1 with YAP, KIBRA, and NEDD4 result in different downstream effects in the cell. However, all three interactions are mediated through the PPXY motifs of AMOTL1 and the WW domains of its partners. In the work presented herein, I characterize the multivalent binding interactions of AMOTL1 to YAP, KIBRA, and NEDD4 to address the following questions: (1) how do these multivalent complexes assemble in context of all putative domains and motifs? (2) Do binding mechanisms vary from partner to partner? And (3) How does AMOTL1 discriminate between different partners present in the same environment? Answers to these questions will help uncover how selective interactions occur in a network of degenerate recognition sequences, and how these interactions have evolved to respond to the needs of the cell. Dissecting these mechanisms will also reveal information needed to manipulate these protein-protein interactions with therapeutics to treat a myriad of disease states.

In Chapter 2, we examine the AMOTL1-YAP interaction and show that binding occurs through a dynamic ensemble of transient complexes. This transient nature is conferred by a combination of stabilizing and destabilizing interaction sites, which results in overall attenuated affinity when all sites are present. Although dynamic protein interactions are critical in cell signaling, characterizing these interactions has proven to be a challenge in the field. Our work provides experimental evidence that multivalent disordered complexes can assemble via dynamic interactions.

In Chapter 3, we characterize the AMOTL1-KIBRA interaction. Unlike the AMOTL1-YAP interaction, stability of this complex relies on synergy between the WW domains of KIBRA. Furthermore, we show that binding is initiated by stable and selective interactions at one PPXY site, followed by destabilizing interactions with adjacent sites. Alternate site specificity allows stable formation of a ternary complex composed of YAP, AMOTL1, and KIBRA. This work provides mechanistic insight into the scaffolding action of AMOTL1.

In Chapter 4, we transition from studying proteins with tandem WW domains to NEDD4-1 with four WW domains. Using clever construct design and segmental domain labeling approaches, we both characterize the solution properties of the complete WW domain segment of NEDD4-1 and show that the presence of adjacent domains modulates these solution properties. This work provides insight into the interconnectivity between WW domains and how multivalency serves to modulate domain behavior, which may be an important feature for finetuning substrate recognition.

In Chapter 5, we expand on our findings from Chapter 4 and characterize the binding mechanism of NEDD4-1 and AMOTL1. Here, complex formation relies on specific domain-motif pairings and dynamic assembly involving WW domain cooperativity. This work confirms that the tetravalent domains of NEDD4-1 are functionally active and important for complex assembly. This work highlights the versatility of multivalency and disorder in the assembly of protein complexes in signaling pathways.

In Chapter 6, I summarize the biological implications of this work and define questions which remain to be addressed. Additional unfinished work is described in Appendices.

Chapter 2

Multivalent Angiotensin-like 1 and Yes-associated protein form a dynamic complex

Amber Vogel, Alexandra Crawford, and Afua Nyarko

Published in *Protein Science* (2022), 31(5), e4295

Copyright © 2022 The Protein Society

Abstract

Multivalent complexes formed between the cancer-promoting transcriptional co-activator, Yes-associated protein (YAP), and proteins containing short linear motifs of type PPXY modulate cell proliferation and are attractive therapeutic targets. However, challenges producing PPXY polypeptides containing the full binding domain has limited understanding of the assembly process. Here, we successfully produced a polypeptide containing the complete set of three PPXY binding sites of Angiomotin-like 1 (AMOTL1), a scaffolding protein that regulates the nucleo-cytoplasmic shuttling of YAP via WW-PPXY interactions. Using an array of biophysical techniques including isothermal titration calorimetry, size-exclusion chromatography coupled to multi-angle light scattering, and solution nuclear magnetic resonance spectroscopy, we show that the AMOTL1 polypeptide is partially disordered, and binds the YAP WW domains to form an ensemble of complexes of varying stabilities. The binding process is initiated by the binding of one YAP WW domain to one AMOTL1 PPXY motif and is completed by transient interactions of the second YAP WW domain with a second AMOTL1 PPXY motif to form an equilibrating mixture composed of various species having two YAP sites bound to two conjugate AMOTL1 sites. We rationalize that the transient interactions fine-tune the stability of the complex for rapid assembly and disassembly in response to changes in the local cellular environment.

Abbreviations

CD, circular dichroism spectroscopy; HSQC, heteronuclear single quantum coherence; ITC, isothermal titration calorimetry; NMR, nuclear magnetic resonance spectroscopy; WT, wild type.

Introduction

Yes-associated protein (YAP) is a transcriptional co-activator which promotes cell survival by activating cell proliferation and anti-apoptosis[144]. YAP-mediated processes are linked to various cancers in which cellular levels and nuclear localization of YAP are increased, and subsequent interactions of DNA-binding transcription factors with YAP are important first steps[145]. Given their central role in cell survival, we have undertaken foundational structure-function studies to elucidate YAP regulatory mechanisms. Results will elucidate fundamental scientific questions as well as inform the design of novel small molecules for therapeutic intervention.

Multiple proteins are implicated in the regulation of YAP cellular concentrations and localization. Among them are the Motin family of proteins Angiomotin (AMOT), Angiomotin-like 1 (AMOTL1) and Angiomotin-like 2 (AMOTL2) which function in cell differentiation, proliferation, and migration[33, 123, 146-148]. The three proteins bind directly to YAP to regulate YAP-promoted cell proliferation[124, 147, 149-152]. Key to Motin interactions with YAP are 2-3 short linear motifs (SLiMs) of type L/PPXY (hereafter referred to as PPXY) which recognize tandem YAP WW domains, autonomous folding units characterized by the presence of two conserved tryptophans[60]. Here, we focus on molecular interactions between YAP and AMOTL1.

AMOTL1 is a 106 kDa protein with three PPXY motifs (P1, P2 and P3) which precede three structured domains: a Bin/Amphiphysin/Rvs (BAR) domain, a coiled-coil and a post synaptic density protein (PDZ)-binding domain[117, 146, 153] (Fig. 2.1). High structural disorder is predicted for the PPXY segment and challenges producing large primarily disordered fragments may explain why biophysical studies of fragments containing all three PPXY sites have been limiting. An attractive workaround to produce shorter polypeptides with one or two PPXY sites has led to conflicting reports; some studies suggest that P2 is the most relevant motif [124, 147, 151]; and others demonstrate that both P1 and P2 sites function in complex formation [149].

Our experiments resolve the question of which PPXY motifs are required for *in vitro* assembly of the YAP-AMOTL1 complex. A battery of formidable and mutually

reinforcing molecular biophysical methodologies are employed, and the experimental results are collectively interpreted in the context of the function of the YAP-AMOTL1 complex.

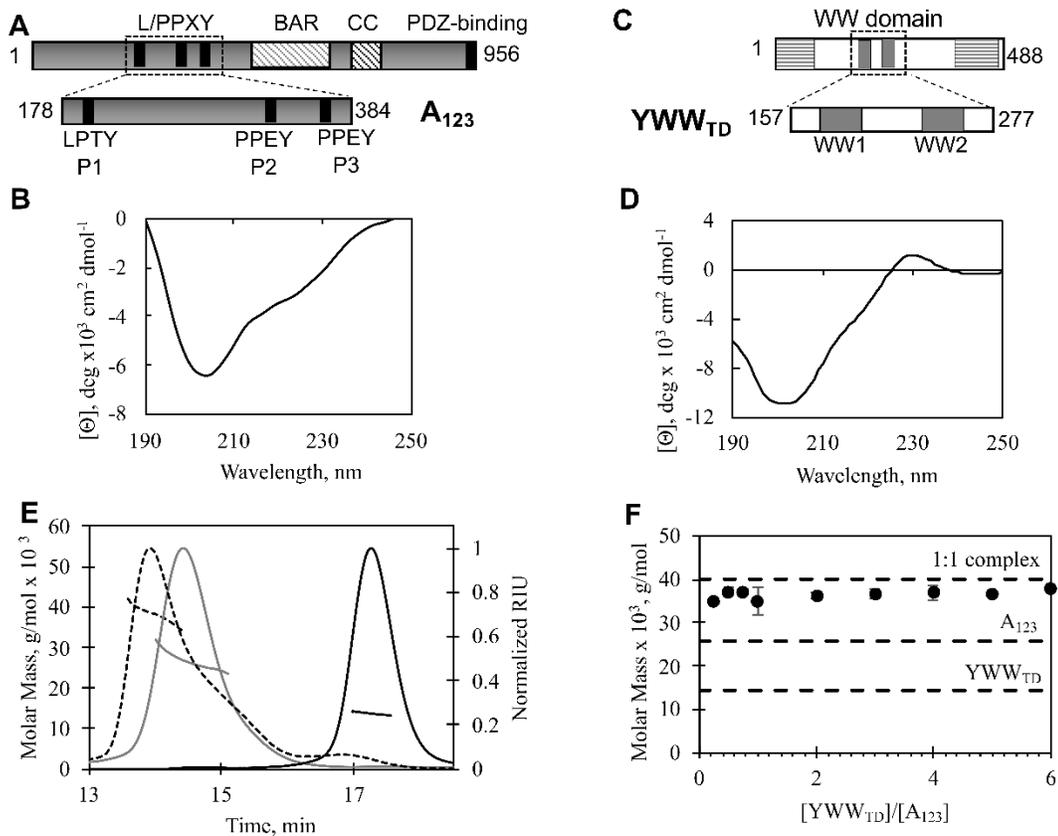


Figure 2.1: Domain architecture and solution properties of apo and bound AMOTL1 and YAP polypeptides. (A) Full-length AMOTL1 (956-residues) contains three PPXY motifs which precede, a Bin/Amphiphysin/Rvs (BAR), coiled-coil (CC) and a postsynaptic density protein (PDZ)-binding domains. The AMOTL1 A₁₂₃ construct (residues 178–384) includes all three PPXY sites. (B) Far UV CD spectrum of A₁₂₃ is consistent with a partially disordered polypeptide. (C) YAP, a multidomain protein, contains two WW domains, WW1 and WW2. The YAP construct, YWW_{TD} (residues 157–277) was designed to include both WW domains. (D) Far UV CD spectrum of YWW_{TD} shows the characteristic spectrum of a WW domain. (E) SEC-MALS elution profiles of A₁₂₃ (gray line), YWW_{TD} (black line), and a 1:1 mixture of both proteins (dashed black line). (F) A plot of the SEC-MALS-detected molar mass of varying ratios of A₁₂₃:YWW_{TD} complexes. The average molar mass values range between 34.9–37.8 kDa. The theoretical molar masses of A₁₂₃, YWW_{TD}, and the 1:1 A₁₂₃-YWW_{TD} complex are indicated by dashed lines.

Results

SEC-MALS and CD experiments of apo and bound proteins. In these experiments we use A₁₂₃, a multivalent 207-residue AMOTL1 polypeptide (residues 178-384) with three PPXY motifs – LPTY, PPEY and PPEY (Fig. 2.1A). The far ultraviolet (UV) circular dichroism (CD) spectrum of A₁₂₃ (Fig. 2.1B) shows a strong signal at 204 nm, suggestive of random coil-like structure, and a relatively weak signal at 222 nm indicative of nascent helical structure. The experimental CD spectrum was analyzed with the web server, DichroWeb [154, 155], to provide calculated secondary structure contents of 24.7% helix, 0.6 % strands, 15.9 % turns and 43.7 % disordered segments.

The YAP polypeptide, YWW_{TD}, (residues 157-277, Fig. 2.1C) contains both WW domains (WW1 and WW2). The far UV CD spectrum of YWW_{TD} (Fig. 2.1D) shows a positive peak at 230 nm and a negative peak at 220 nm which are signature WW domain peaks attributed to packing of the two tryptophans and the β -strands, respectively. Taken together, the CD data indicate that the A₁₂₃ polypeptide is partially folded and the YWW_{TD} polypeptide adopts the folded WW domain structure.

Multi-angle light scattering (MALS) measures the proportion of light scattered by an analyte and when combined with size exclusion chromatography can be used to determine molar mass without comparison to globular protein standards[156]. Completely or partially disordered proteins have relatively large hydrodynamic radii, which precludes the use of SEC alone to infer molecular mass, whereas SEC-MALS is particularly attractive for computing their molar mass. The average molecular mass of A₁₂₃ probed by SEC-MALS is 26.1 ± 1.2 kDa, a value close to the theoretical monomer molecular mass of 25.9 kDa. The SEC-MALS calculated molar mass of YWW_{TD} is 14.1 ± 1.1 kDa (expected monomer molar mass is 14.2 kDa) (Fig. 2.1E; Table 2.1). An equimolar mixture of A₁₂₃ and YWW_{TD} elutes earlier than unbound A₁₂₃ or YWW_{TD} with an SEC-MALS-measured molecular mass of 35 kDa, slightly less than the theoretical molar mass of 40.1 kDa expected for a 1:1 complex. When resolved on SDS-PAGE, the peak for the equimolar mixture migrates as two bands with migration rates similar to unbound A₁₂₃ and YWW_{TD} (data not shown). A₁₂₃ and YWW_{TD} complexes formed by mixing 0.25–6-fold molar excess of YWW_{TD} with A₁₂₃ give SEC-MALS measured molecular masses in the range 34.9–37.8 kDa (Fig. 2.1F), values

slightly less than the theoretical molar mass of 40.1 kDa for a 1:1 complex (Table 2.1). Taken together, the SEC-MALS data indicate that under the experimental conditions, the A₁₂₃ and YWW_{TD} polypeptides are monomeric in solution, and the YWW_{TD}-A₁₂₃ complex is formed between one YWW_{TD} molecule and one A₁₂₃ molecule.

Table 2.1: Experimental and theoretical molar masses of apo and bound AMOTL1 A₁₂₃ and YAP YWW_{TD}. *Reported value is average of triplicate experiments.

Protein	Stoichiometry	Molar mass (kDa)	
		<i>Experimental</i>	<i>Theoretical</i>
A ₁₂₃ (AMOTL1)	-	26.1 ± 1.2*	25.9
YWW _{TD} (YAP)	-	14.1 ± 1.1*	14.2
YWW _{TD} -A ₁₂₃ complex	0.25:1	34.9	-
	0.5:1	37.3	-
	0.75:1	37.2	-
	1:1	35 ± 3*	40.1
	2:1	36.4 ± 0.7*	54.3
	3:1	36.8	68.5
	4:1	37.0	82.7
	5:1	36.8	96.9
	6:1	37.8	111.1

Solution NMR experiments of apo and YAP-bound AMOTL1 A₁₂₃. NMR experiments were used to determine for A₁₂₃ the local secondary structure propensity, backbone dynamics and binding interface residues. The ¹H-¹⁵N heteronuclear single quantum coherence (HSQC) spectrum of A₁₂₃ shows poor dispersion in the NH-region, as expected for a protein with random coil-like and/or helical structure (Fig. 2.2A). Despite significant peak overlap, we assigned 79% (145 of 183 non-proline residues) of amide backbone nuclei, which enabled residue-specific analyses. Resonance assignments for the 26 additional residues that are part of the expression vector are not included in the analysis.

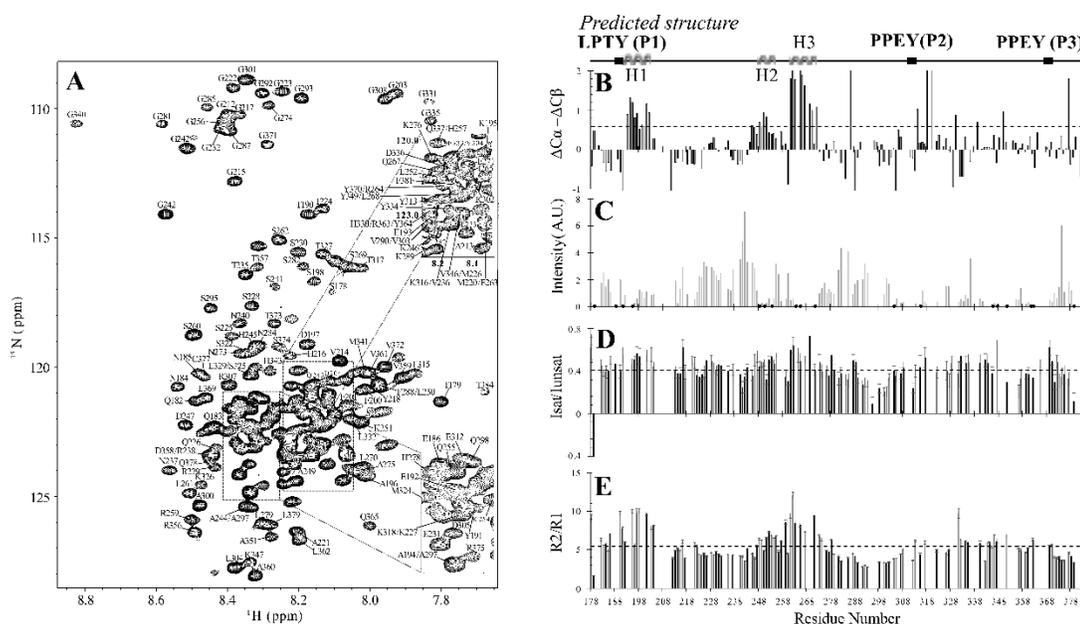


Figure 2.2: NMR backbone assignments and solution dynamics of AMOTL1 A₁₂₃ polypeptide. (A) ^1H - ^{15}N -HSQC of A₁₂₃ showing assignments for 145 of 183 non-proline residues. (B) A plot of the deviation of the chemical shifts from random coil values ($\Delta 13\text{C}\alpha - \Delta 13\text{C}\beta$). (C) CLEANEX-HSQC analysis of A₁₂₃. The gray bars represent the rapidly exchanging amide protons, and the black dots are the relatively slow exchanging amide protons. (D) A plot of the steady-state heteronuclear NOE (Het-NOE) values. (E) The ratio of the transverse and longitudinal relaxation ($R2/R1$) as a function of the residue number. Average values are shown as dashed lines in plots B, D, and E. The predicted secondary structure of A₁₂₃ is shown above the plots as a line diagram where disordered segments are represented by a solid line. Three predicted helices, H1 (residues 191–202) and H2 (residues 247–254), and H3 (residues 261–272) are located between the first (P1) and second (P2) PPXY sites.

PSIPRED[157], a sequence-based secondary structure prediction algorithm, shows that the A₁₂₃ polypeptide is largely disordered except for three short helical segments designated H1, H2 and H3 (Fig. 2.2). The experimental secondary structure propensities of A₁₂₃ were determined from the relative deviations of assigned CA and CB chemical shift values from standard random coil values ($\Delta\text{CA} - \Delta\text{CB}$). A plot of the $\Delta\text{CA} - \Delta\text{CB}$ as a function of the residue number (Fig. 2.2B), shows sequential and substantial positive values for residues 193–197, 261–263, 265–267, and 269–270, within the predicted helical segments H1 and H3. An additional stretch of residues (245–257) within the predicted H2 α -helical segment, have positive deviations slightly below the computed average value, suggestive of a relatively weak or nascent helix.

Using the backbone chemical shifts and the $\delta 2D$ analysis software[158], we determined the probability distribution of the secondary structural elements as 69.1% coil, 15.8% polypropylene II helix, 8.0% regular helix, and 7.1% beta strands, which is consistent with the conclusion from CD analysis that A_{123} , has limited folded structure.

To identify solvent-exposed residues in the A_{123} polypeptide, a clean chemical exchange (CLEANEX)-HSQC experiment which measures fast backbone amide proton exchange with water was recorded. The plot of assigned residues in Figure 2.2C shows residues with rapidly exchanging amide protons (grey bars) and solvent protected residues (black circles) which are not detected in the CLEANEX-HSQC spectrum (Fig. 2.3). As expected for a partially folded polypeptide, most residues in A_{123} are accessible to solvent, with the most rapidly exchanging amide protons (residues with the highest intensities) within the linker segments connecting the P1 and P2 sites. Consistent with a predicted helical segment, backbone amide protons for residues in the H2 segment are protected from solvent.

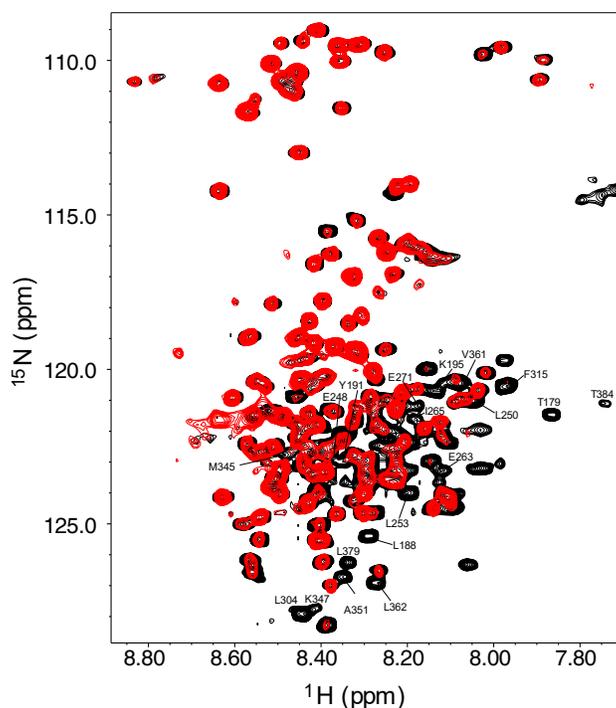


Figure 2.3: A_{123} solvent accessibility. 1H - ^{15}N TROSY- (black) and CLEANEX- (red) HSQC spectra of apo A_{123} . Labeled residues have relatively slow exchanging amide protons and are not detected in the CLEANEX-HSQC spectrum.

To determine backbone dynamics at multiple timescales, we measured a suite

of NMR experiments, including steady-state ^1H - ^{15}N heteronuclear NOE (HetNOE), ^{15}N longitudinal (R_1), and transverse (R_2) relaxation. HetNOE measurements are sensitive to the strength of the magnetic field and report backbone dynamics in the pico- to nano-second (ps-ns) timescale. Generally, positive and negative HetNOE values are indicative of restricted and mobile residues respectively, but at high magnetic field strengths, mobile residues may display low positive values[159]. HetNOE values are mostly positive with an average of 0.4 (Fig. 2.2D). At our experimental field strength of 800 MHz, HetNOE values below the average of 0.4 are indicative of mobile segments. The most mobile residues, 179 and 384 (negative NOE value), are at the N and C-termini. Intriguingly, the first residue of the native protein sequence (178) has an NOE value greater than the average value of 0.4. A possible explanation for this anomaly is that the non-native poly-histidine sequence adjacent to 178 restricts the motion of this residue.

The R_2/R_1 values reflect motions in the nanosecond timescale and identifies segments with slower tumbling. The plot in Figure 2.2E shows relatively high values for residues 190-198 and 259-271 which implies relatively slower tumbling for the two segments.

A_{123} residues involved in binding YWW_{TD} were mapped by NMR titration experiments in which unlabeled YWW_{TD} and isotopically labeled A_{123} were mixed at molar ratios in the range of 0.25:1–2:1 ($\text{YWW}_{\text{TD}}:A_{123}$). As a reporter of the binding interactions, we monitored changes in the intensities of peaks corresponding to the tyrosine residues, Y191, Y313, and Y370, at the 3 PPXY sites. The plot in Figure 2.4A shows a decrease in the intensity of each tyrosine peak with increasing concentrations of added YWW_{TD} . Note that peaks for all three tyrosine residues are not detected in the partially bound 0.75:1 ($\text{YWW}_{\text{TD}}:A_{123}$) complex. Loss in peak intensity is attributed to binding-induced exchange broadening, binding-induced conformational changes and/or slower tumbling of the complex[160]. Line broadening occurs, and peaks disappear when the rate of exchange between the bound and unbound conformations is intermediate on the NMR timescale.

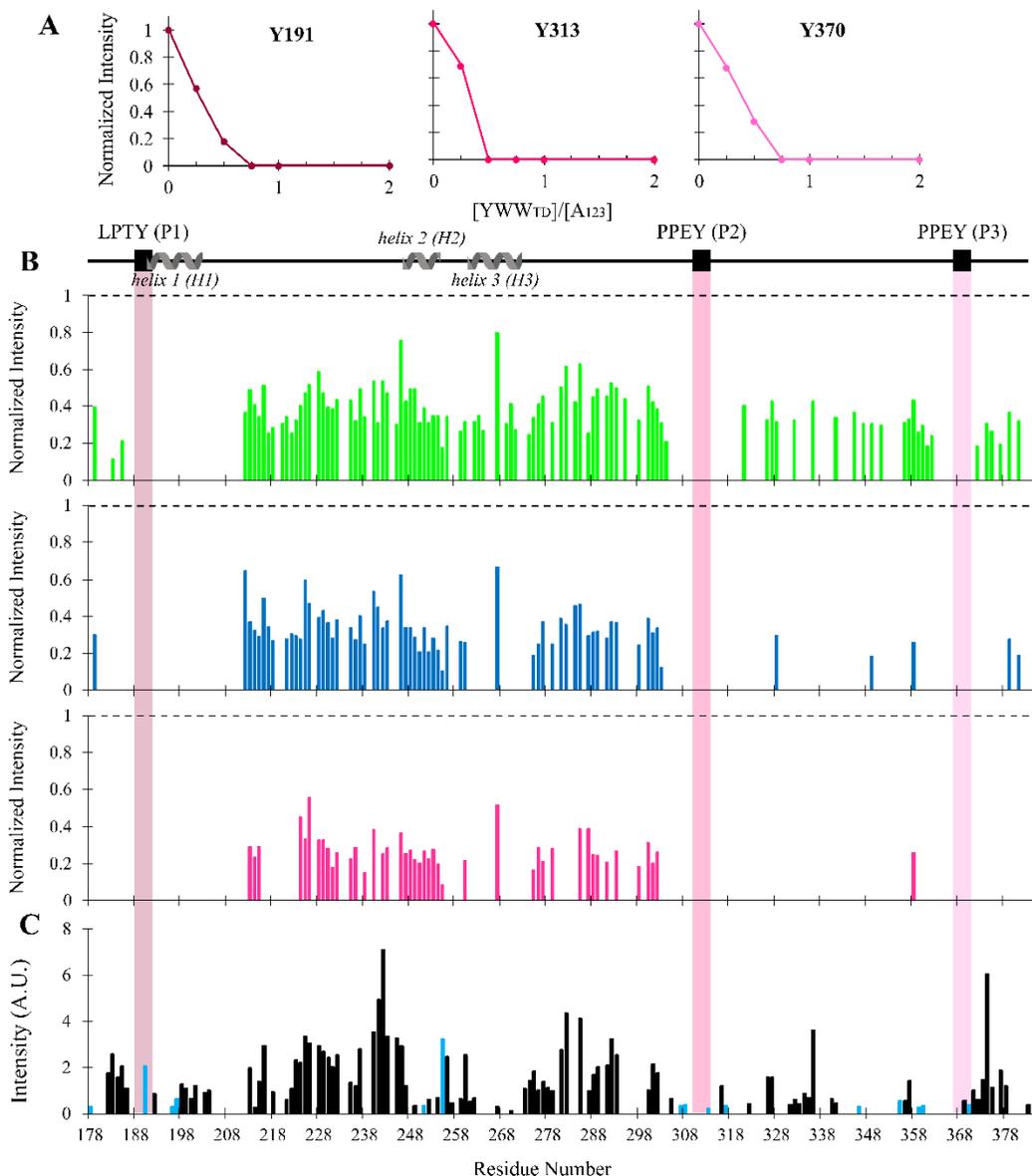


Figure 2.4: NMR mapping of the A₁₂₃-YWW_{TD} binding interface. (A) Changes in peak intensities of the three tyrosine residues, Y191 (P1), Y313 (P2), Y370 (P3), of each motif as a function of YWW_{TD}-A₁₂₃ molar ratio. (B) Normalized intensities of assigned A₁₂₃ residues plotted as a function of the residue number for apo A₁₂₃ (black dashed lines), and upon addition of 0.75 (top panel, green), 1 (middle panel, blue), and 2 (bottom panel, magenta) molar equivalents of YWW_{TD}. The predicted secondary structure is shown above the plots. At the final molar equivalents of 2:1 (bottom panel), only residues primarily located in the P1-P2 linker segment are observed in the spectrum. (C) A bar plot of the intensities of peaks in the CLEANEX-HSQC spectrum of unbound A₁₂₃. Peaks that disappear upon addition of 0.25 molar equivalents of YWW_{TD} are shown in cyan and include residues at the three PPEY sites (shaded in pink).

Complex formation also induces resonance intensity changes in other residues and leads to many missing peaks in spectra of the YAP-bound protein. These missing peaks imply that the residues are close to binding interfaces and/or involved in binding-induced allosteric changes[161]. At the 0.75:1 molar ratio (Fig. 2.4B, upper panel), missing crosspeaks map to residues in the sequence vicinity of P1 (186-204), H3 (261, 265, 268 and 273), P2 (305-317), P3 (365-371), and the linker segment between P2 and P3 (322, 335, 340, 342). Additional cross-peaks corresponding to residues 245, 262-264, 269-274 and C-terminal residues 302-384 (except 328, 349, 358) are missing in the spectrum of the 1:1 YWW_{TD} - A_{123} complex (Fig. 2.4B, middle panel). At the final molar ratio of 2:1 (YWW_{TD} : A_{123}), the only remaining spectral peaks correspond to residues primarily located in the P1-P2 linker segment (Fig. 2.4B, lower panel).

To identify protected segments in the partially bound- A_{123} polypeptide, CLEANEX-HSQC experiments were recorded at a YWW_{TD} : A_{123} molar ratio of 0.25:1. Peaks that disappear upon addition of 0.25 molar equivalents of YWW_{TD} are mapped onto the CLEANEX-HSQC data of unbound A_{123} and are shown as cyan bars in Figure 2.4C. The least solvent exposed amide protons (most protected residues) map to residues in the sequence vicinity of P1, P2, P3, and the predicted H2 helical segment.

To summarize the NMR titration and CLEANEX-HSQC data, the partially bound A_{123} polypeptide formed by addition of a sub-stoichiometric concentration of YWW_{TD} to A_{123} results in disappearance of peaks in the sequence vicinity of all three PPXY motifs, a clear indication that YWW_{TD} binds all three sites concurrently. Most of the missing peaks in the stoichiometric complex are in the sequence vicinity of the P1 site, a clear indication of more favorable interactions at the P1 site. Further, peaks corresponding to most residues in the predicted helix (H3), which is not a putative binding site, also disappear, possibly due to binding-induced conformational changes. Finally, spectral peaks remaining in the fully bound complex map to the linker segment between P1 and P2; this clearly shows that residues in this linker region of A_{123} remain largely disordered in the YAP-bound protein.

ITC experiments of WT and site-directed mutants. Reaction thermodynamics, stoichiometry, and effective binding affinity of the A_{123} - YWW_{TD} interaction were

determined by ITC. Binding of the two proteins is enthalpically driven, occurs with a binding stoichiometry (N) close to 1:1 and has an effective dissociation constant (K_d) of $0.26 \pm 0.01 \mu\text{M}$ (Figure 2.5A, Table 2,2). Because there are two putative binding sites on the YWW_{TD} polypeptide (2 WW domains) and three putative binding sites on

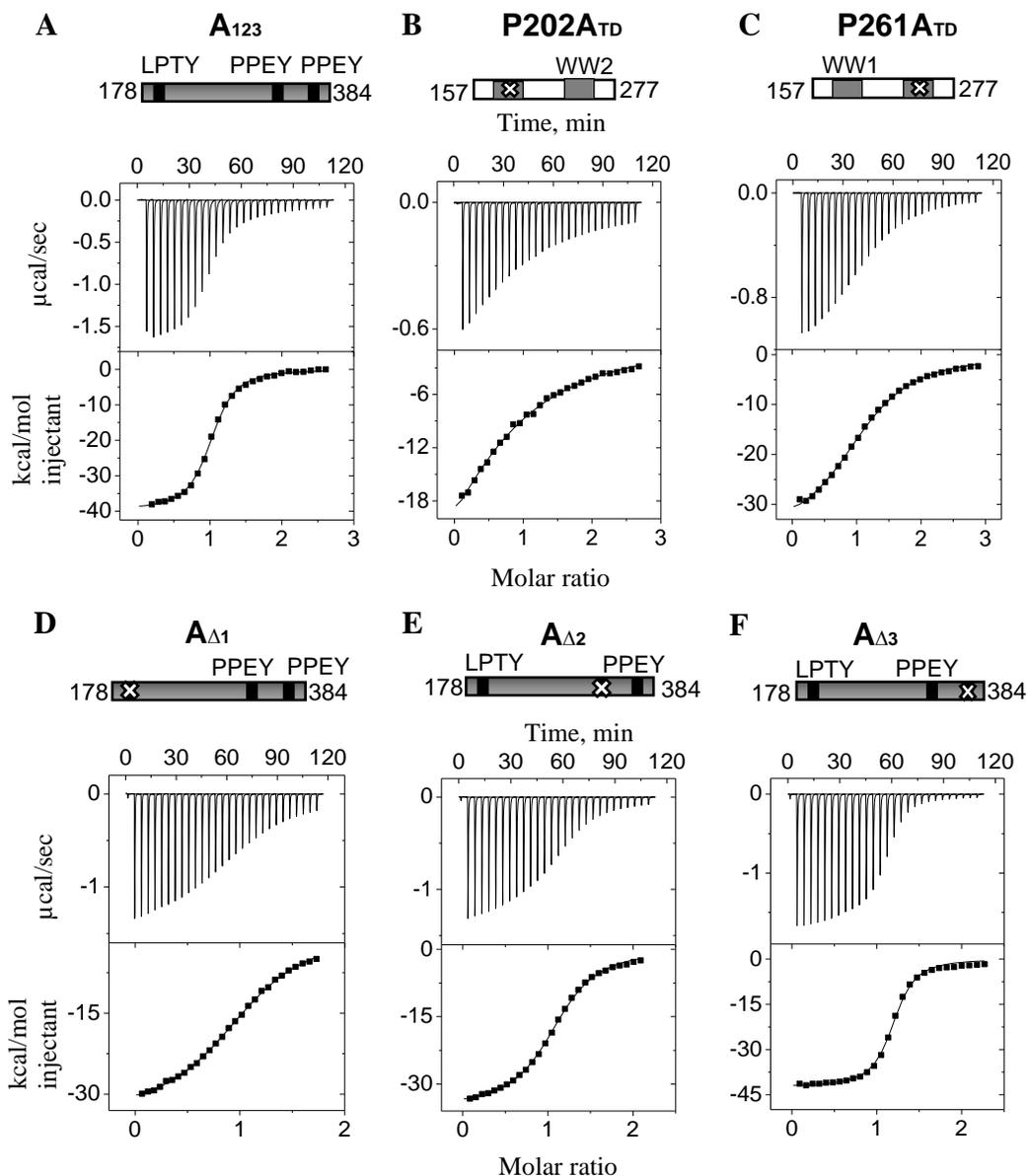


Figure 2.5: ITC binding isotherms of the A_{123} - YWW_{TD} interaction. Representative binding isotherms for interactions of A_{123} with (A) YWW_{TD} (B) P202A_{TD} and (C) P261A_{TD} or YWW_{TD} with A_{123} mutants (D) $\text{A}_{\Delta 1}$, (E) $\text{A}_{\Delta 2}$, and (F) $\text{A}_{\Delta 3}$ in which sites P1, P2 and P3, respectively are inactivated. Active domains or motifs are labeled in the schematic of the constructs above each binding isotherm. ITC data were collected at 25 °C in pH 7.5 buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 5 mM β -mercaptoethanol, and 0.5 mM NaN_3 .

the A₁₂₃ polypeptide (PPXY motifs), binding of a single WW-PPXY site or of both WW sites to two PPXY sites will result in a binding stoichiometry of 1:1.

Table 2.2: Thermodynamics parameters for the YAP-AMOTL1 interactions.
*Reported value is the average of duplicate experiments.

Titrant	Cell	N	K _d (μ M)	Δ H (kcal/mol)	T Δ S (kcal/mol)	Δ G (kcal/mol)
YWW _{TD}	A ₁₂₃	1.1	0.26 \pm 0.01	-39.6 \pm 0.9	-30.3 \pm 0.9	-8.98 \pm 0.03
	A Δ 1*	1.0	1.8	-33.8	-26.0	-7.8
	A Δ 2	1.1	0.6 \pm 0.1	-35.5 \pm 0.3	-27.1 \pm 0.3	-8.47 \pm 0.01
	A Δ 3*	1.2	0.11	-41.1	-31.8	-9.5
P202A _{TD}	A ₁₂₃	1.0	7.6 \pm 0.2	-37 \pm 4	-30 \pm 4	-6.98 \pm 0.01
P261A _{TD}	A ₁₂₃	1.2	0.70 \pm 0.02	-30 \pm 2	-22 \pm 2	-8.40 \pm 0.02

To determine whether a single YAP WW domain is sufficient for binding, we introduced mutations that inactivate either the first (P202A_{TD}) or second (P261A_{TD}) YAP WW domain. Binding of A₁₂₃ to the P202A_{TD} mutant occurs with a K_d of 7.6 μ M (Figure 2.5B), a value \sim 10-fold higher than 0.7 μ M computed for the A₁₂₃-P261A_{TD} interaction (Figure 2.5C). The binding stoichiometry of 1:1 or 1:1.2 implies that in each case the active WW domain binds at least one A₁₂₃ PPXY. However, the weaker but significantly different binding affinities of the mutants relative to the WT YWW_{TD} polypeptide suggest that both YAP WW domains contribute to the stability of the WT YWW_{TD}:A₁₂₃ complex.

To determine if inactivating specific PPXY sites destabilizes the A₁₂₃-YWW_{TD} complex, we used A₁₂₃ ‘knockout’ variants, A Δ 1, A Δ 2, and A Δ 3, each with one inactivated motif (indicated by subscript). Figures 2.5D-F show that the reactions are enthalpically driven and have a binding stoichiometry in the range of 1:1 to 1:1.2 for each ‘knockout’ site (Table 2.2). The effective K_d varies from 1.8 μ M in A Δ 1 (Figure 2.5D) to 0.6 μ M in A Δ 2 (Figure 2.5E) and 0.1 μ M in A Δ 3 (Figure 2.5F, Table 2.2). To summarize the ITC data, binding of YWW_{TD} to A₁₂₃ is more stable (lower K_d) when both the YAP WW1 domain and AMOTL1 P1 sites are active. Inactivating the P2 and P3 sites have modest effects on the stability of the complex. Finally, the non-integer binding stoichiometries are indicative of inhomogeneity in the complexes formed, possibly due to variations in the number of occupied binding sites.

Discussion

Direct binding of the AMOTL1 scaffold protein to YAP regulates cell proliferation. While it is known that assembly of the complex occurs via multivalent WW-PPXY interactions, and studies with shorter polypeptides containing one or two PPXY sites have provided some insight of the binding interactions, it remains unknown how polypeptides containing the full binding domain interact, specifically which pairs of the three PPXY sites bind the two YAP WW domains. We have clarified the uncertainty as to which AMOTL1 PPXY motifs are required for assembly of its complex with YAP. Our first advancement was to produce a construct of AMOTL1 (A₁₂₃), which contains the entire multivalent PPXY segment rather than only one or two PPXY sites, and to use site-directed mutagenesis of A₁₂₃ to inactivate specific PPXY sites. Then we characterized interactions of the tandem WW domains of YAP (YWW_{TD}) with A₁₂₃ and its PPXY variants. Our conclusions are drawn from collective analyses of data measured by an extensive array of methodologies applied both to apo proteins and to the complex formed by A₁₂₃ and YWW_{TD}.

We conclude that the apo AMOTL1 PPXY domain, A₁₂₃, has limited folded structure, localized in two short helical segments, residues 193-197 and 245-257 as shown by CD and NMR. This first experimental demonstration that the full AMOTL1 PPXY segment is partially disordered is consistent with the structures of other multivalent PPXY proteins [84, 90, 162]. Second, the AMOTL1-YAP PPXY-WW complex is formed by one molecule of AMOTL1 bound to one molecule of YAP, as shown by the hydrodynamic method of SEC-MALS and ITC. Third, binding of the YAP tandem WW domains to A₁₂₃ perturbs residues at all three PPXY sites as shown by solution NMR spectroscopy. More residues are perturbed in the sequence vicinity of the P1 site, a clear indication of more intermolecular interactions at P1 relative to P2 or P3, and evidence that P1 is critical for assembly of the AMOTL1-YAP complex. The P1 site in the AMOTL1 paralog, AMOT, is also indispensable for binding specific WW domain targets[149] and plays a critical role in the function of the protein[163]. Likewise, the P1 site in AMOTL1 may provide some functional advantages to the protein. Fourth, inactivating the P3 site results in a ~2-fold binding enhancement which implies that P3 contributes negative entropy to the overall interaction. We speculate

that the negative entropy may be eliminated when the P3 site is occupied by another WW domain protein such as the Kidney and Brain expressed protein, KIBRA which is reported to preferentially bind the P3 site in AMOT[123]. Inactivating the YAP WW1 domain also de-stabilized the complex to a much greater degree than the WW2 domain, findings in agreement with the different binding preference of the YAP WW domains [78].

Considered together, the novel information from these complementary methodologies indicate that the AMOTL1-YAP PPXY-WW complex is structurally dynamic, and the YAP WW tandems interact with all three PPXY sites in AMOTL1 as illustrated by the model in Figure 2.6. The dynamic structure is composed of different complexes of varying stabilities as previously noted for other WW-PPXY complexes [90, 162]. Complexes in which both the YAP WW1 domain, and the N-terminal AMOTL1 LPTY (P1) sites are bound are the most stable and likely initiate the binding process. The binding process is completed by a relatively weak or transient binding of the second YAP WW domain (WW2). This step leads to simultaneous binding of both YAP WW domain sites to varying pairs of AMOTL1 PPXY sites and to formation of an equilibrating mixture of interconverting species transiently formed by two YAP sites bound to two conjugate AMOTL1 sites.

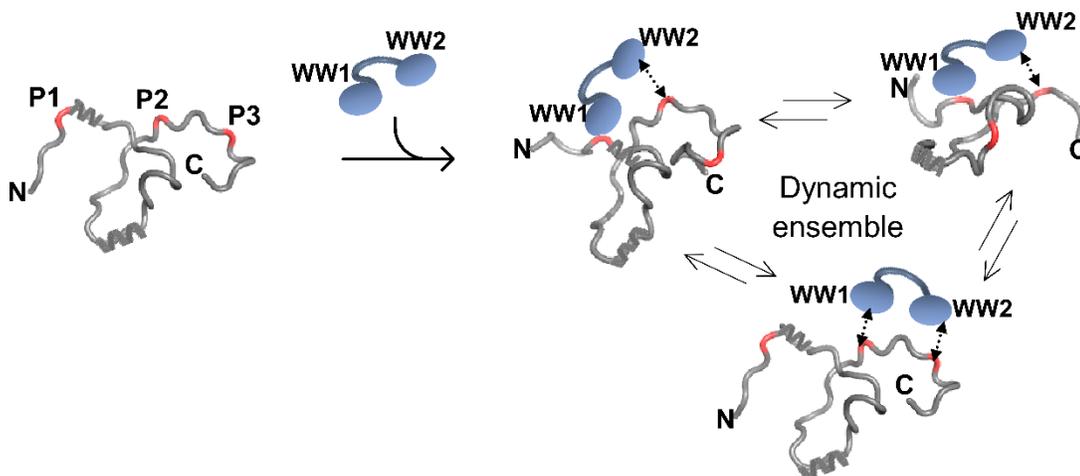


Figure 2.6: Model of the species formed in the AMOTL1-YAP complex. The YAP tandem WW domains and the predominantly unstructured AMOTL1 PPXY domain-containing three PPXY sites, designated P1, P2, and P3, form a dynamic ensemble of interconverting complexes. Complexes in which AMOTL1 sites P1-P2 or P1-P3 are bound are relatively more stable than the complex in which sites P2-P3 are bound.

Functional implications of a dynamic YAP-AMOTL1 complex. A hallmark of regulatory complexes in signaling pathways is their rapid and spontaneous response to cellular signals. The modest micromolar binding affinity of the dynamic complex can facilitate rapid assembly and disassembly of the complex in response to changes in the local environment. Further, recent findings that link YAP [164], multivalency, intrinsic disorder and low complexity regions, to phase separation and the formation of membraneless organelles in human cells[165-168] could organize the ensemble in space. Phase separation may be triggered by changes in the local environment such as a change in the local concentration of YAP. The presence of another WW domain-containing protein could also shift the dynamic equilibria in favor of specific complexes. For instance, AMOTL1 forms a ternary via WW-PPXY interactions with both YAP and the Kidney and Brain expressed protein KIBRA[2]; and simultaneous binding to both WW domain partners may shift the dynamic equilibria in favor of specific complexes in the ensemble. We propose that forming dynamic complexes underlie AMOTL1 function.

Materials and Methods

Cloning of constructs and recombinant protein production are described in Appendix 1.

Circular dichroism. Far UV circular dichroism (CD) measurements were recorded at 25 °C on a JASCO 720 spectrophotometer using a pathlength of 1 mm, and a bandwidth of 1.0 nm. Prior to data collection, the protein was dialyzed against 10 mM sodium phosphate, pH 7.5 with or without 100 mM NaF. The final concentration of the protein for CD analysis was 3 or 5.4 μ M and reported CD data are the average of experimental repeats.

Size exclusion chromatography-multi-angle light scattering (SEC-MALS). The average molar masses and association states of the proteins were determined from size exclusion chromatography (SEC) (AKTA FPLC, GE Healthcare) connected to multi-angle light scattering (DAWN, Wyatt Technology) and refractive index (Optilab, Wyatt Technology) detectors. 100 μ L of proteins in the concentration range of 25-150 μ M were injected at a flow rate of 0.7 or 1.0 mL/min onto a Superdex²⁰⁰ 10/300 (Cytiva

life sciences) chromatographic column pre-equilibrated with a pH 7.5 buffer composed of 50 mM sodium phosphate, 0.4 M NaCl, 1 mM NaN₃, 5 mM β-mercaptoethanol. Average molar masses were computed with the ASTRA software package, version 8 (Wyatt Technologies).

NMR data collection and analysis. NMR experiments were performed at 10 °C on a Bruker Avance III, 800 MHz spectrometer (Bruker BioSpin) equipped with a triple resonance cryogenic probe. Data were collected on isotopically labeled A₁₂₃ at concentrations of 75 or 200 μM and in a pH 6.8 buffer composed of 50 mM sodium phosphate, 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM NaN₃, and 2 mM tris(2-carboxyethyl) phosphine. All samples contained 10% D₂O as the lock nucleus and 0.5% DSS for internal referencing.

Backbone resonances were assigned from BEST-TROSY ¹H-¹⁵N HSQC experiments [169] and triple resonance (3D) experiments HNCACB, HN(CO)CACB, HNCO, HN(CA)CO, and ¹H-¹⁵N HSQC-TOCSY. The TROSY-based pulse sequences were used to reduce line broadening due to rapid signal decay. All 3D experiments used non-uniform sampling (NUS) to reduce data collection times. The NUS data was reconstructed using the iterative shrinkage thresholding approach in NMRPipe [170]. NMR spectra were processed in NMRPipe [170] and visualized with Sparky [171] or NMRView [172].

Per-residue secondary structure propensities were calculated from the deviations of experimental CA and CB chemical shifts from the random coil values of Poulsen et al (https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/) [173-175]. Substantial positive deviations greater than a pre-determined mean value, for four or more sequential residues indicate helical propensity while negative deviations for three or more sequential residues suggest extended structure.

Heteronuclear-NOE (HetNOE), longitudinal (R₁) and transverse (R₂) spin relaxation rates were measured using TROSY-based interleaved pulse sequences [176]. HetNOE experiments were collected with and without proton saturation using a relaxation delay of 8 s. NOE errors were calculated as previously reported [84]. R₁ spin relaxation rates were measured with relaxation delays of 0.02, 0.06 (x3), 0.1, 0.2, 0.4, 0.6, 0.8, and 1.2 s. R₂ spin relaxation rates were obtained with delays of 17, 34 (x3),

51, 68, 85, 140, 170, and 240 ms with a recycle delay of 1.5 s. R_1 and R_2 data were fit to a single exponential decay function, $I(t) = I_0e^{-Rt}$, where t is the variable relaxation delay and R is the relaxation rate. Rates were computed using the rate analysis tool in NMRView [172]. Experimental errors were estimated from standard deviations of triplicate (x3) experiments. Results are reported only for unambiguously assigned resonances with reliably quantified peak intensities.

CLEANEX NMR experiments were collected with a mixing time of 100 ms using a recycle delay of 1.5 s. The A_{123} - YWW_{TD} complex was formed by adding 0.25 molar equivalents of YWW_{TD} to A_{123} .

For NMR titration experiments, unlabeled YWW_{TD} was added to ^{15}N -labeled A_{123} to final molar ratios (unlabeled: ^{15}N -labeled:) of 0.25:1, 0.5:1, 0.75:1, 1:1, and 2:1. To correct for minor variations in sample concentrations at each molar ratio, peak intensities (measured as peak height) were normalized as the ratio of the intensity of the peak in the bound spectrum to the intensity of the peak in the unbound spectrum.

Isothermal titration calorimetry (ITC). A VP-ITC instrument (Malvern instruments Inc, MA) set to 25 °C, was used to record ITC data. Prior to the titrations all samples were extensively dialyzed against a pH 7.5 buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol. 27 or 28 injections of 96-148 μM WW proteins were titrated into 6-16 μM A_{123} or its variants. Data were collected in duplicates, triplicates or quadruplicates, using proteins from two different preparations. Reported data are the average of experimental repeats. Isotherms were analyzed by single-site fits of the thermograms using the *Origin 7.1* software. The free energy of binding (ΔG) was calculated from the equation $\Delta G = -RT\ln(K_a)$, where R is the universal gas constant, T is temperature in Kelvin, and K_a is the association constant.

delta2D Method. The residue-level disorder probability of A_{123} was calculated with CA, CB, CO, N, and HN chemical shifts using the delta2D web server[158].

DichroWeb Method. The percent helicity of A_{123} was calculated using DichroWeb with data collected at 25 °C on a 5.4 μM sample in buffer composed of 10 mM sodium phosphate and 100 mM NaF at pH 7.5[177]. Reported helicity was obtained the Contin-LL method and the reference set SP175[178].

Acknowledgements

The authors wish to thank Profs. Clare Woodward and Elisar Barbar for valuable discussions, and Prof. Joachim Kremerskothen (University Hospital, Münster, Germany) for the AMOTL1 cDNA. This work is supported in part by the National Science Foundation (MCB-2114544 to Afua Nyarko) and the Christopher and Catherine Matthews Graduate Fellowship (to Amber Vogel). NMR experiments were collected at the Oregon State University NMR Facility funded in part by the National Institutes of Health, HEI Grant 1S10OD018518, and by the M.J. Murdock Charitable Trust grant # 2014162.

Chapter 3

Domain and motif-specific roles in assembly of the multivalent WW-PPXY complex of Angiotensin-Like 1, Yes-associated protein and Kidney and Brain-expressed protein

Amber Vogel, Matthew McWhorter, and Afua Nyarko

In preparation for submission

Abstract

Multivalent WW domain-PPXY motif interactions are abundant in biological processes and implicated in numerous disease states, making them desirable therapeutic targets. While the functional importance of multiple multivalent WW-PPXY complexes is well-documented, mechanistic insight into the assembly process at the residue level is largely undetermined. One example is the interaction between the multivalent scaffolding cell-growth regulator proteins, Kidney and Brain-expressed protein (KIBRA) and Angiotensin-Like 1 (AMOTL1), which contain two WW domains and three PPXY motifs, respectively. The assembled complex protects AMOTL1 from proteasomal degradation, resulting in decreased cell proliferation. Here, we utilized constructs of KIBRA and AMOTL1 containing all binding-competent sequences to unravel the binding mechanism in context of multivalency and in the presence of an alternate WW domain partner. Using isothermal titration calorimetry, multi-angle light scattering, and nuclear magnetic resonance, we reveal specific roles for the multiple domains/motifs in the assembly process. For the KIBRA WW domains, we show that both domains are equally important for the stability of the complex. For the AMOTL1 PPXY motifs, we demonstrate that the third PPXY site initiates binding, and the first/second PPXY sites modulate the stability of the complex. In the presence of a second WW domain partner, the transcription regulator Yes-associated protein (YAP), the role of the first/second PPXY sites become less pronounced since the YAP WW domains bind these sites. Domain/motif-specific roles likely fine-tune the assembly process by facilitating rapid assembly/disassembly of multivalent WW-PPXY complexes in response to changes in the local cellular environment.

Introduction

Multivalent protein interactions between intrinsically disordered proteins (IDPs) and regions (IDRs) and their protein binding partners play essential roles in signaling networks and multiple disease conditions, making them attractive molecular targets for therapeutic intervention[66, 138, 179]. A prominent pathway regulated by such multivalent interactions is the Hippo signaling pathway, a phosphorylation cascade which targets the transcriptional coactivator and cell growth regulator, Yes-associated protein (YAP)[111]. Subcellular localization of YAP is a determinant of cell growth and death, whereby nuclear and cytosolic localization result in increased and decreased cell growth, respectively. Critical for the regulation of YAP are two tandem WW domains, a 35-40 residue triple antiparallel β -sheet structure characterized by the presence of two conserved tryptophan residues which bind the PPXY motif (where X is any amino acid)[60, 75, 180]. WW domain-PPXY motif interactions are pervasive throughout the Hippo signaling pathway and understanding how these complexes assemble is critical to the development of cancer therapeutics.

Angiomotin-Like 1 (AMOTL1) is a 107 kDa multivalent scaffolding protein from the Motin family implicated in numerous cell functions including regulation of YAP subcellular localization[33, 117, 146, 147]. The Motin family contains three members: Angiomotin (AMOT), Angiomotin-Like 1 (AMOTL1), and Angiomotin-Like 2 (AMOTL2), which contain conserved modular domain architecture including 2-3 PPXY motifs (three in AMOTL1: P1, P2, and P3) which bind numerous WW domain proteins. Two of the WW domain-containing binding partners of AMOT(L1) are YAP and Kidney and Brain-expressed protein (KIBRA, also known as WWC1)[2, 33, 117, 147, 181, 182]. KIBRA is a multidomain scaffolding protein containing two WW domains with essential roles in directional migration of podocytes, memory and cognition, and upstream regulation of the Hippo signaling pathway[112, 183-185]. Both the AMOTL1-YAP and AMOTL1-KIBRA complexes increase the stability of AMOTL1 and downregulate cell growth and cancer development[2, 126, 146]. Furthermore, work by Kremerskothen et al. showed that AMOTs can bind YAP and KIBRA simultaneously[2], which suggests that multivalent PPXY motif interactions allow AMOTL1 to act as a molecular scaffold in cell signaling. Uncovering the

mechanisms by which AMOTL1 discriminates between its partners on the molecular level could inform the development of cancer drugs.

Limited structural information based on polypeptides containing one or two AMOT(L1) PPXY sites implicate the N- and C-terminal sites in complex assembly with YAP and KIBRA, respectively[3, 123]. However, how these complexes form in context of the complete set of binding-competent motifs has previously been unclear, due to the difficulties in reconstructing and characterizing large, disordered protein complexes. Recent work by our group successfully characterized the full PPXY motif segment of AMOTL1 and its multivalent interaction with YAP[4]. In this study, we showed that YAP forms the strongest interaction with P1, but transient interactions with P2 and P3 result in a dynamic complex with reduced overall affinity. However, a comparable study of complex assembly between AMOTL1 and KIBRA with all binding sites available does not exist. The most complete study to date solved a structure of the tandem WW domains of KIBRA in complex with the C-terminal PPXY motif of AMOT, which showed binding of both domains to this motif[3]. How the complex assembles in context of all available binding sites remains unclear.

In this work, we use a multitude of complimentary biophysical techniques and recombinantly expressed AMOTL1, KIBRA, and YAP polypeptides containing the full complement of binding sites to understand how the binary AMOTL1-KIBRA and ternary YAP-AMOTL1-KIBRA complexes assemble in context of multivalency. We show that AMOTL1-KIBRA complex assembly depends on WW domain cooperativity and is initiated by relatively stable interactions at the third PPXY motif. These initial contacts are followed by transient interactions at the first and second PPXY sites which destabilize overall complex stability. Furthermore, we successfully reconstitute the ternary complex and show how alternate motif preference results in a stable scaffold anchored by AMOTL1. We propose that forming dynamic complexes and scaffolding underlie the functions of AMOTL1 and its paralogs.

Results

Rationale for construct design. The tandem WW domains of KIBRA correspond to residues 6-39 (WW1) and 53-86 (WW2). The KIBRA KWW_{TD} polypeptide used in this study contains both WW domains and spans residues 1-91[84]. Constructs P37A_{TD} and P84A_{TD} contain proline to alanine substitutions which inactivate the WW1 or WW2 domain, respectively[84]. The AMOTL1 A₁₂₃ polypeptide (residues 178-384) contains all three PPXY motifs (P1, P2, and P3); constructs A_{Δ1}, A_{Δ2}, A_{Δ3}, and A₃ have sites P1, P2, P3, and both P1 and P2, respectively inactivated and were designed to test the importance of each PPXY site to the stability of assembled complexes. Schematics of constructs are shown in Figure 3.1A-B.

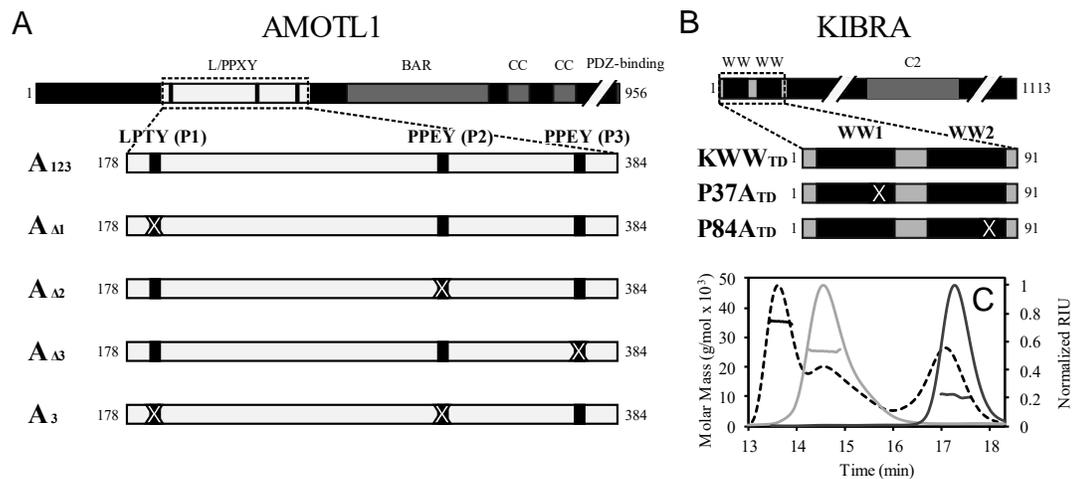


Figure 3.1: KWW_{TD} and A₁₂₃ construct design and complex assembly. (A) Domain architecture of full-length (FL) AMOTL1 (top) containing three N-terminal PPXY motifs, a bin/amphiphysin/rvs (BAR) domain, two coiled-coil (CC) domains, and a C-terminal PDZ-binding domain. Protein constructs used in this study include residues 178-384 of AMOTL1 with all PPXY motifs active (A₁₂₃). Additionally, A₁₂₃ mutant constructs were generated in which the PPXY motif tyrosine was mutated to alanine (denoted by a white 'X') to abrogate binding in P1 (A_{Δ1}), P2 (A_{Δ2}), P3 (A_{Δ3}), and P1 and P2 (A₃). (B) Domain architecture of FL-KIBRA containing two N-terminal WW domains and a C2 domain. Protein constructs used in this study include residues 1-91 of KIBRA with both WW domains active (KWW_{TD}) and mutant constructs which abrogate binding to the first (P37A_{TD}) or second (P84A_{TD}) WW domain through mutation of a highly conserved proline to alanine, represented by a white 'X'. (C) SEC-MALS of KWW_{TD} (solid dark grey line), A₁₂₃ (solid light grey line), and a 1:2 mixture of A₁₂₃:KWW_{TD} (dotted black line).

SEC-MALS experiments to determine the stoichiometry of the AMOTLI-KIBRA complex. The SEC-MALS elution profiles and molecular masses of A₁₂₃, KWW_{TD} and a 1:2 mixture of A₁₂₃:KWW_{TD} are shown in Figure 3.1C. The absolute molar mass of KWW_{TD} was 10.7 ± 0.6 kDa, which agrees with the theoretical monomeric mass of 11.3 kDa (Table 3.1). Complexes constructed by mixing A₁₂₃ and KWW_{TD} at molar ratios of 1:1 and 1:2 gave absolute molar masses of 35.2 and 34.6 ± 0.9 kDa, respectively, slightly lower than the theoretical molar mass of 37.2 kDa for a 1:1 A₁₂₃:KWW_{TD} complex. From these results, we conclude that the predominate A₁₂₃:KWW_{TD} complex is formed between one polypeptide of A₁₂₃ and one polypeptide of KWW_{TD}.

Table 3.1: Average molar masses of apo and bound KWW_{TD}. *Reported values are the average of duplicate experiments. ^a Values from reference [4].

Sample	Observed MW (kDa)	Theoretical MW (kDa)
KWW _{TD}	10.7 ± 0.6	11.3
A ₁₂₃ ^a	26.1 ± 1.2	25.9
1:1 A ₁₂₃ :KWW _{TD}	35.2*	37.2
1:2 A ₁₂₃ :KWW _{TD}	34.6 ± 0.9	48.5

ITC experiments to determine the relative stabilities of WT and mutant AMOTLI-KIBRA complexes. ITC-measured binding thermodynamic parameters were determined for the interaction between KWW_{TD} and WT or mutant A₁₂₃ polypeptides. (Table 3.2). The experimentally measured binding stoichiometries were close to 1 which is interpreted as one KWW_{TD} polypeptide binds one WT or mutant A₁₂₃ polypeptide. The K_d for the WT A₁₂₃-KWW_{TD} interaction is 0.7 ± 0.2 μ M (Fig. 3.2A).

Table 3.2: ITC of KWW_{TD} titrated into A₁₂₃/mutants collected at 25 °C. *Average calculated from duplicate runs.

Syringe	Cell	N	K _d (μ M)	Δ H (kcal/mol)	T Δ S (kcal/mol)	Δ G (kcal/mol)
KWW _{TD}	A ₁₂₃	0.9 ± 0.1	0.7 ± 0.2	-37 ± 6	-29 ± 6	-8.4 ± 0.1
KWW _{TD}	A _{Δ1}	1.0	0.4*	-38	-30	-8.7
KWW _{TD}	A _{Δ2}	0.92 ± 0.02	0.25 ± 0.04	-40 ± 1	-31 ± 2	-9.07 ± 0.01
KWW _{TD}	A _{Δ3}	1.11 ± 0.04	22 ± 1	-49.3 ± 0.3	-42.9 ± 0.2	-6.37 ± 0.03
KWW _{TD}	A ₃	0.9 ± 0.1	0.27 ± 0.02	-40.0 ± 0.7	-31 ± 0.6	-9.0 ± 0.1

The interactions between KWW_{TD} and A_{123} mutant polypeptides $A_{\Delta 1}$ (P1 inactivated, Fig. 3.2B), $A_{\Delta 2}$ (P2 inactivated, Fig. 3.2C), and A_3 (both P1 and P2 inactivated, Fig. 3.2E) resulted in slightly higher affinities than WT, with K_d values of 0.4, 0.25, and 0.27 μM , respectively. In contrast, binding was relatively weak (high K_d) for A_{123} mutant polypeptide $A_{\Delta 3}$ (P3 inactivated, Fig. 3.2D), with a K_d value of 22 μM . Taken together, the ITC data show that the KWW_{TD} - A_{123} complex is relatively more stable when the P3 site alone or in combination with P1 or P2 is available for binding to KWW_{TD} .

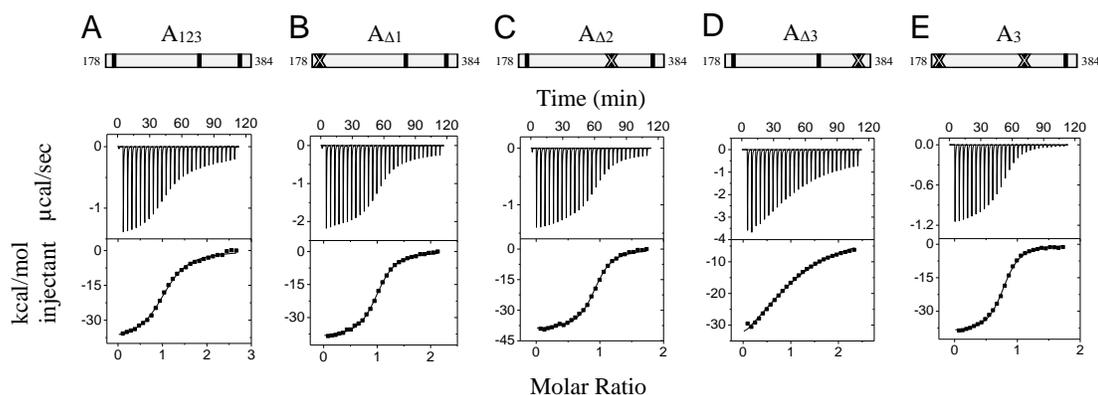


Figure 3.2: Binding of KWW_{TD} to A_{123} /mutants monitored by ITC. Representative isotherms of KWW_{TD} titrated into AMOTL1 polypeptides (A) A_{123} , (B) $A_{\Delta 1}$, (C) $A_{\Delta 2}$, (D) $A_{\Delta 3}$, (E) A_3 collected at 25 $^{\circ}\text{C}$.

To determine the contribution of each WW domain to the stability of the A_{123} - KWW_{TD} complex, ITC experiments were performed with the KWW_{TD} mutants, $P37A_{TD}$ (inactive WW1 domain) and $P84A_{TD}$ (inactive WW2 domain). The binding thermograms did not reach sufficient saturation to compute reliable thermodynamic parameters. However, by fixing the stoichiometry (N) to 1 in the data analysis process, rather than allowing it to be optimized as an additional fitting parameter, we obtained K_d values ≥ 40 μM (Fig. 3.3, Table 3.3). Further, when either $P37A_{TD}$ or $P84A_{TD}$ were

Table 3.3: ITC of KWW_{TD} mutants titrated into A_{123} or A_3 collected at 25 $^{\circ}\text{C}$. N set to 1 in the data analysis process.

Syringe	Cell	N	K_d (μM)
$P37A_{TD}$	A_{123}	1	$> 63 \pm 3$ μM
$P84A_{TD}$	A_{123}	1	$> 44 \pm 5$ μM
$P37A_{TD}$	A_3	1	> 63 μM
$P84A_{TD}$	A_3	1	> 40 μM

titrated into A_3 , binding was too weak to accurately compute K_d values. To summarize, inactivating either WW domain weakens binding, a clear indication that both WW domains must be binding competent for the formation of a stable A_{123} -KWW_{TD} complex.

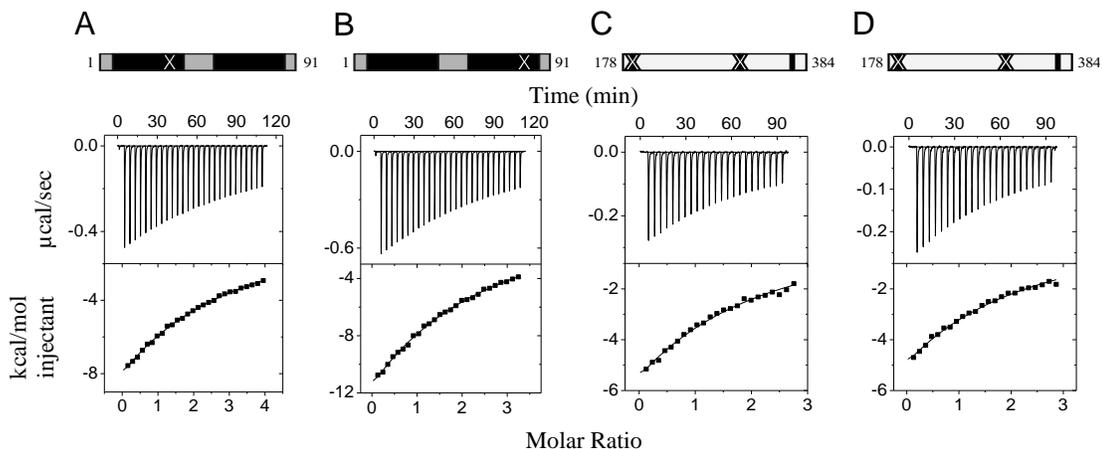


Figure 3.3: Binding of KWW_{TD} WW domain mutants to A_{123} and A_3 monitored by ITC. Representative isotherms of (A) P37A_{TD} titrated into A_{123} , (B) P84A_{TD} titrated into A_{123} , (C) P37A_{TD} titrated into A_3 , and (D) P84A_{TD} titrated into A_3 collected at 25 °C.

NMR studies to map binding interface residues. Chemical shift changes in a ^1H - ^{15}N HSQC spectrum, and/or broadening of peaks are sensitive indicators of residues directly involved in binding and/or residues which undergo binding-induced structural changes. Early binding events such as binding initiation sites are best observed at sub-stoichiometric ratios of ^{15}N -labeled:unlabeled proteins, while saturation concentrations of the unlabeled partner map all residues of the entire binding interface and/or residues impacted by binding-induced conformational changes.

NMR titration data of ^{15}N -labeled KWW_{TD} with unlabeled A_{123} is shown in Figure 3.4. Upon addition of 0.4 equivalents A_{123} , several peaks disappear and are labeled in Figure 3.4A. These missing peaks map to residues in both the WW1 and WW2 domains of the crystallographic structure of AMOT-bound mouse KIBRA WW domains[3], which share 98.9% sequence identity with human KIBRA WW domains (Fig. 3.4B). This data indicates that both WW domains bind AMOTL1 simultaneously.

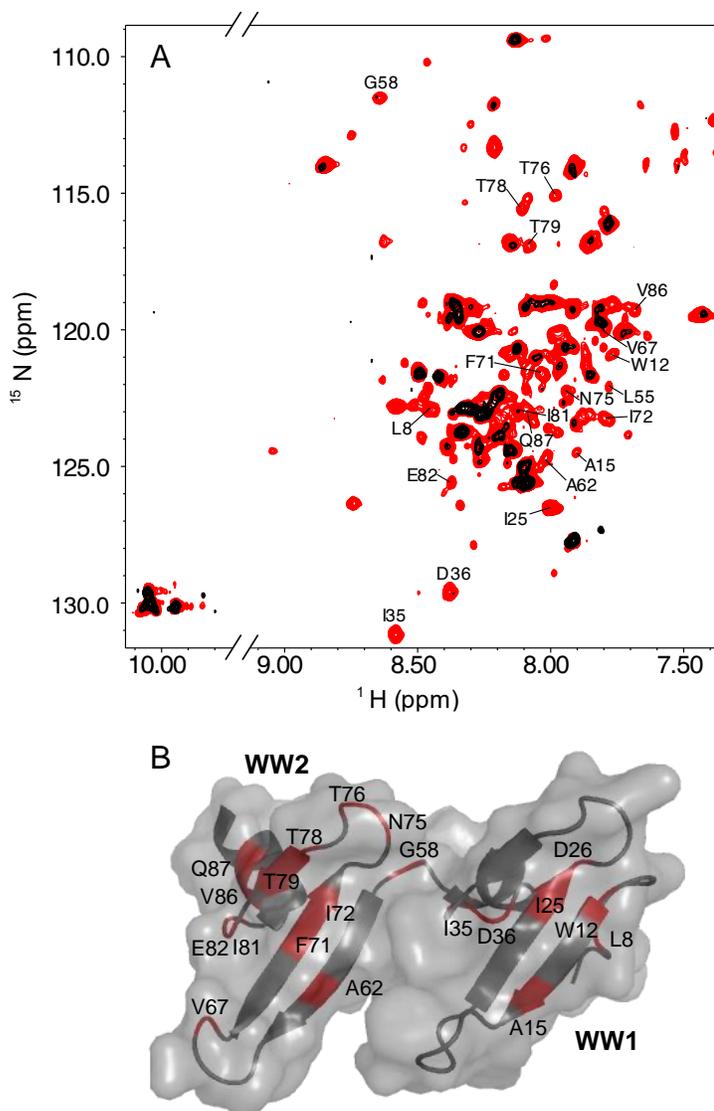


Figure 3.4: Binding-induced chemical shift changes of the KWW_{TD} WW domains. (A) ^1H - ^{15}N -HSQC spectra of 200 μM apo KWW_{TD} (red) overlaid with KWW_{TD} bound to 0.4 equivalents unlabeled A₁₂₃ (black). Assigned peaks which disappear at each titration point are labeled. (B) Structure of the tandem WW domains of KIBRA (PDB: 6JJX) with residues that disappear upon addition of 0.4 equivalents A₁₂₃ colored in red and labeled. All experiments were collected at 25 °C in buffer composed of 10 mM sodium phosphate (pH 6.8), 10 mM NaCl, 1 mM NaN₃, and 5 mM TCEP.

NMR titration data of ^{15}N -labeled A₁₂₃ with unlabeled KWW_{TD} is shown in Figure 3.5. Upon addition of sub-stoichiometric ratios of KWW_{TD}, 12 peaks disappeared, 9 of which are in the vicinity of P3. The localization these peak perturbations strongly imply that binding is initiated at the P3 site. Upon addition of 2-fold molar excess KWW_{TD}, a total of 57 peaks disappeared. These residues were primarily localized to P1, P2, and the P1-P2 linker. From these results, we conclude that all three PPXY motifs are involved in A₁₂₃-KWW_{TD} complex assembly, either by direct interactions or binding-induced conformational changes.

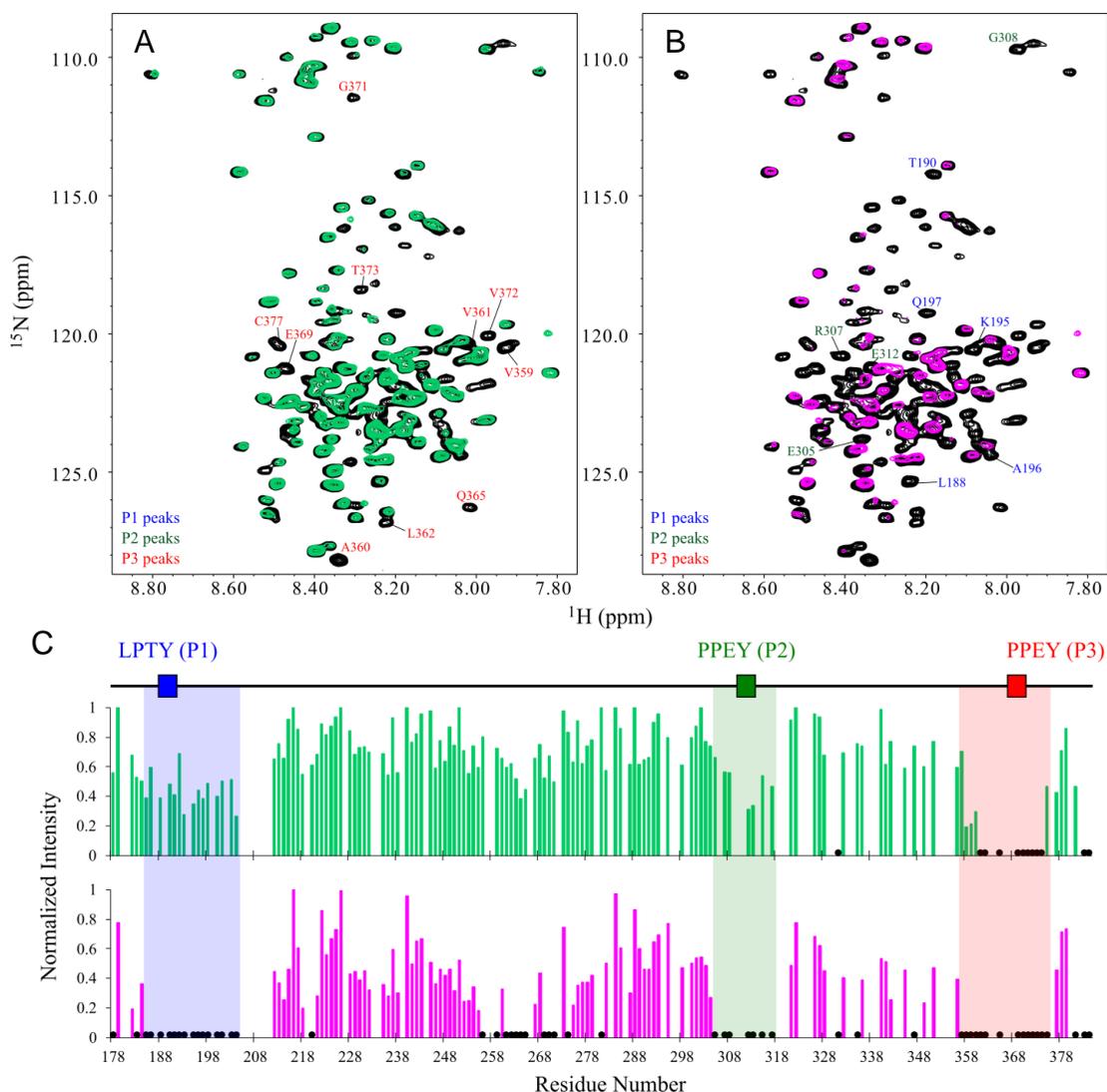


Figure 3.5: The binding interface of A₁₂₃ mapped by NMR shows binding of all three PPXY motifs initiated by P3. ^1H - ^{15}N -HSQC spectra of apo A₁₂₃ (black) overlaid with spectra of A₁₂₃ bound to (A) 0.75 (green) and (B) 2 (magenta) equivalents KWW_{TD}. Representative peaks which disappear in the vicinity of P1, P2, and P3 are labeled in blue, green, and red, respectively. (C) ^1H - ^{15}N -A₁₂₃ peak intensity normalized to initial peak height plotted versus residue number bound to (top) 0.75, or (bottom) 2 equivalents KWW_{TD}. Peaks which disappear are indicated by black circles. Stretches of residues perturbed upon binding in the vicinity of P1, P2, and P3 are shaded in blue, green, and red, respectively. All NMR data were collected at 10 °C.

Taken together, the NMR data indicate that both WW domains of KWW_{TD} bind A₁₂₃. KWW_{TD} binding to A₁₂₃ occurs first through extensive interactions with residues in the sequence vicinity of P3, followed by interactions with P1 and P2.

Solution studies of the YAP-AMOTL1-KIBRA ternary complex. To determine if KIBRA, AMOTL1, and YAP form a ternary complex in vitro, we mixed equimolar concentrations of A₁₂₃, KWW_{TD}, and a YAP polypeptide containing both WW domains (residues 157-277, YWW_{TD}). Figure 3.6A shows the analytical SEC elution profiles of the individual proteins, A₁₂₃, KWW_{TD}, YWW_{TD}, and an equimolar mixture of the three proteins. Formation of a ternary complex is evidenced by the presence of a single peak which elutes earlier than peaks corresponding to the binary complexes and resolves as three bands on SDS-PAGE (insert). The average molar mass of the ternary complex measured by SEC-MALS was 45.5 ± 0.5 kDa (Fig. 3.6B), slightly less than the theoretical molar mass of 51.4 kDa for a 1:1:1 ternary complex but higher than the theoretical and experimentally determined molar masses for each binary complex (Table 3.1).

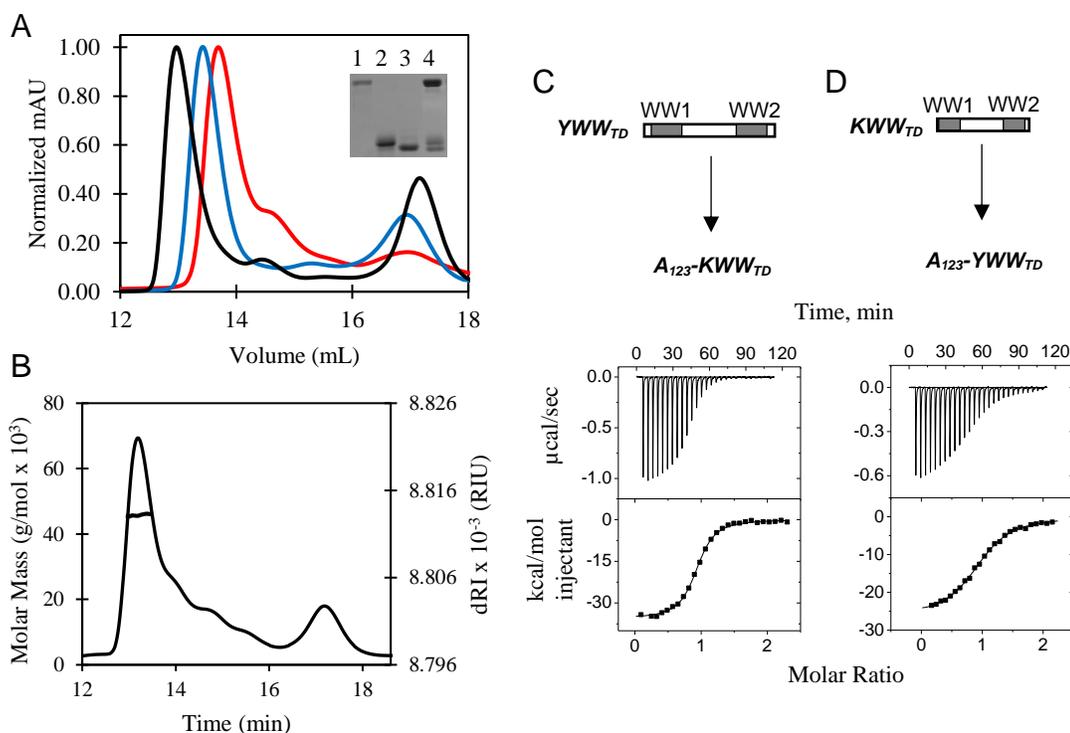


Figure 3.6: Assembly of the YWW_{TD}-A₁₂₃-KWW_{TD} ternary complex monitored by analytical SEC, SEC-MALS, and ITC. (A) Analytical SEC elution profiles of A₁₂₃ in complex with YWW_{TD} (red), KWW_{TD} (blue), and both YWW_{TD} and KWW_{TD} (black). SDS denaturing acrylamide gel of (1) A₁₂₃, (2) YWW_{TD}, (3) KWW_{TD}, and (4) a mixture of all three proteins. (B) SEC-MALS of the YWW_{TD}-A₁₂₃-KWW_{TD} ternary complex shown in black. Representative isotherms of (C) YWW_{TD} titrated into the A₁₂₃-KWW_{TD} binary complex, and (D) KWW_{TD} titrated into the A₁₂₃-YWW_{TD} binary complex collected at 25 °C.

To determine thermodynamic parameters for the ternary system, we conducted ITC experiments in which YWW_{TD} was titrated into a preformed $KWW_{TD}-A_{123}$ complex or KWW_{TD} was titrated into a preformed $A_{123}-YWW_{TD}$ complex. These experiments yielded K_d values of 0.2 and 0.7 μM , respectively, close to values for the binary complexes (Table 3.4, Fig. 3.6C-D). However, the enthalpic term was lower (less negative value) when KWW_{TD} was titrated into the pre-formed $YWW_{TD}-A_{123}$ complex relative to apo A_{123} . This indicates fewer hydrogen bonds and/or van der Waals interactions in the ternary complex relative to the $KWW_{TD}-A_{123}$ complex. Titration of KWW_{TD} into YWW_{TD} resulted in no observable binding (Fig. 3.7). Altogether, this data indicates that A_{123} , KWW_{TD} , and YWW_{TD} form a ternary complex anchored by A_{123} .

Table 3.4: ITC of YWW_{TD} or KWW_{TD} titrated into the alternate A_{123} -WW domain complex collected at 25 °C. ^a Reported values are the average of duplicate experiments. ^b Values from reference [4]. ^c Values from Table 2 shown for comparison.

(syringe)	(cell)	N	K_d (μM)	ΔH (kcal/mol)	TAS (kcal/mol)	ΔG (kcal/mol)
^a YWW_{TD}	A_{123} - KWW_{TD}	0.9	0.2	-35.3	-26.0	-9.2
^a KWW_{TD}	A_{123} - YWW_{TD}	1.0	0.7	-25.8	-17.4	-8.4
<i>For comparison</i>						
^b YWW_{TD}	A_{123}	1.1	0.26 ± 0.01	-39.6 ± 0.9	-30.3 ± 0.9	-8.98 ± 0.03
^c KWW_{TD}	A_{123}	0.9 ± 0.1	0.7 ± 0.2	-37 ± 6	-29 ± 6	-8.4 ± 0.1

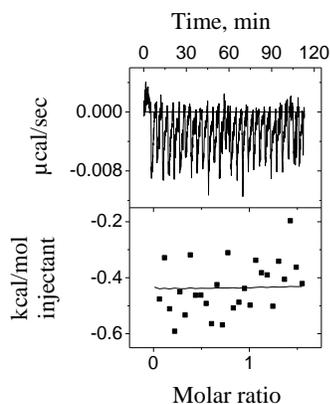


Figure 3.7: ITC isotherm of KWW_{TD} titrated into YWW_{TD} at 30 °C.

Discussion

While cooperative binding of the tandem WW domains of KIBRA to bivalent PPXY motifs is well-documented[3, 84, 85], their binding behavior for a trivalent partner has not been reported. Our work examines binding of KIBRA to trivalent AMOTL1 in context of all potential recognition sequences. In many systems, multivalency increases binding affinity through avidity[34, 39, 186]. Here, we show that the tandem WW domains of KIBRA bind AMOTL1 with strong cooperativity. However, for AMOTL1, one motif anchors the interaction while the other motifs attenuate overall complex stability. Finally, we show that YAP and KIBRA can bind AMOTL1 simultaneously, producing a stable ternary complex. The following paragraphs detail the findings and implications of this work.

The KIBRA WW domains are equally important for the stability of the AMOTL1-KIBRA complex. In this study, we show that the tandem WW domains of KIBRA are highly cooperative. Our ITC data show that abolishing either domain results in weak binding, indicating that both domains are compulsory for stable complex assembly. Furthermore, NMR titration data show peak disappearance in both WW domains at sub-stoichiometric ratios of A_{123} , indicating that both domains participate in binding simultaneously. The KIBRA WW domain cooperativity may be explained by several binding models. One domain could make contacts with P3, followed by cooperative interactions of the other domain at P1/P2. Alternatively, cooperativity may be conferred by interactions of both domains at P3, as seen in the crystal structure of mouse KIBRA and the third PPXY motif of AMOT (the parent isoform)[3]. Finally, cooperativity may be necessary to induce folding of the natively unfolded WW2 domain[84]. In this later scenario, the folded WW2 may subsequently bind other PPXY sites or play the role of a “chaperone” by facilitating more stable interactions of WW1 as has been reported for other bivalent WW domain proteins[67].

In addition to AMOTL1, the WW domains of KIBRA bind cooperatively to Synaptopodin, an essential organizer of the podocyte[84, 183], and neural and renal protein Dendrin[3, 85]. However, the contributions of each WW domain to complex assembly vary between these partners. In the case of KIBRA-Synaptopodin, WW1 initiates complex assembly and has > 12-fold higher affinity relative to WW2[84].

Similarly, the WW domains of KIBRA bind cooperatively to the second and third PPXY motifs of the partner Dendrin[85]. Multiple WW domain communication strategies underscore the diverse, partner-dependent mechanisms by which KIBRA complexes assemble.

Unique roles of the AMOTL1 PPXY motifs in the assembly of the AMOTL1-KIBRA complex. Contrary to previous studies that use incomplete PPXY constructs which implicate P3 in binding KIBRA[3, 123], we show that all three motifs play important roles in complex assembly. ITC and NMR data show that P3 has the highest affinity for KIBRA and is the first site to be perturbed upon binding. However, P1 and P2 also play important roles. NMR titration data show that in the presence of excess KIBRA, peaks disappear in all three motifs. Furthermore, in the absence of P3, AMOTL1 can still bind KIBRA, albeit weakly. Intriguingly, the affinity of the complex with all three sites is weaker than when P1 and P2 are abrogated. It is clear from these findings that while P3 initiates complex assembly, the role of P1/P2 is to modulate the stability of the complex.

Signaling protein complexes must bind transiently to ensure rapid responses to subtle cellular messages. Affinity attenuation caused by P1 and P2 may be important for ease of assembly and disassembly of the complex in response to cellular cues (Fig. 3.8A). This phenomenon has been observed for other AMOTL1 complexes and may extend to other multivalent proteins[4]. For example, in the KIBRA-Dendrin interaction, the second and third PPXY motifs have the highest affinity, while the first PPXY motif of Dendrin in isolation exhibits weak affinity for KIBRA[85]. As such, this motif may serve to attenuate overall affinity for KIBRA and/or act as a recognition site for an additional WW domain partner.

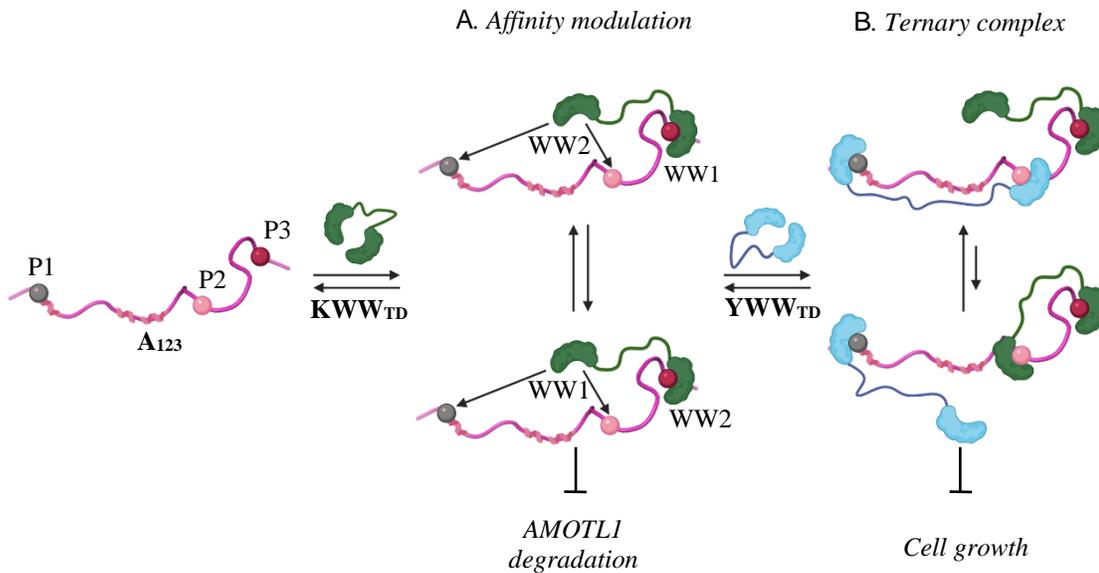


Figure 3.8: Proposed model of complex assembly. (A) Assembly of the A₁₂₃-KWW_{TD} complex is initiated through WW domain binding to the third PPXY motif of A₁₂₃, followed by weak binding at the first and second PPXY motifs which modulate binding affinity. The AMOTL1-KIBRA interaction protects AMOTL1 from proteasomal degradation[2]. (B) YWW_{TD} and KWW_{TD} can bind A₁₂₃ simultaneously, making contacts with the first and third PPXY motifs, respectively. The role of the second motif in ternary complex assembly remains elusive. The ternary complex may result in enhanced regulation of YAP subcellular localization, ultimately leading to decreased cell growth. The A₁₂₃ cartoon was adapted from the AMOT polypeptide crystal structure in complex with KIBRA (PDB: 6JJX)[3]. Created with BioRender.com.

Domain/motif specific roles in the assembly of the ternary KIBRA-AMOTL1-YAP complex. In a cellular context, many AMOTL1 WW domain partners may be present in the same sub-cellular location simultaneously. A clear example is colocalization of AMOTL1, KIBRA, and YAP in kidney cells[2]. Reconstructing the ternary complex from recombinant proteins reveals that both WW domain proteins, YAP and KIBRA, can bind simultaneously to the AMOTL1 PPXY sites. Importantly, binding of one partner (YAP or KIBRA) does not detract from the affinity of the second partner (Fig. 3.8B). More studies are needed to expound the specific roles of the WW domain/PPXY motifs in the ternary complex.

Conclusions and impact. Our work shows that the KIBRA and AMOTL1 multivalent WW domains/PPXY motifs have specific roles and use different

communication strategies to finetune complex assembly. While these findings provide unique structural insights in the assembly of multivalent WW-PPXY complexes, they also underscore the importance of studying multivalent interactions in context of all potential binding sites, not just as isolated domains or motifs. The role of destabilizing interactions likely serves to modulate the stability of AMOTL1 in the cell. When cellular concentrations of YAP are low, KIBRA may interact with all three PPXY motifs of AMOTL1, resulting in a less stable complex and increased degradation of AMOTL1. In the presence of high YAP concentrations, a ternary complex is more likely to form, which may stabilize both the YAP and KIBRA interactions with AMOTL1. This likely serves to simultaneously enhance AMOTL1 stability and increase cytoplasmic retention of YAP, resulting in decreased cell growth even when YAP concentrations are high. Future studies are needed to demonstrate the biological significance of affinity attenuation and ternary complex assembly.

Materials and Methods

Cloning of constructs and recombinant protein production are described in Appendix 1.

Analytical size exclusion chromatography (SEC) and multi-angle light scattering (MALS). Analytical SEC experiments were conducted using injections volumes of 100 μ L protein samples at 25-50 μ M which were injected onto a Superdex200 10/300 GL column (Cytiva life sciences) connected to an AKTA FPLC (GE Healthcare). For SEC-MALS experiments, this system was coupled to a Dawn multi-angle light scattering/Optilab refractive index system (Wyatt Technologies). Experiments were collected at room temperature at a flow rate of 1 mL/min in buffer containing 50 mM sodium phosphate, 0.4 M NaCl, 1 mM NaN₃, and 5 mM β -mercaptoethanol at pH 7.5. Elution profiles were recorded at 280 nm. Unless otherwise stated, molar masses were determined from triplicate experiments using ASTRA software package, version 8 (Wyatt Technologies).

Isothermal Titration Calorimetry (ITC). ITC data were recorded on a VP-ITC instrument (Malvern instruments Inc, MA) set to 25 °C unless otherwise stated. Twenty-seven or twenty-eight injections of 84-410 μ M KWW_{TD}/mutants or YWW_{TD}

were titrated into 8-37 μM A_{123} /mutants. Protein samples were dialyzed against 50 mM sodium phosphate, 50 mM sodium chloride, 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol at pH 7.5. Experiments for the ternary complex were collected with A_{123} pre-bound to molar excess KWW_{TD} or YWW_{TD} . The binding isotherms were fit using Origin 7.0 software to a single-site binding model which assumes independent binding of each site to give the average dissociation constant (K_d), a measure of the stability of the complex, and the binding stoichiometry (N) which gives the average molar ratio of the complex. Unless otherwise stated, reported data are the average of triplicate titrations. The free energy of binding (ΔG) was calculated from the equation $\Delta G = -RT\ln(K_a)$, where R is the universal gas constant, T is temperature in Kelvin, and K_a is the association constant.

Nuclear Magnetic Resonance Experiments (NMR). NMR experiments were performed on a Bruker Avance III, 800 MHz spectrometer (Bruker BioSpin) equipped with a triple resonance cryogenic probe at 10 or 25 $^{\circ}\text{C}$. ^1H - ^{15}N TROSY-HSQC spectra were collected using Echo-AntiEcho phase discrimination with 256 increments defined by 128 scans and 1024 points, with or without the BEST (band selective excitation short transient) sequences of Favier and Brutscher[169]. Binding studies of unlabeled KWW_{TD} titrated into 50-60 μM ^{15}N -labeled A_{123} at molar ratios (unlabeled:labeled) of 0.25, 0.5, 0.75, 1, and 2 were acquired at 10 $^{\circ}\text{C}$ in pH 6.8 buffer composed of 50 mM sodium phosphate, 100 mM NaCl , 50 mM arginine, 50 mM glutamate, 1 mM NaN_3 , and 2 mM tris(2-carboxyethyl) phosphine (TCEP). Binding studies of unlabeled A_{123} titrated into 200 μM ^{15}N -labeled KWW_{TD} at molar ratios (unlabeled:labeled) of 0.1, 0.2, 0.3, and 0.4 were acquired at 25 $^{\circ}\text{C}$ in pH 6.8 buffer composed of 10 mM sodium phosphate, 10 mM NaCl , 1 mM NaN_3 , and 5 mM TCEP. All samples contained 10% D_2O as the lock nucleus and 0.5% DSS for internal referencing. All NMR spectra were processed in NMRPipe[170] and visualized with Sparky[171] or NMRView[187].

Chapter 4

Domain-domain communication in multivalent WW proteins: insights from the tetravalent NEDD4-1 WW domains

Amber Vogel, Matthew McWhorter, Hannah Stuwe, and Afua Nyarko

Abstract

The NEDD4 protein family is an important set of E3 ubiquitin ligases with diverse roles in cell signaling and protein homeostasis. Dysregulation of these proteins can result in a variety of disease states, making understanding of their function and regulation critical to the control of these maladies. All NEDD4 proteins contain 2-4 WW domains involved in substrate recognition and autoinhibition of catalytic activity. While the mechanisms of binding and autoinhibition are characterized for some NEDD4 proteins, the purpose of having multiple domains remains unclear for the founding member of the family: NEDD4-1. In this study, we undertake detailed biophysical experiments to elucidate the solution properties of the complete WW domain segment of NEDD4-1. We show that the chemical environment of the primary substrate binding site – WW3 – is modulated by adjacent domains. This work shows that the WW domains of NEDD4-1 are not isolated units, but in fact behave as interconnected modules of a multivalent system. These connections are propagated by disordered linkers between adjacent domains. Interdomain communication may regulate the function of NEDD4-1.

Introduction

Neural precursor cell-expressed developmentally downregulated 4-1 (NEDD4-1) is the founding member of the NEDD4 family of E3 ubiquitin ligases which are essential to a plethora of biological functions and pathologies[129, 131]. The nine NEDD4 family members (NEDD4-1, NEDD4-2 or NEDD4L, NEDL1, NEDL2, SMURF1, SMURF2, WWP1, WWP2, and ITCH) share a conserved modular architecture composed of an N-terminal C2 lipid/Ca²⁺ binding domain, 2-4 middle WW domains (4 in NEDD4-1), and a C-terminal enzymatic HECT domain used to transfer ubiquitin from E2 ligases to substrate molecules (Fig. 4.1A). Ubiquitination of substrates results in various downstream effects including regulation of neurodevelopment, organ development, wound repair, tumor suppression, endocytosis of membrane proteins, transcription, autophagy, electrolyte balance, and protection from UV-damage to DNA[131, 188, 189]. These proteins also participate in pathologies such as tumorigenesis, Parkinson's disease, lung disorders, cardiovascular diseases, and viral infection, making them important targets for therapeutic intervention[131, 134, 138, 190, 191].

The presence of multiple WW domains in all nine NEDD4 family members suggests that WW domain multivalency plays a functional role in these proteins. Of the interactions that are characterized, the third WW domain of NEDD4-1 is the most important for complex stability with most of its partners, including organic anion transporter 1[192], commissureless[193], endothelial Na⁺ channel[194], p53-binding protein 2[88], and viral late domains from Ebola, Marburg, HTLV-1, and Rabies[88]. Despite its enhanced binding affinity for partners, the WW3 domain shares 40-61.8% sequence identity with the other domains (Fig. 4.1C). The WW3 domain has also been shown to exhibit dynamic exchange processes between folded and unfolded states both in isolation and in context of neighboring domains[195]. Thus, in terms of complex assembly, sequence, and chemical exchange, it is unclear why NEDD4-1 needs four WW domains.

Autoinhibition is reversed by phosphorylation of residues adjacent to WW1 or by binding of adaptor proteins to WW2. For SMURF2, NEDD4-1, and NEDD4-2, the autoinhibited state is stabilized through intramolecular interactions of the HECT domain with the C2 domain, WW1, and the WW1-2 linker[197-199]. However, while the WW1-2 linker of NEDD4-1 is predicted to contain a helical segment involved in autoinhibitory contacts, to our knowledge, this linker has never been structurally characterized[197]. Furthermore, of the WW domains of NEDD4-1, only WW1 is involved in autoinhibition, leaving the roles of WW2 and WW4 unclear.

Understanding the role of all four WW domains of NEDD4-1 requires a more holistic biophysical approach that includes all multivalent domains. However, dynamics, structural heterogeneity, and the large size of the WW domain region preclude residue-level characterization using traditional crystallographic and NMR approaches. In this study, we overcome these challenges by generating NEDD4-1 polypeptides containing monovalent (one domain), bivalent (two domains), and tetravalent (four domains) WW domains. We then use a structure-guided de novo design approach based on Sortase-mediated transpeptidation and complementary biophysical techniques to characterize the solution properties of NEDD4-1 WW domains in isolation and in the context of the entire multidomain region. We show that the WW domain segment is highly dynamic and structurally heterogeneous. Additionally, we show that adjacent WW domains regulate the chemical environment of their neighbors. In particular, the chemical environment of WW3 is modulated by WW2 and WW4. Domain-domain communication may fine-tune binding of the WW3 domain to its many partners. Our findings provide insight into the physical role of WW domain multivalency in NEDD4-1 and highlight the importance of characterizing domain solution properties in context of more native polypeptide segments.

Results

Construct design. NEDD4-1 is a multivalent protein which contains four WW domains. To study structural features of the complete WW domain region of NEDD4-1, a construct was designed to include all four WW domains, spanning residues 600-935 (hereafter referred to as WW1-4). Isolated WW domain constructs of WW1

(residues 600-647), WW2 (residues 645-809), WW3 (residues 807-881), and WW4 (residues 879-935) were also generated. An additional shorter construct of WW2 (residues 700-809, WW2_{short}) was generated to aid in NMR assignments of WW2. To gain insight into the effect of neighboring domains on WW domain structure, NEDD4-1 constructs containing consecutive pairs of WW domains were designed: WW1-2 (residues 600-809) and WW3-4 (residues 807-935). All NEDD4-1 constructs contained additional N- and/or C-terminal residues designed for use in Sortase-A ligation reactions (Table 4.1). Constructs and domain organization are shown in Figure 4.1B.

Table 4.1: NEDD4-1 construct tags.

Construct	Residues	N-terminal tag (non-native residues)	C-terminal tag (non-native residues)
WW1-4	600-935	GAHM-	---
WW1	600-647	SG-	-PETGGHHHHHH
WW2	647-809	GGG-	-AETGGHHHHHH
WW2_{short}	700-809	GGG-	-AETG
WW3	807-881	GGG-	-LPETGGHHHHHH
WW4	879-935	GGG-	---
WW1-2	600-809	SG-	-AETG
WW3-4	807-935	GGG-	---

Solution properties of the multivalent NEDD4-1 WW domain region. To determine the oligomerization state of WW1-4, we performed size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) (Fig. 4.2A). Triplicate experiments yielded an average molar mass of 37.7 ± 0.2 kDa in good agreement with the theoretical monomeric mass of 38.3 kDa (Table 4.2).

Table 4.2: SEC-MALS of WW1-4 and WW1-2-¹⁵N(3-4).

NEDD4	Theoretical MW, kDa	Experimental MW, kDa
WW1-4	38.3	37.7 ± 0.2
WW1-2- ¹⁵ N(3-4)	38.9	39.1 ± 0.2

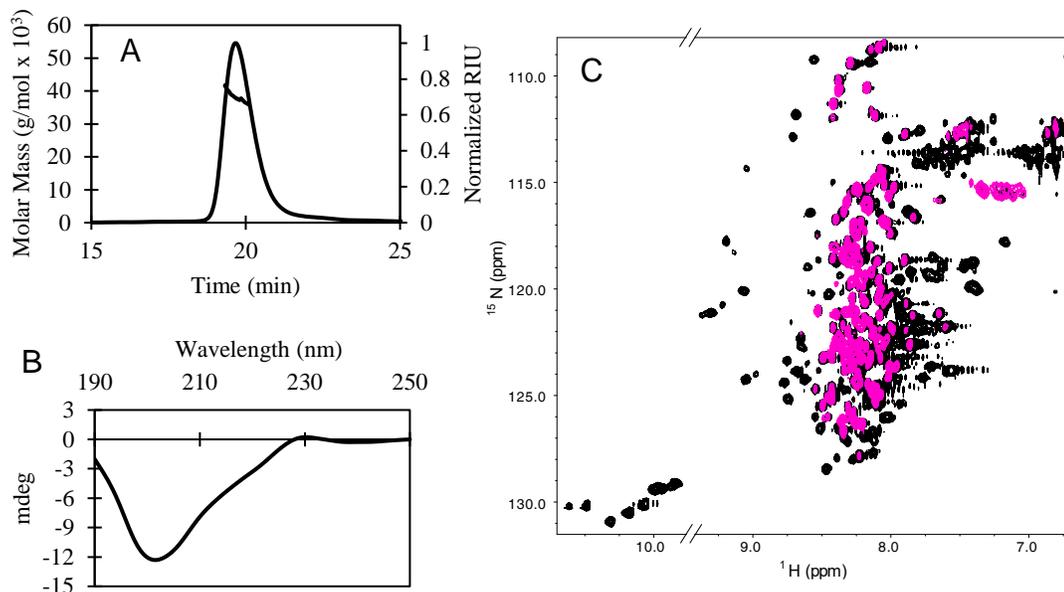


Figure 4.2: Average secondary structure and oligomerization state of WW1-4. (A) SEC-MALS elution profile of WW1-4. (B) CD spectrum of WW1-4. (C) ^1H - ^{15}N BEST-TROSY HSQC (black) and CLEANEX (magenta) spectra of WW1-4. All experiments were performed at 25 °C.

To obtain low-resolution insight into the structure of the complete WW domain region of NEDD4-1, we performed far UV circular dichroism (CD) analysis of WW1-4 (Fig. 4.2B). The CD spectrum showed a weak positive peak at 230 nm indicative of aromatic sidechain interactions which are characteristic of WW domain CD spectra[84, 200]. The large negative peak around 200 nm is consistent with disorder. Analysis of the CD data using DichroWeb showed that WW1-4 is 7.8% helical, 39.2% strands, and 53% disordered. The calculated β -strand content is consistent with the domain structure, as 40.5% of the WW1-4 residues compose WW domains, which are made of β -strands. Thus, the disordered and helical components likely correspond to linker regions.

To confirm the presence of mixed structure in WW1-4, we collected ^1H - ^{15}N -HSQC and CLEANEX experiments on ^{15}N -labeled WW1-4 (Fig. 4.2C). Of the 306 native, non-proline residues, only ~197 resonances (64%) were observed in the HSQC. Although a few resonances were well dispersed in the HSQC, the significant peak overlap in the 7.5-8.5 ppm range suggests that the multivalent polypeptide contains disorder and/or helical segments. Of the 197 peaks observed in the HSQC,

approximately 123 peaks remained in the CLEANEX spectrum, indicating that a large portion of observable peaks were solvent exposed. Altogether, this data indicates that our tetravalent WW construct, WW1-4, is monomeric in solution and contains a mixture of structured and disordered segments.

Solution properties of the isolated NEDD4-1 WW domains. Since the WW1-4 NMR spectrum was missing ~36% of peaks, and most of those present appeared as a congested blob, we did not attempt to obtain residue-specific assignments for the tetravalent domains. Instead, we split WW1-4 into four pieces, each containing a WW domain and connecting linker residues. To see if splitting WW1-4 into smaller pieces changed its solution properties, we compared each monovalent domain spectrum to that

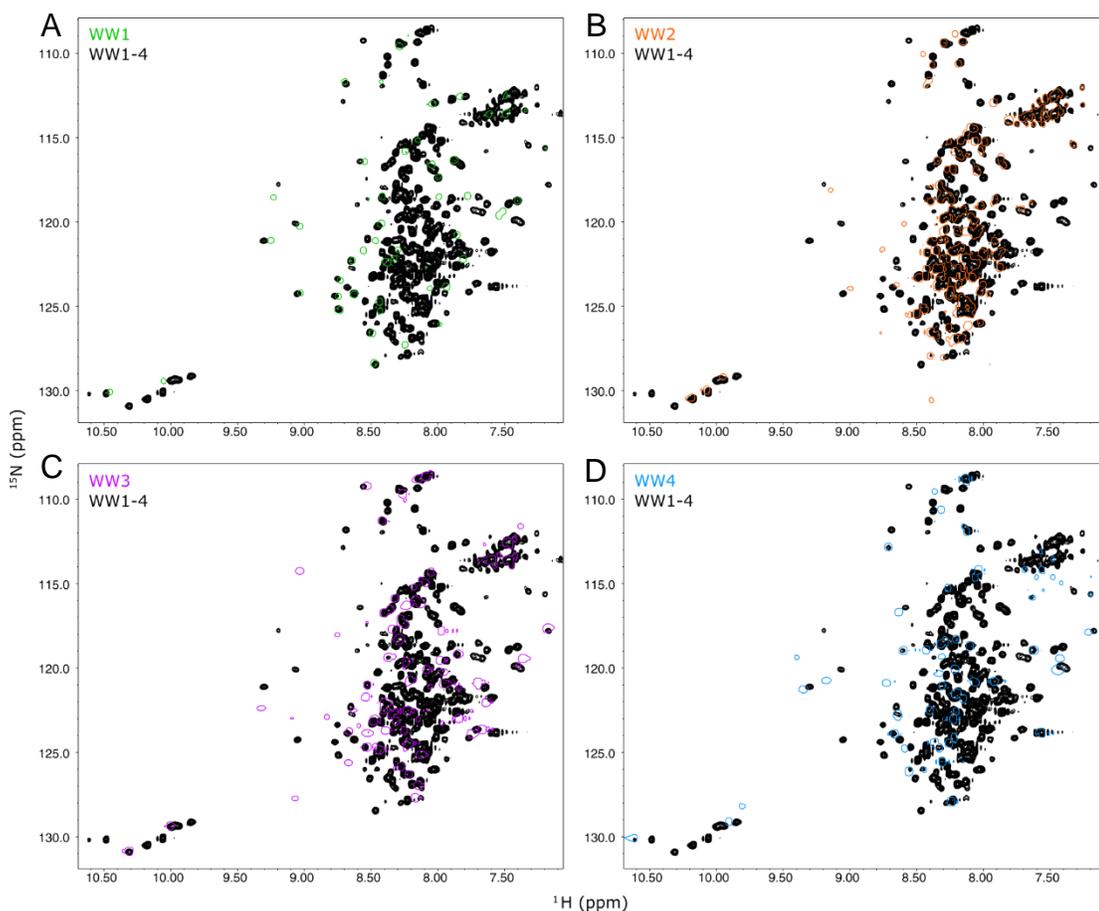


Figure 4.3: NMR comparison of isolated NEDD4-1 domains and WW1-4. ^1H - ^{15}N -BEST-TROSY HSQC spectra of (A) WW1 (green), (B) WW2 (orange), (C) WW3 (purple), and (D) WW4 (blue) overlaid on WW1-4 (black). All spectra were collected at 25 °C.

of tetravalent WW1-4 (Fig. 4.3). Slight changes were observed in all isolated domains, but particularly in WW3. Isolated domains showed new peaks and chemical shift perturbations (CSPs) relative to the tetravalent domains. These differences reflect changes in the chemical environment and/or exchange processes of individual domains when put in context of their neighbors.

To map residue-specific changes in the chemical environment of the isolated versus multivalent WW domains, resonance assignments were obtained for the monovalent proteins. Unambiguous assignments were made for 37 (88%), 103 (67%), 55 (81%), and 41 (79%) native, non-proline residues in WW1, WW2, WW3, and WW4, respectively (Fig. 4.4). Unfortunately, since many peaks in the WW domain of the WW2 construct were missing, likely due to dynamics of this domain on the intermediate timescale, only 11 (35%) domain residues were assigned. Intriguingly, several residues flanking the WW2 and WW3 domains showed up as two distinct peaks in their corresponding HSQCs, indicating multiple conformations in slow chemical exchange.

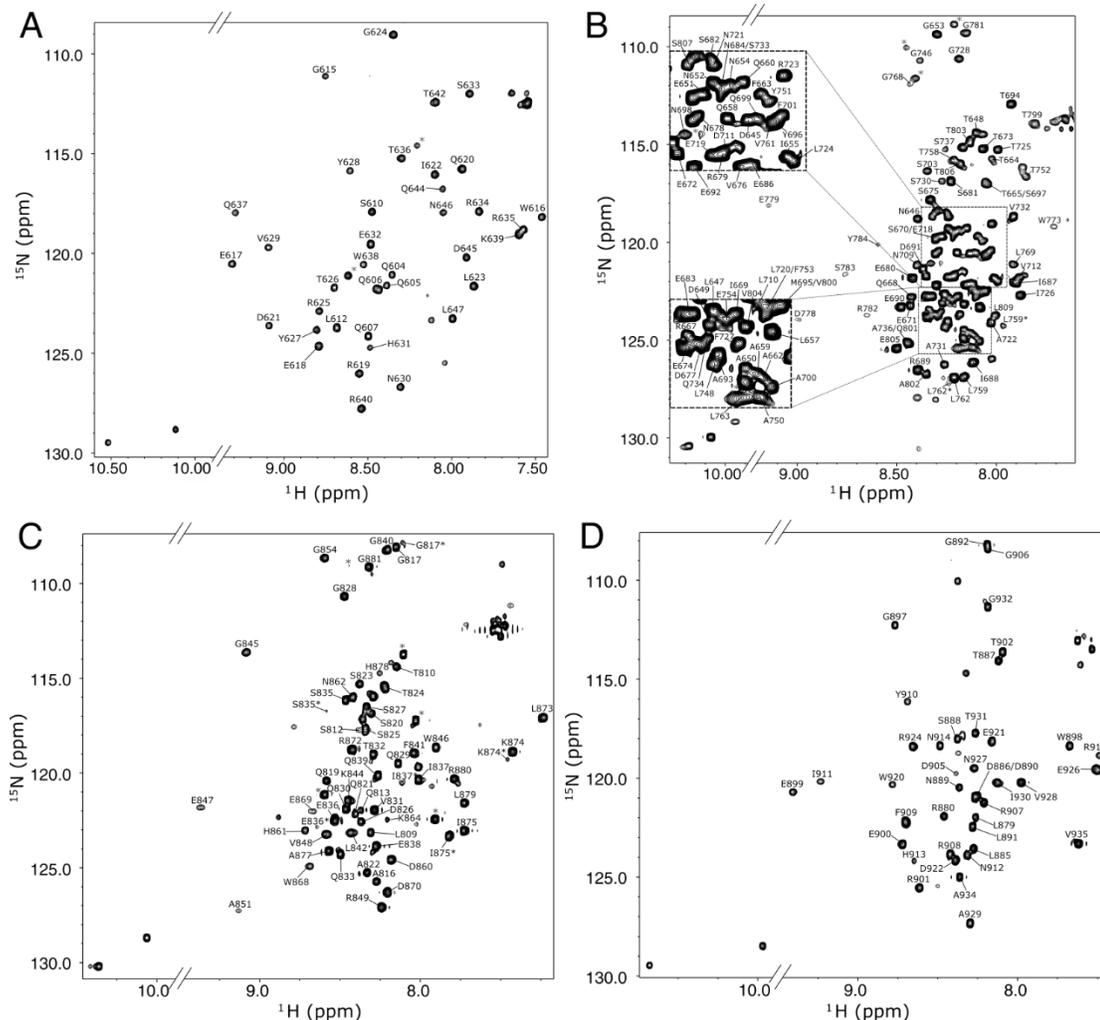


Figure 4.4: ^1H - ^{15}N -HSQC spectra and assignments for isolated NEDD4-1 WW domains. ^1H - ^{15}N -HSQC spectra showing assignments for (A) 37 of 42 non-proline WW1 residues, (B) 103 of 154 non-proline WW2 residues, (C) 55 of 68 non-proline WW3 residues, and (D) 41 of 52 non-proline WW4 residues. Several residues flanking the WW2 and WW3 domains show up as two distinct peaks in the HSQC spectra and are noted with asterisks beside the residue name: I726*, L759*, and L762* near WW2; and G817*, S835*, E836*, I837*, K874*, and I875* near WW3. Peaks identified as belonging to non-native residues from the expression vector are noted with asterisks. All spectra were collected at 25 °C.

Residue-specific analyses were conducted to characterize the secondary structure of each isolated WW domain and linker region. Aside from one helix in the WW1-2 linker (residues 650-665), PSIPRED secondary structure predicted the WW domain linkers to be disordered. Figure 4.5 shows deviations of experimentally determined CA and CB chemical shifts relative to a random coil ($\Delta\text{Ca}-\Delta\text{Cb}$) and Talos-

N calculated secondary structure and confidence plotted versus residue number. Helices and strands determined using $\Delta\text{Ca}-\Delta\text{Cb}$ values were defined as stretches of \geq four consecutive positive values and \geq three consecutive negative values, respectively. WW1 clearly shows three stretches of sequential negative $\Delta\text{Ca}-\Delta\text{Cb}$ values, consistent with the triple, anti-parallel β -strand fold of WW domains. However, since WW domains are well characterized, and missing assignments precluded the analysis of consecutive residues in the other domains, further inspection focused on linker residues.

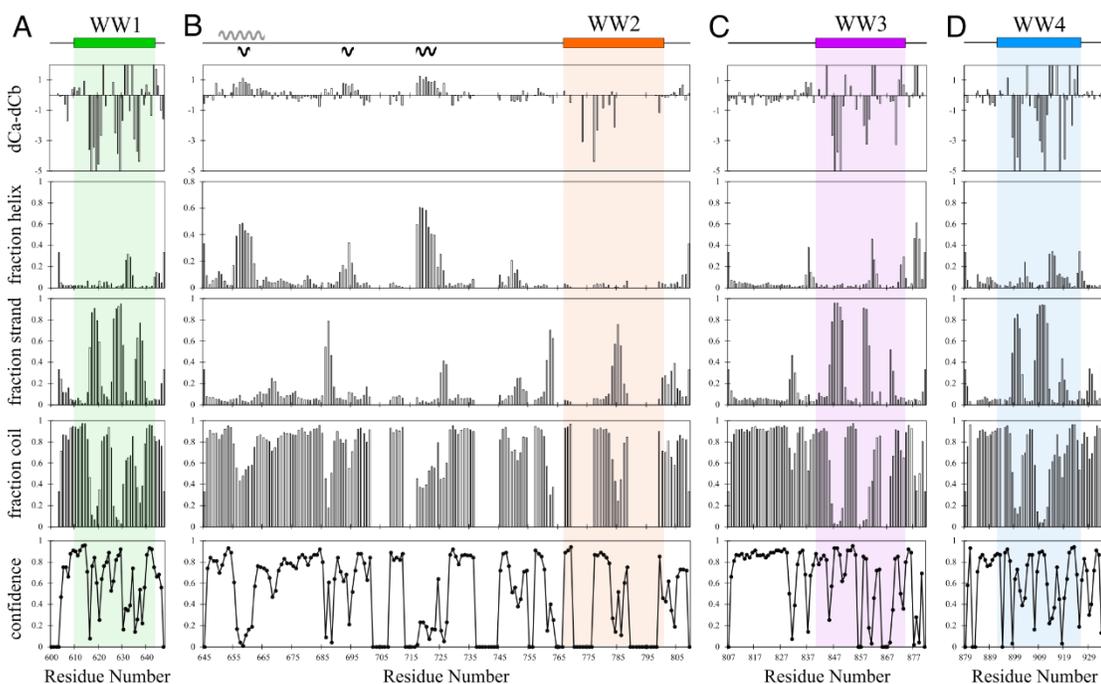


Figure 4.5: Residue-level secondary structure characterization of isolated NEDD4-1 WW domains. A plot of the deviations of CA and CB chemical shifts relative to those of a random coil (top plot, cut off at +2 and -5), Talos-N calculations of fraction helix, strand, and coil (middle plots), and Talos-N confidence level (bottom plot) plotted versus residue number for (A) WW1, (B) WW2, (C) WW3, and (D) WW4. For clarity, only values determined from experimental chemical shift data are included. Line diagrams are shown above each plot with regions belonging to WW domains shaded in green (WW1), orange (WW2), purple (WW3), and blue (WW4). The predicted and experimentally determined secondary structural elements of linker regions are shown above and below the line diagrams, respectively.

Residues flanking WW1, WW4, and within the WW2-3 linker showed only small deviations in $\Delta\text{Ca}-\Delta\text{Cb}$ values, indicating a lack of structure. In agreement, Talos-

N calculations show that these regions have a high degree of disorder (Fig. 4.5). In the WW1-2 linker, three short stretches of residues showed positive carbon chemical shift deviations (Fig. 4.5B). The average positive $\Delta\text{Ca}-\Delta\text{Cb}$ value of 0.474 for WW2 was used as a cutoff to define helical propensities. Regions with helical propensity corresponded to residues 657-660 (helix 1), 692-695 (helix 2), and 717-723 (helix 3). Otherwise, $\Delta\text{Ca}-\Delta\text{Cb}$ values and Talos-N analysis showed that the WW1-2 linker is disordered.

Next, we performed HetNOE, R_2/R_1 , and CLEANEX NMR experiments which measure dynamics on different NMR timescales (Fig. 4.6). Since unambiguous assignment of WW2 peaks proved challenging because of extensive peak overlap (Fig. 4.4B), only data for WW1, WW3, and WW4 were analyzed. Residues within the WW1, WW3, and WW4 domains had average HetNOE values of 0.70, 0.69, and 0.73, respectively, indicating moderately rigid structure on the ps-ns timescale. Residues flanking the domains showed tapered decreases in HetNOE values, indicating a loss of structure moving away from the domains. As expected, HetNOE values for residues at the N-terminus of WW1 dropped rapidly, indicating high flexibility. Decreases in HetNOE values C-terminal to WW1 and WW3 were less pronounced due to the presence of 12 non-native residues left behind from the expression vector.

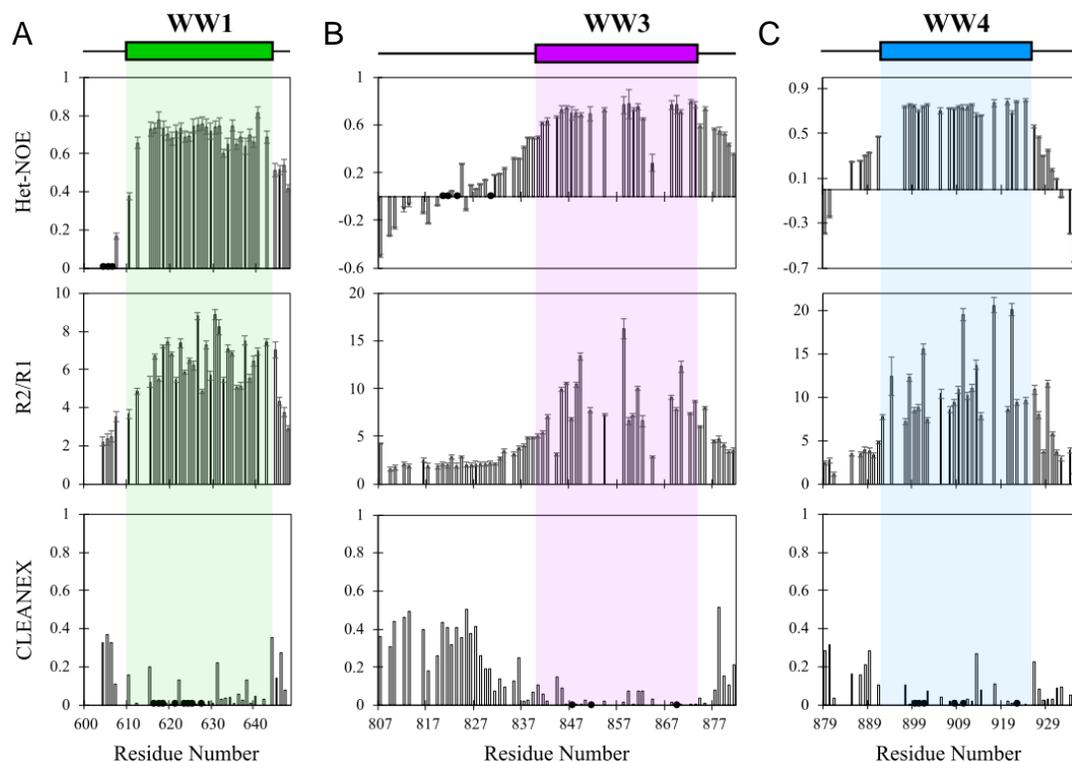


Figure 4.6: Residue-level dynamics of WW1, WW3, and WW4 determined by NMR. HetNOE (top plot), R_2/R_1 (middle plot), and normalized CLEANEX intensity values (bottom plot) plotted versus residue number for (A) WW1, (B) WW3, and (C) WW4. Regions corresponding to the WW domains are shaded in green (WW1), purple (WW3), and blue (WW4). Only data for unambiguously assigned, non-overlapping peaks are shown. Values of zeros are shown as black dots along the X-axes. All data were collected at 25 °C.

Next, we analyzed R_2/R_1 values, which report on the protein tumbling time. As expected, residues within the domains showed higher R_2/R_1 values than linkers, indicating slower tumbling times. Finally, we measured the solvent accessibility of amide protons using CLEANEX NMR experiments. Aside from a few outliers within each domain, CLEANEX values were low or zero within the WW domains themselves, indicating protection from solvent. Moving away from the domains, residues showed higher CLEANEX values, particularly in the WW2-3 linker, indicating more solvent accessibility. To summarize, (i) WW1, WW3, and WW4 show characteristic WW domain structure, and (ii) domain linker residues are predominantly unstructured except for three short regions which show helical propensity in the WW1-2 linker.

Residue-level WW domain comparison show subtle differences in dynamics.

Since most NEDD4-1-substrate interactions are mediated through by the WW3 domain, we next asked if are differences in apo WW domain dynamics or structure. Since few assignments were made in the WW2 domain, we again focused only on WW1, WW3, and WW4. Figure 4.7 shows HetNOE, R_2/R_1 , and normalized CLEANEX peak intensities plotted versus relative WW domain position (1-34), with data for seven flanking residues (-1 to -7 and +1 to +7) shown on either side for context. Differences between the domains are shown on the apo NEDD4-1 WW3 domain structure.

Within the domains themselves, no major deviations in HetNOE values were apparent (Fig. 4.7A-B). However, residues on either side of the WW3 domain showed significantly higher HetNOE values than WW1 or WW4, indicating reduced ps-ns flexibility of residues flanking WW3. Intriguingly, five residues flanking WW3 produced two peaks in the ^1H - ^{15}N -HSQC spectrum, indicating slow chemical exchange of residues in the vicinity of WW3 (Fig. 4.4C). In comparing the R_2/R_1 values, residues G845 and V848, which bracket the first Trp residue important for domain stability[201], and two residues in loop 1 showed enhanced rigidity in WW3 relative to WW1 and WW4 (Fig. 4.7C-D). The tip of the second β strand showed more flexibility in WW3 than WW1 and WW4. In analyzing the CLEANEX data, we found that residues with enhanced solvent protection include residues in loop 1 and C-terminal to WW3 (Fig. 4.7E-F). In contrast, I859 in strand 2 is more solvent exposed than in WW1 or WW4.

In summary, we show differences in dynamics of WW3 relative to WW1 and WW4 on several timescales which may explain the favored binding of WW3 to substrates. First, these data indicate that WW3-flanking residues fall within a unique exchange regime relative to WW1 and WW4. The reduced ps-ns timescale dynamics and slow chemical exchange of WW3-flanking residues may reduce the entropic cost of binding. Second, residues in the vicinity of strand 1 of WW3 show increased rigidity and decreased solvent exposure relative to the other domains, which likely help stabilize the overall WW domain fold. Third, loop 1 residues show reduced flexibility and solvent exposure, which could reduce entropy loss associated with binding. Finally,

residues at the tip of strand 2 show greater flexibility and solvent exposure, which may be important for partner recognition.

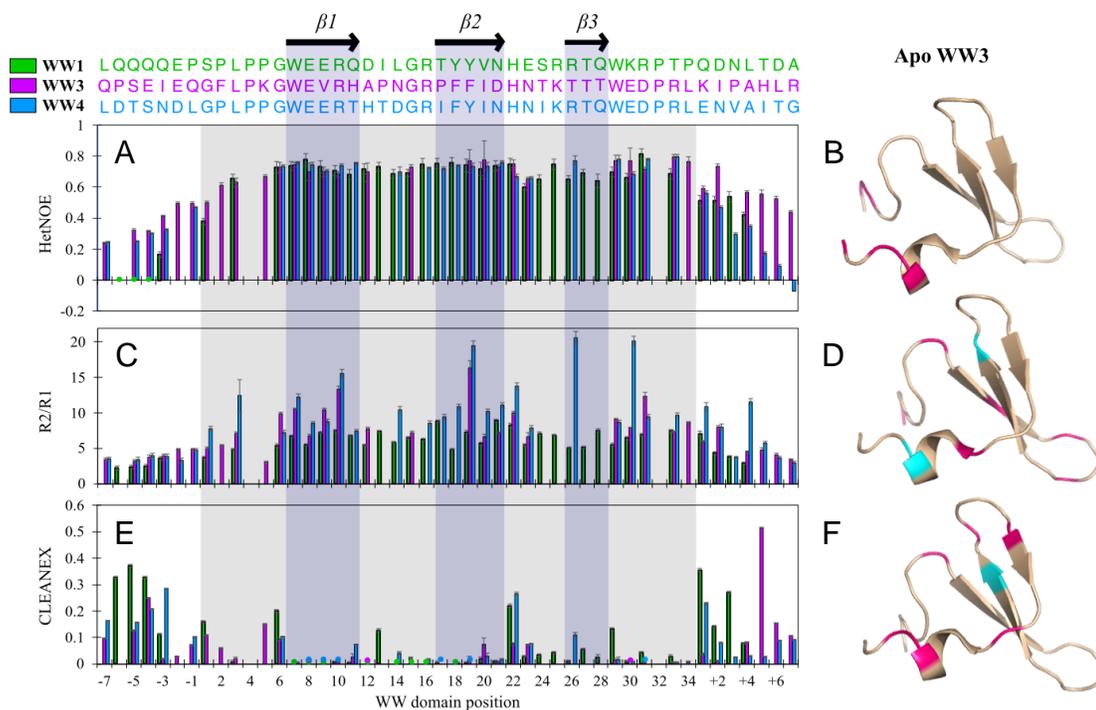


Figure 4.7: Interdomain differences in dynamics of WW1, WW3, and WW4. (A) HetNOE, (C) R_2/R_1 , and (E) normalized CLEANEX intensity values plotted versus relative WW domain position for WW1 (green), WW3 (purple), and WW4 (blue). Values of zero are indicated by circles on the X-axes of plots A and E. The WW domain region is shaded in grey. β -strand segments are shaded in blue. The WW domain sequences are shown above plot A. (B, D, F) Structures of apo NEDD4-1 WW3 (PDB: 5AHT). (B) WW3 residues with HetNOE values higher than WW1 or WW4 are shown in pink. (D) WW3 residues with R_2/R_1 values higher than WW1 or WW4 are shown in pink. Residues with R_2/R_1 values lower than WW1 or WW4 are shown in cyan. (F) WW3 residues with CLEANEX values lower than WW1 or WW4 are shown in pink. WW3 residues with CLEANEX values higher than WW1 or WW4 are shown in cyan.

WW2 modulates the chemical environment of WW1. Peak heterogeneity, overlap, and missing resonances posed challenges in characterizing the NEDD4-1 WW domains in a multivalent context. Thus, to determine changes in chemical environment of isolated versus tandem domains, we compared the chemical shifts and peak intensities of isolated domains to their multivalent counterparts. First, we compared the spectra of monovalent WW1 and WW2 to bivalent WW1-2 (Fig. 4.8A-B). Several

WW1 peaks were missing or showed CSPs relative to the WW1-2 spectrum. In contrast, peaks in WW2 and the WW1-2 linker showed only small CSPs in the isolated domain constructs relative to the tandem domains. Next, we compared the spectrum of bivalent WW1-2 to tetravalent WW1-4 (Fig. 4.8C). Almost every peak in the WW1-2 spectrum had an exact matching peak in the WW1-4 spectrum, indicating that WW3-4 does not affect the chemical environment of WW1-2. Altogether, these data indicate that the chemical environment of WW1 is changed by the presence of WW2/the linker segment but is unaffected by WW3 and WW4.

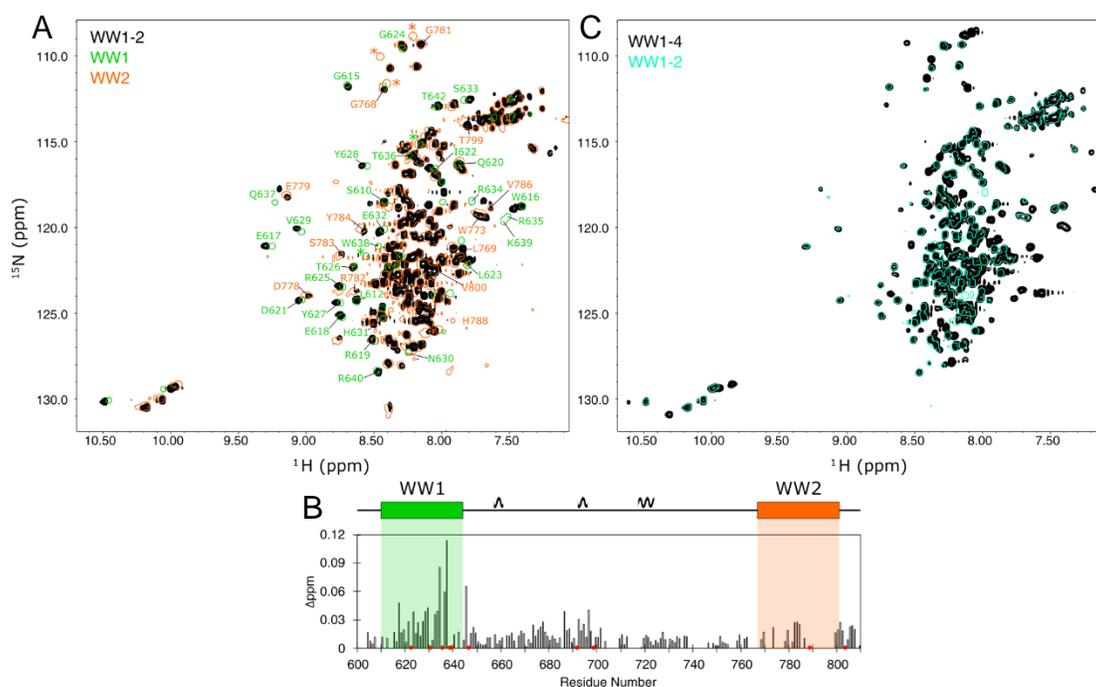


Figure 4.8: ^1H - ^{15}N -HSQC comparison of NEDD4-1 WW1 and WW2 to tandem WW1-2. (A) ^1H - ^{15}N -BEST-TROSY HSQC spectra of WW1 (green) and WW2 (orange) overlaid on WW1-2 (black). Peaks corresponding to the WW domains are labeled. Peaks corresponding to non-native residues from the expression vector are indicated with asterisks. (B) Chemical shift perturbations of WW1 and WW2 relative to WW1-2 plotted versus residue number. The locations of the domains and experimentally determined helices are indicated. Peaks in the isolated domains with no corresponding peak in the tandem domain spectrum are indicated with red dots on the X-axis. (C) ^1H - ^{15}N -BEST-TROSY HSQC spectra of WW1-2 (turquoise) and overlaid on WW1-4 (black). All spectra were collected at 25 °C.

WW3 and WW4 chemical environments are modulated by adjacent domains. Of all the domains, WW3 experienced the most spectral changes when in context of its neighbors (Fig. 4.3C). This finding prompted us to examine which neighboring domains are responsible for modulating the chemical environment of WW3. To see if WW3 is modulated by WW4, we examined differences in the chemical shifts and peak intensities of monovalent WW3 and WW4 relative to bivalent WW3-4 (Fig. 4.9). As expected, numerous peaks at the junction between the isolated domain constructs

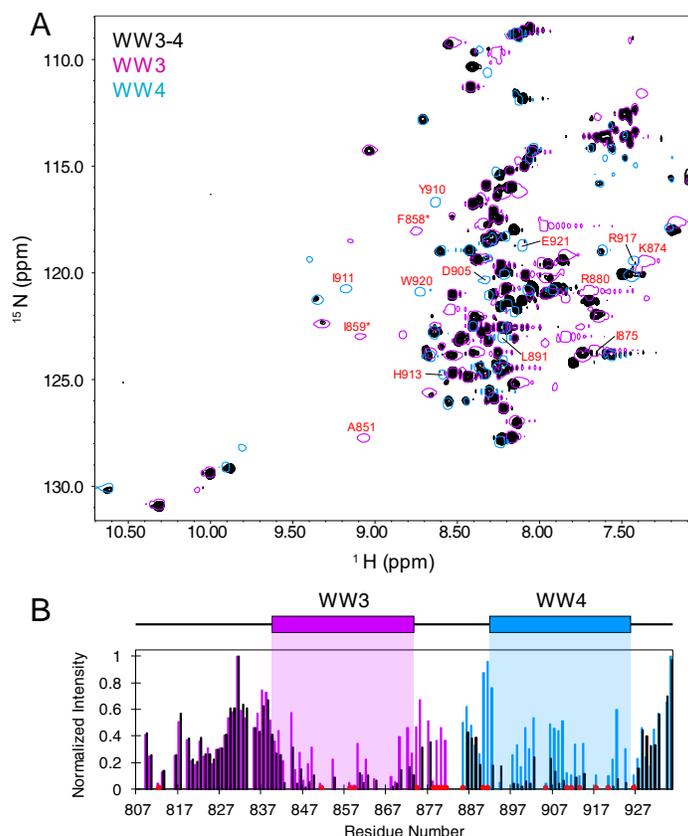


Figure 4.9: ^1H - ^{15}N -HSQC comparison of WW3 and WW4 to WW3-4. (A) ^1H - ^{15}N -HSQC overlay of WW3 (purple, open peaks) and WW4 (blue, open peaks) on WW3-4 (black, filled peaks). Isolated WW domain assigned peaks without a corresponding peak in the WW3-4 spectra are labeled in red. Ambiguous assignments are shown with asterisks. (B) Normalized peak intensity of isolated WW3 (purple bars), WW4 (blue bars), and WW3-4 (black bars) plotted versus residue number. Peak intensities were normalized to the most intense peak in each spectrum (V831 in WW3, V935 in WW4, and V831 in WW3-4). Residues in the isolated domains which lack a corresponding peak in the WW3-4 spectrum are shown as red dots along the X-axis. The WW domain architecture is shown above each bar plot, where WW3 and WW4 are shaded in purple and blue, respectively. All data were collected at 25 °C.

lacked a clear counterpart in the bivalent spectrum, since these previously terminal residues were now in the middle of the construct bridging WW3 and WW4. Intriguingly, thirteen isolated domain peaks lacked a clear counterpart in the WW3-4 spectrum: four and nine peaks in the vicinity of WW3 and WW4, respectively. Monovalent WW3 and WW4 domain normalized peak intensities were more intense than in the bivalent spectrum, while residues on the N- and C-termini showed little change in intensity. This data indicates that WW3 and WW4 modulate the chemical environment of each other, while terminal residues are unchanged.

Next, we were curious if the presence of WW1-2 caused additional changes in the WW3-4 domains. However, extensive peak overlap in the tetravalent spectrum made comparison of bivalent WW3-4 to tetravalent WW1-4 challenging. To overcome this challenge, we used evolved Sortase A (eSrt) transpeptidation chemistry to ligate ^{15}N -labeled WW3-4 with an N-terminal GGG tag to unlabeled WW1-2 with a C-terminal LPETGG- H_{6x} tag, yielding the segmentally labeled WW1-2-LPET- ^{15}N (GGG-WW3-4) product, hereafter referred to as WW1-2- ^{15}N (3-4) (Fig. 4.10A-B). This approach allowed residue-specific analysis of WW3-4 in context of NMR-invisible WW1-2. To validate ligation of the intended WW domain substrates, we performed SEC-MALS and obtained a molar mass of 39.1 ± 0.2 kDa in good agreement with the theoretical mass of 38.9 kDa for the ligated product (Table 4.2). Resonances present in the WW1-2- ^{15}N (3-4) spectrum closely matched those in the WW1-4 spectrum, indicating no drastic change in the solution structure of WW1-2- ^{15}N (3-4) due to the seven-residue LPETGGG linker relative to WW1-4 (Fig. 4.10C).

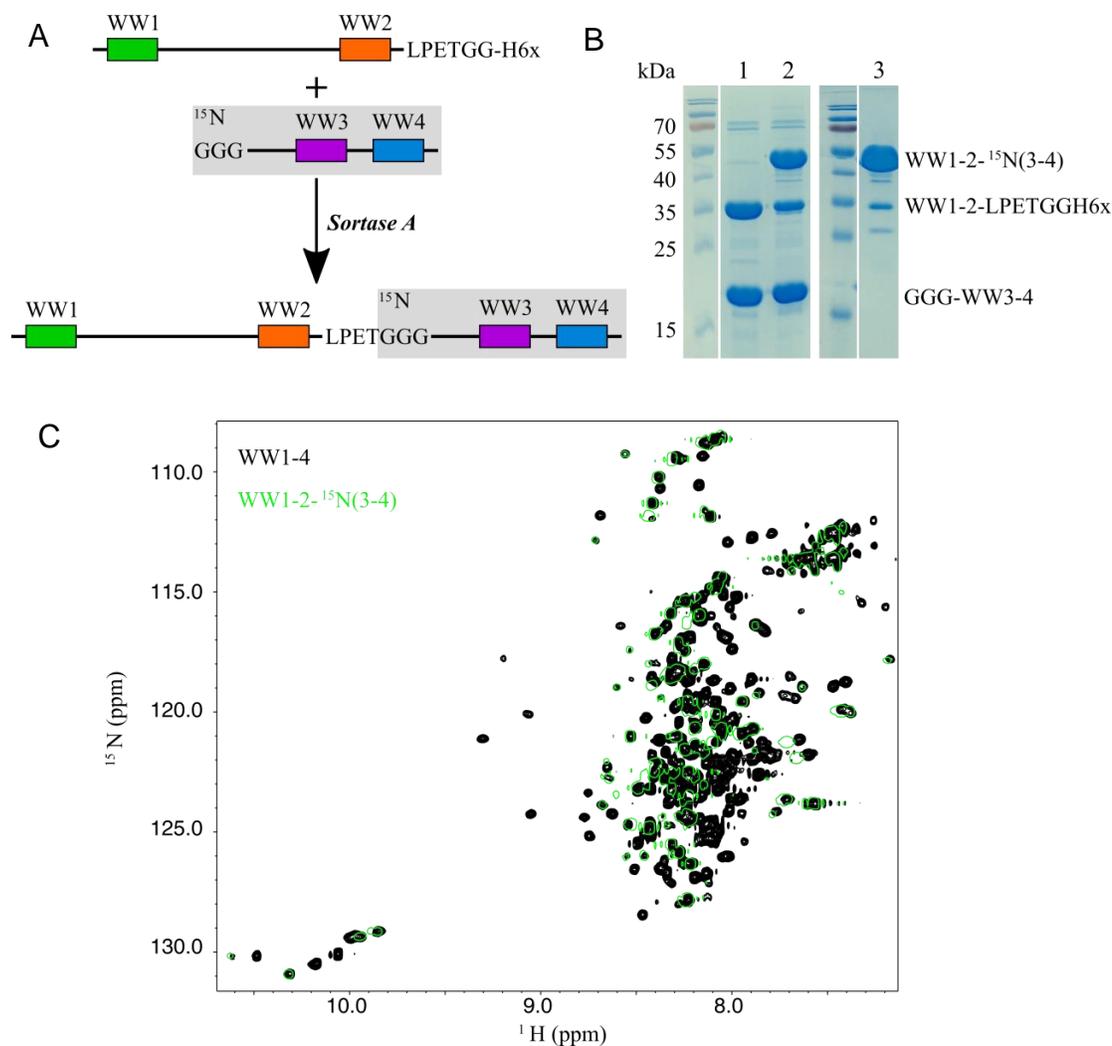


Figure 4.10: Sortase A-mediated ligation and NMR validation of ^{15}N -WW3-4 to unlabeled WW1-2. (A) Sortase-mediated transpeptidation reaction scheme showing substrates WW1-2-LPETGG- H_{6x} + ^{15}N -labeled GGG-WW3-4 \rightarrow product WW1-2-LPET- ^{15}N (GGG-WW3-4). (B) SDS-PAGE 15% acrylamide denaturing gel showing (lane 1) WW1-2-LPETGG- H_{6x} and ^{15}N -GGG-WW3-4 unligated substrates, (lane 2) reaction progress after incubation with eSrt for 2 h at 4 °C, and (lane 3) purified segmentally labeled eSrt product, WW1-2- ^{15}N (3-4). (C) ^1H - ^{15}N -HSQC overlay of segmentally labeled WW1-2- ^{15}N (3-4) (green, open peaks) on uniformly ^{15}N -labeled WW1-4 (black, filled peaks) collected at 25 °C.

Next, we compared the HSQC spectra of ^{15}N -labeled WW3-4 and WW1-2- ^{15}N (3-4) (Fig. 4.11). Almost all peaks observed in the HSQC spectrum of ^{15}N -labeled WW3-4 were present in the WW1-2- ^{15}N (3-4) spectrum with a few exceptions (Fig. 4.11A). Figure 4.11B is a bar graph of normalized peak intensity versus residue number

for WW3-4 and WW1-2- ^{15}N (3-4). Interestingly, residues in the vicinity of the WW3 domain showed notably weaker/missing peak intensities in the WW1-2- ^{15}N (3-4) spectrum relative to WW3-4, while disordered residues showed little change in intensity. Although these changes could result from the large size of the tetravalent

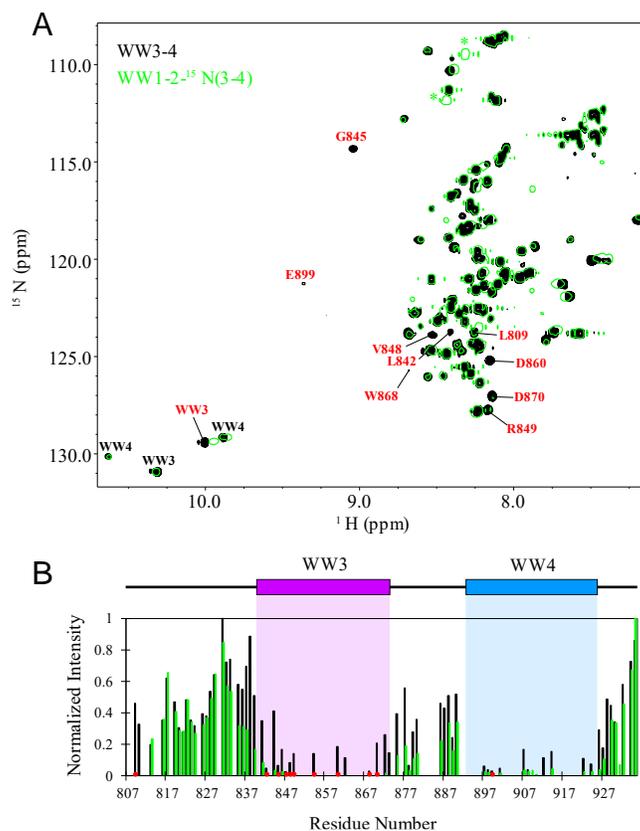


Figure 4.11: ^1H - ^{15}N -HSQC comparison of WW3-4 to WW1-2- ^{15}N (3-4). (A) ^1H - ^{15}N -HSQC overlay of WW1-2- ^{15}N (3-4) (green, open peaks) on ^{15}N -labeled WW3-4 (black, filled peaks). Eleven WW3-4 amide backbone resonances lack a corresponding peak in the WW1-2- ^{15}N (3-4) spectrum, nine of which are labeled on the spectrum in red. Gly residues from the eSrt recognition tag show up only in the WW1-2- ^{15}N (3-4) spectrum and are noted with green asterisks. The Trp sidechain peaks are labeled according to which domain they belong to. One Trp sidechain peak from WW3 is significantly shifted in the WW1-2- ^{15}N (3-4) spectrum (labeled in red). (B) Normalized peak intensity of WW3-4 (black bars) and WW1-2- ^{15}N (3-4) (green bars). Peak intensities are normalized to the most intense peak in the spectrum: V831 in WW3-4 and V935 in WW1-2- ^{15}N (3-4). Residues in WW3-4 which lack a corresponding peak in the WW1-2- ^{15}N (3-4) spectrum are shown as red dots along the X-axis. The WW domain architecture is shown above each bar plot, where WW3 and WW4 are shaded in purple and blue, respectively. All data were collected at 25 °C.

domain protein, another explanation is that WW3 experiences changes in chemical environment in the presence of WW1-2. Consistent with this explanation, a Trp sidechain peak belonging to WW3 is shifted in the WW1-2-¹⁵N(3-4) spectrum relative to its position in the WW3-4 spectrum. These spectral changes indicate that WW3 is modulated by the NEDD4-1 segment containing WW1 and WW2.

In summary, chemical environments of the NEDD4-1 WW domains are modulated by their nearest neighbors: WW2 modulates WW1, WW3 modulates WW4, and WW3 is modulated by both WW2 and WW4.

Discussion

The presence of 2-4 WW domains within all nine NEDD4 family members suggests a role for multivalency in the biological activities of these proteins. However, structural heterogeneity and the relatively large sizes of intact multivalent domains make *in vitro* characterization challenging. Therefore, studies tend to focus on isolated WW domains. In some cases, this is an appropriate workaround in obtaining insight into structure and binding. However, in most cases, this approach is inadequate, as relevant information about domain-domain communication is lost. In this work, we use a structure-guided de novo design approach based on Sortase-mediated transpeptidation to study the NEDD4-1 WW domains in isolation and as part of the intact multivalent WW domain construct. As discussed below, significant differences in NMR-detected chemical environment of the intact multivalent domain segment relative to the isolated domains suggest domain-domain communication that could modulate the structure and binding of NEDD4-1.

The NEDD4-1 WW domain segment is large (336 residues), contains > 50% disorder, and is highly dynamic[195]. These features make structural characterization immensely challenging. To work around these factors, we produced NEDD4-1 segments of varying length and prepared segmentally labeled NEDD4-1 WW domains to reduce spectral crowding for NMR analysis. Comparison of the chemical shifts and peak intensities of monovalent, bivalent, and tetravalent WW domains revealed that the chemical environments of each domain are modulated by adjacent domains (Fig. 4.12).

Interdomain communication is propagated by connecting linker segments, which are disordered in the multivalent domain context.

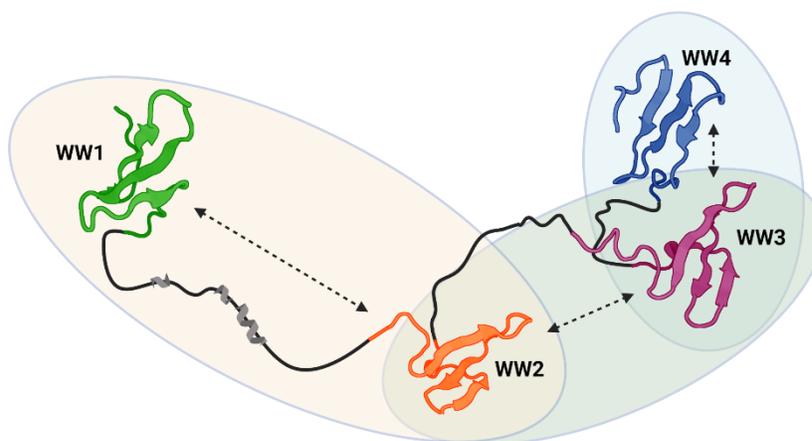


Figure 4.12: Proposed model of NEDD4-1 WW domain interconnectivity. Structures of representative WW domains connected by disordered/helical linkers, interconnectivity between WW1 and WW2, WW2 and WW3, and WW3 and WW4. Structures shown are of rat NEDD4 WW1 (green, PDB: 2N8S), rat NEDD4 WW2 (orange, PDB: 2N8U), human NEDD4-1 WW3 (purple, PDB: 5AHT) and human WWP1 WW4 (blue, PDB: 2OP7). Created with Biorender.com.

WW1 and WW4, which bracket the WW domain region, show minor spectral changes when linked to WW2 and WW3, respectively. This indicates that only adjacent domains are needed for stabilization of WW1 and WW4. In contrast, WW3 experienced greater spectral changes, particularly reduced peak intensity and peak disappearance. These changes occurred both upon ligation of WW4 and WW1-2, indicating that multiple neighboring domains modulate WW3. Since WW1 was unaffected by WW3-4, it is likely that WW3 is only modulated by the WW2 domain within the WW1-2 segment. The disappearance of WW3 peaks in the presence of neighboring domains could result from a gain in structure reflected in slower tumbling and faster T₂ relaxation. Alternately, peak disappearance could result from enhanced chemical exchange processes on the intermediate NMR timescale. Future work is needed to determine the specific effects of neighboring domains on WW3 dynamics. However, it is clear that adjacent domains modulate the chemical environment of WW3. This work underscores the importance of studying multivalent domains in the context of all relevant domains.

Functional implications. Many studies highlight WW3 as the predominant binding site for NEDD4-1 substrates[88, 192-194]. However, WW domain multiplicity is known to participate in substrate binding through multiple mechanisms: cooperativity through additive but separate interactions, cooperativity through interdependent interactions, and intramolecular domain chaperoning[67]. Our work shows that neighboring domains significantly impact the chemical environment of WW3. This domain interconnectivity may serve to regulate the binding behavior of WW3 to substrate molecules. The functions of NEDD4-1 substrates range from enhancing cell growth and tumorigenesis to promoting apoptosis[131]. Modulation of affinity through multivalent interactions may be important for controlling the stability of growth regulator proteins.

Materials and Methods

Cloning and protein production and purification are described in Appendix 1.

Circular Dichroism. Far UV circular dichroism (CD) data of 3 μ M WW1-4 dialyzed against 10 mM sodium phosphate, pH 7.5 were recorded on a JASCO 720 spectrophotometer, using a path length of 1 mm and a bandwidth of 1.0 nm. Reported data are the average of 5 scans collected at 25 °C. Secondary structure content was calculated using DichroWeb with the Contin-LL method and SP175 reference set [154, 177, 178].

Multi-Angle Light Scattering. Absolute molecular weights were obtained by size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) using an AKTA FPLC (GE Healthcare), a DAWN multiple-angle light scattering, and Optilab refractive index system (Wyatt Technology). Data were collected at room temperature, on a Superdex²⁰⁰ 10/300 GL column (Cytiva life sciences) pre-equilibrated with 50 mM sodium phosphate, 50-400 mM NaCl, 1 mM NaN₃, and 5 mM β -mercaptoethanol at pH 7.5 buffer. Protein samples at concentrations of 25 μ M were injected at a flow rate of 0.7-0.8 mL/min. The data were analyzed with ASTRA software package, version 8 (Wyatt Technology).

Preparation of Sortase A-ligated products. Sortase A reactions for production of segmentally labeled NMR samples were prepared on a 20-25 mL scale using 1:2

ratio WW1-2:WW3-4 at concentrations of 50 μ M (WW1-2) and 100 μ M (WW3-4) in buffer composed of 50 mM MOPS, 150 mM NaCl, 5 mM tris (2-carboxyethyl) phosphine (TCEP), and 1 mM NaN_3 at pH 7.5 (buffer A). Substrates were combined with 0.25 equivalents evolved Sortase A (eSrt)[202] with LPXTG specificity[203] (where X is E in our system) and incubated at 4 °C for 2 hr at 100 rpm on an Innova 200 platform shaker (New Brunswick™). Reactions were quenched by Ni-NTA affinity chromatography. Specifically, 3 mL Ni-NTA resin were added to the reaction and incubated at 4 °C for 1 hr at 100 rpm as described above. Sortase-ligated product and unreacted WW3-4 eluted in buffer A containing 0 – 20 mM imidazole. Unreacted WW1-2 and eSrt contain His_{6x}-affinity tags eluted in buffer A containing 350 mM imidazole. The Sortase-ligated product was further purified using anion exchange chromatography in buffer composed of 20 mM Tris, pH 8.0, at a concentration of 0.22-0.26 M NaCl.

Nuclear Magnetic Resonance data collection and analysis. NMR experiments were performed on a Bruker Avance III, 800 MHz spectrometer (Bruker BioSpin) equipped with a triple resonance cryogenic probe. Isotopically labeled NEDD4-1 proteins at concentrations of 0.14-1 mM were dialyzed against 50 mM MOPS, 50 mM NaCl, 1 mM NaN_3 , and 2 mM TCEP, pH 6.8, and data were collected at 20 or 25 °C. All samples contained 10% D₂O as the lock nucleus and 0.5% DSS for internal referencing.

Backbone resonance assignments were determined from band selective excitation short transient (BEST) ¹H-¹⁵N HSQC experiments and triple resonance (3D) HNCACB, HNCOCACB, HNCO, and HNCACO experiments (TROSY [169] or non-TROSY versions). To reduce collection times, all 3D experiments used non-uniform sampling (NUS). NUS reconstructions were performed using the Sparse Multidimensional Iterative Lineshape-Enhanced (SMILE) approach[204] in NMRPipe[170]. Spectra were visualized in Sparky[171] or NMRView[205].

Residue-specific secondary structure propensities were calculated using Talos-N[206] and from the deviations of the experimental CA and CB chemical shifts relative to the random coil values of Poulsen et al. (https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/)[173-175, 207]. Three or

more sequential negative values below a pre-determined average chemical shift value and four or more sequential positive values above a pre-determined average chemical shift value suggest propensity to form strand and helical secondary structure, respectively.

Heteronuclear NOE (HetNOE), longitudinal (R_1), and transverse (R_2) spin relaxation rates were measured using TROSY[176] or non-TROSY based interleaved pulse sequences[208]. HetNOE experiments were collected with and without proton saturation with a relaxation delay of 8 s. NOE values and errors were calculated in NMRView[205] as previously reported[84]. R_1 spin relaxation rates were determined using relaxation delays of 0.02, 0.06 (x3), 0.1, 0.2, 0.4, 0.6, 0.8, and 1.2 s. R_2 spin relaxation rates were measured using relaxation delays of 17, 34 (x3), 51, 68, 85, 136, 170, and 237 ms, with a recycle delay of 1.5 s. R_1 and R_2 data were fit using the rate analysis tool in NMRView as previously described[4]. Experimental errors were calculated from triplicate experiments (x3). CLEANEX NMR experiments were collected with a mixing time of 100 ms using a recycle delay of 1.5 s.

Chemical shift perturbations (CSPs) of the isolated and bivalent WW domains relative to WW1-4 were calculated using the following equation,

$$\Delta\delta \text{ (ppm)} = \sqrt{(\Delta\delta_{HN})^2 + (\Delta\delta_N \times 0.14)^2}$$

where HN and N are the ^1H and ^{15}N chemical shifts, respectively.

Chapter 5

Biophysical characterization of the AMOTL1-NEDD4 complex

Amber Vogel, Matthew McWhorter, and Afua Nyarko

Abstract

Angiomotin-Like 1 (AMOTL1) is a multivalent scaffolding protein involved in cell growth and morphogenesis, and improper regulation of its cellular concentrations is implicated in cancer. Turnover of AMOTL1 is controlled by direct PPXY motif-WW domain interactions and ubiquitination by the multivalent E3 ubiquitin ligase, NEDD4 (neural precursor cell-expressed developmentally downregulated 4). Uncovering the mechanism of complex assembly will reveal how AMOTL1 is controlled on the molecular level. While specific WW-PPXY pairings are implicated in complex assembly of closely related NEDD4 and AMOTL1 family members, binding of tetravalent NEDD4 and trivalent AMOTL1 in context of all putative binding sites has never been biophysically characterized. In this study, we produced the full binding domains of AMOTL1 and NEDD4 and used complementary molecular biophysics approaches including multi-angle light scattering, solution nuclear magnetic resonance (NMR), and isothermal titration calorimetry to investigate the mechanism of AMOTL1-NEDD4 complex assembly. Our findings reveal that binding of NEDD4 to AMOTL1 occurs through stable contacts between WW3 and PPXY1, and dynamic interactions of WW2 and WW4 with PPXY2 and PPXY3. This work highlights the importance of multivalency in assembly of NEDD4 complexes and begins to uncover finetuned mechanisms that regulate the stability of AMOTL1.

Introduction

Angiomotin-Like 1 (AMOTL1) is a 107 kDa tight junction (TJ)-associated, multidomain scaffolding protein from the Motin family. It is involved in numerous morphogenic functions including cell polarity, cell-cell contacts, migration, and regulation of cell growth[146]. AMOTL1 and its paralogs Angiomotin (AMOT) and Angiomotin-Like 2 (AMOTL2) are implicated in both tumor suppression and oncogenesis in a context-dependent manner[119, 146]. Regulating the cellular concentrations of these proteins is an important checkpoint in cell growth and proliferation, and in some cases, degradation of Angiomotin proteins (AMOTs) results in cancer[2]. Thus, understanding the mechanisms by which AMOTs are regulated on the molecular level is an important step in the pursuit of cancer therapeutics.

AMOT protein turnover is controlled by several E3 ubiquitin ligases including neural precursor cell-expressed developmentally downregulated 4-1 (NEDD4-1). NEDD4-1 is the founding member of the NEDD4 family of E3 ubiquitin ligases. These proteins all contain an N-terminal C2 lipid/Ca²⁺ binding domain, 2-4 WW domains, and a C-terminal enzymatic HECT domain responsible for ubiquitination of substrates[129]. NEDD4-1 has four WW domains, and of the > 40 identified substrates of NEDD4-1, most bind one or more of its WW domains through cognate L/PPXY motifs (hereafter referred to as PPXY), of which AMOTL1 has three (P1, P2, and P3). The interaction of AMOTL1 with NEDD4 causes dislodgement from TJs, polyubiquitination, and 26S proteasomal degradation of AMOTL1, resulting in enhanced cell growth[126, 141, 209]. However, questions remain as to how the AMOTL1-NEDD4-1 complex assembles.

Most binding studies show that NEDD4 WW3 is the primary binding site for substrates[88, 192-194]. However, several recent pieces of evidence suggest that the other domains play a role in substrate recognition. First, we recently conducted a study showing that adjacent apo WW domains modulate the chemical environments of their neighbors (Chapter 4). Second, work by Rheinmann and coworkers showed that binding of NEDD4L (a closely related NEDD4-1 paralog) to AMOT is mediated predominantly by the first PPXY motif of AMOT and the third WW domain of NEDD4L, having an affinity of 4.5 μ M[163]. However, other PPXY motifs and WW

domains add strength to the interaction, as binding of constructs containing all cognate binding sites produced an affinity of 2.7 μM , and abolishing binding of the third WW domain alone did not disrupt complex-mediated cellular function.

Although the preferred binding sites are known for NEDD4L and AMOT, such information is absent for the NEDD4-1-AMOTL1 interaction. The WW domain regions of NEDD4-1 and NEDD4L and the PPXY motif regions of AMOTL1 and AMOT are fairly similar (55.6% and 57.3%, respectively). However, subtle differences in cellular function suggest that these closely related paralogs may behave differently. For example, AMOT functions primarily in regulating cell shape, whereas AMOTL1 controls cell polarity and cell-cell junctions[119]. Furthermore, the expression levels of the AMOT and NEDD4 isoforms are different in various disease states[119, 134]. While multivalency appears to play a role in the NEDD4L-AMOT interaction, the specific contributions of each domain/motif in context of their neighbors remain unknown.

In this study, we generate constructs of NEDD4-1 containing all combinations of adjacent WW domains and use conjointly reinforcing biophysical techniques to dissect the role of multivalency in NEDD4-1-AMOTL1 complex assembly. Our findings reveal that the NEDD4-1 WW domains bind AMOTL1 with cooperativity, whereby WW3 and P1 anchor binding but require engagement of WW2 and WW4 with additional PPXY sites to achieve full binding affinity. This work unravels some of the finetuned complexity of WW domain-PPXY motif assemblies in cell signaling and protein homeostasis and highlights the importance of studying multivalent WW domain-PPXY motif interactions in context of all potential recognition sites.

Results

Construct design. NEDD4-1 and AMOTL1 are multivalent proteins which contain four WW domains and three PPXY motifs, respectively. To study binding in context of the complete WW domain region of NEDD4-1, a construct was designed to include all four WW domains, spanning residues 600-935 (hereafter referred to as WW1-4). To study the contributions of each domain to binding, multiple permutations of the WW domains were designed: WW1 (residues 600-647), WW2 (residues 645-

809), WW3 (residues 807-881), and WW4 (residues 879-935), WW1-2 (residues 600-809), WW2-3 (residues 645-881), and WW3-4 (residues 807-935), WW1-3 (residues 600-881), WW2-4 (residues 645-935), and P871A (residues 600-935 with a proline to alanine mutation introduced at residue 871 to abolish binding of WW3). All NEDD4-1 constructs contained additional non-native N- and/or C-terminal residues from the expression vector (Table 5.1).

Table 5.1: Tags of NEDD4-1 constructs. *Alternate design of WW1-2 described in Appendix 1.

Construct	Residues	N-terminal tag (non-native residues)	C-terminal tag (non-native residues)
WW1-4	600-935	GAHM-	---
P871A	600-935	GAHM-	---
WW1	600-647	SG-	-PETGGHHHHHH
WW2	647-809	GGG-	-AETGGHHHHHH
WW3	807-881	GGG-	-LPETGGHHHHHH
WW4	879-935	GGG-	---
WW1-2*	600-809	SG-	-PETGGHHHHHH
WW2-3	647-881	GGG-	-LPETGGHHHHHH
WW3-4	807-935	GGG-	---
WW1-3	600-881	SG-	-SLPETGGHHHHHH
WW2-4	645-935	GGG-	---

Constructs of AMOTL1 were designed to include the complete PPXY region spanning residues 178-384 (designated A₁₂₃). Three PPXY mutant constructs were designed to study the contribution of each motif to binding. These constructs contain Tyr to Ala (YΔA) mutations within the first (P1), second (P2), or third (P3) PPXY motif: Y191A (designated A_{Δ1}), Y313A (designated A_{Δ2}), and Y370A (designated A_{Δ1}). All NEDD4-1 and AMOTL1 constructs are depicted in Figure 5.1.

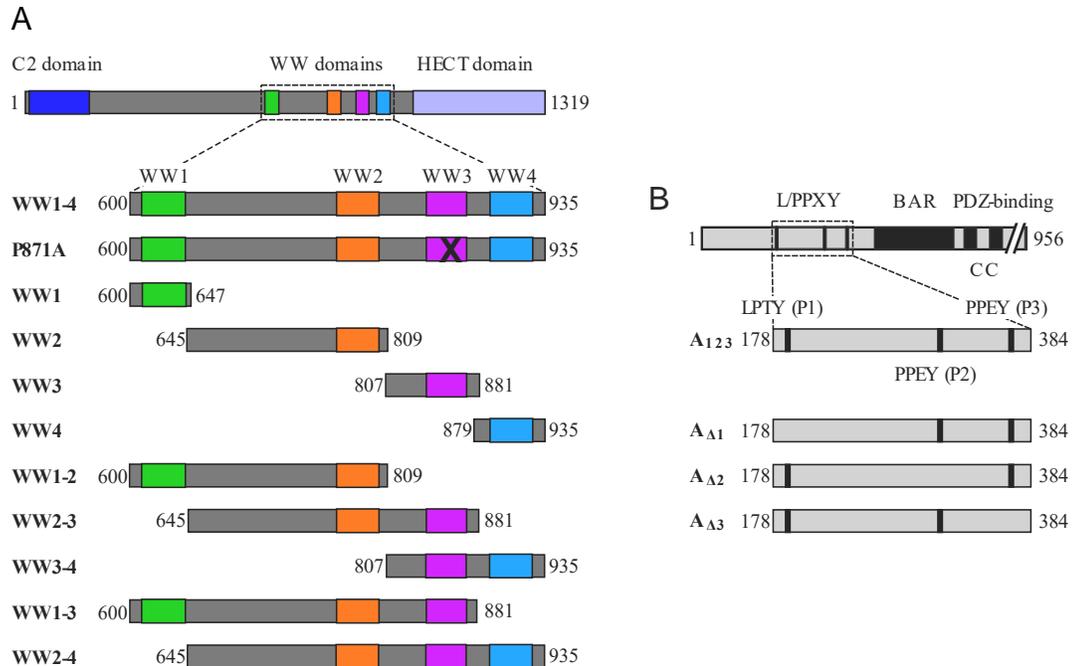


Figure 5.1: Construct Design of NEDD4-1 and AMOTL1 with all putative binding domains. (A) Domain architecture of full-length (FL)-NEDD4 (residues 1-1319) containing an N-terminal C2-domain (dark blue), four WW domains (WW1, green; WW2, orange; WW3, purple; and WW4, blue), and a C-terminal HECT domain (blue-grey). Below, enlarged design for 11 WW domain constructs of NEDD4-1: WW1-4 containing all four WW domains (residues 600-935); P871A (residues 600-935), with a Pro to Ala mutation at position 871 to inactivate the WW3 domain; WW1 (residues 600-647); WW2 (residues 645-809); WW3 (residues 807-881); WW4 (residues 879-935); WW1-2 (residues 600-809); WW2-3 (residues 645-881); WW3-4 (residues 807-935); WW1-3 (residues 600-881); and WW2-4 (residues 645-935). (B) Domain architecture of FL-AMOTL1 (residues 1-956) containing three N-terminal PPXY motifs, a Bin/Amphiphysin/Rvs (BAR) domain, coiled-coil (CC) domains, and a C-terminal postsynaptic density protein (PDZ)-binding domain. Below, enlarged design for four PPXY motif constructs of AMOTL1 (residues 178-384): A₁₂₃ containing all three PPXY motifs; and three mutant constructs, A_{Δ1}, A_{Δ2}, A_{Δ3}, with Tyr to Ala mutations within the first (Y191A) second (Y313A), and third (Y370A) PPXY motifs, respectively.

WW1-4 and A₁₂₃ form a 1:1 complex with sub-micromolar affinity. Due to the presence of four WW domains and three PPXY motifs in NEDD4-1 and AMOTL1, respectively, a range of binding configurations are possible. To determine the overall binding stoichiometry and energetics in context of all domains and motifs, we performed SEC-MALS and ITC with WW1-4 and A₁₂₃ (Fig. 5.2). For SEC-MALS

experiments, A_{123} and WW1-4 were combined at ratios of 1:1 and 1:3, which yielded molar masses of 55.9 ± 0.6 and 54.8 ± 1.0 , respectively (Fig. 5.2A, Table 5.2). These values are below the theoretical molar mass of 64.2 kDa for a 1:1 complex, but higher than apo WW1-4 or A_{123} which have empirical molar mass values of 37.7 and 26.1 kDa, respectively[4]. Furthermore, addition of excess WW1-4 to A_{123} did not result in an increase in observed molar mass or change in elution volume, indicating that the complex formed contains the maximum amount of bound WW1-4. Since SEC-MALS is a macroscopic technique that measures the average molar mass, one possible explanation for an observed molar mass 10 kDa below the empirical mass is that the complex is highly dynamic and elutes as an equilibrium of bound and unbound species.

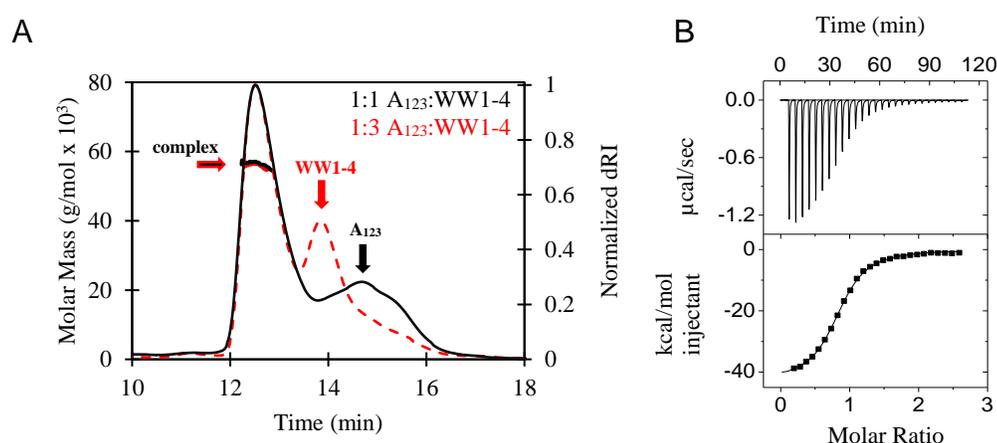


Figure 5.2: SEC-MALS and ITC show WW1-4 and A_{123} form a 1:1 complex. (A) SEC-MALS elution profiles of A_{123} mixed with WW1-4 at molar ratios of 1:1 (black, solid line) and 1:3 (red, dashed line). (B) Representative isotherm of WW1-4 titrated into A_{123} collected at 25 °C in buffer composed of 50 mM sodium phosphate, 50 mM sodium chloride, 5 mM β -mercaptoethanol, and 0.5 mM sodium azide at pH 7.5.

Table 5.2: SEC-MALS of WW1-4 bound to A_{123} .

A_{123} :WW1-4	Theoretical MW, kDa	Experimental MW, kDa
1:1	64.2	55.9 ± 0.6
1:3	140.8	54.8 ± 1.0

To quantify the thermodynamics of binding, WW1-4 was titrated into A_{123} and monitored by ITC (Fig. 5.2B, Table 5.3). The binding interaction yielded a K_d of $0.4 \pm 0.1 \mu\text{M}$ with a favorable enthalpic term of $-42.4 \pm 0.5 \text{ kcal/mol}$ counterbalanced by an

unfavorable entropic term of -33.7 ± 0.6 kcal/mol. The interaction produced a stoichiometry of 0.9 ± 0.1 which agrees with the conclusion that the complex in the SEC-MALS analysis is formed by one monomer of WW1-4 bound to one monomer of A₁₂₃. Altogether, these data show that this interaction occurs between one molecule each of WW1-4 and A₁₂₃ with sub-micromolar binding affinity.

Table 5.3: Thermodynamic parameters for A₁₂₃ bound to varying valency NEDD4-1 WW domains. * N set to 1 in data analysis fitting. ^a Average of duplicate runs. No binding: n.b.

Syringe	Cell	N	K _d (μ M)	Δ H (kcal/mol)	T Δ S (kcal/mol)	Δ G (kcal/mol)
WW1-4	A ₁₂₃	0.9 ± 0.1	0.4 ± 0.1	-42.4 ± 0.5	-33.7 ± 0.6	-8.7 ± 0.1
WW1	A ₁₂₃	1*	n.b.			
WW2	A ₁₂₃	1*	$> 60^a$	<i>weak binding</i>		
WW3	A ₁₂₃	1.16 ± 0.02	5.0 ± 0.8	-35 ± 4	-27 ± 3	-7.2 ± 0.1
WW4	A ₁₂₃	1*	$> 60^a$	<i>weak binding</i>		
WW1-2	A ₁₂₃	1*	$> 40^a$	<i>weak binding</i>		
WW2-3	A ₁₂₃	1.04 ± 0.01	0.57 ± 0.03	-34.8 ± 0.1	-26.3 ± 0.2	-8.52 ± 0.03
WW3-4	A ₁₂₃	1.16 ± 0.03	1.1 ± 0.1	-35 ± 1	-27 ± 1	-8.13 ± 0.03
WW1-3	A ₁₂₃	1.1 ± 0.1	0.4 ± 0.1	-33 ± 1	-25 ± 1	-8.8 ± 0.1
WW2-4	A ₁₂₃	1.00 ± 0.02	0.35 ± 0.02	-40.3 ± 0.5	-31.5 ± 0.5	-8.80 ± 0.03
P871A	A ₁₂₃	0.9 ± 0.2	1.7 ± 0.4^a	-41 ± 4	-33 ± 4	-7.9 ± 0.1

NEDD4 WW domains bind AMOTL1 cooperatively. To determine the level of cooperativity between the WW domains of NEDD4-1 in binding AMOTL1, we performed a series of ITC experiments in which varying combinations and valency of WW domains were titrated into A₁₂₃ (Fig. 5.3). First, we titrated each isolated WW domain into A₁₂₃. Of the four WW domains, only WW3 bound A₁₂₃ with affinity in the range quantifiable by ITC (Fig. 5.3C). The WW3-A₁₂₃ interaction yielded a stoichiometry of 1.16 ± 0.02 , indicating that WW3 predominantly binds one PPXY motif of A₁₂₃, but weak interactions with additional motifs likely account for the non-integer stoichiometry. The interaction resulted in a K_d of $5.0 \pm 0.8 \mu$ M, ~10-fold weaker than WW1-4, indicating that the other domains play a critical role in achieving full binding affinity. Binding of WW2 and WW4 to A₁₂₃ was too weak to accurately compute, having values $\geq 60 \mu$ M (Fig. 5.3B, D). Titration of WW1 into A₁₂₃ produced

heats indistinguishable from the baseline, indicating the absence of an interaction (Fig. 5.3A).

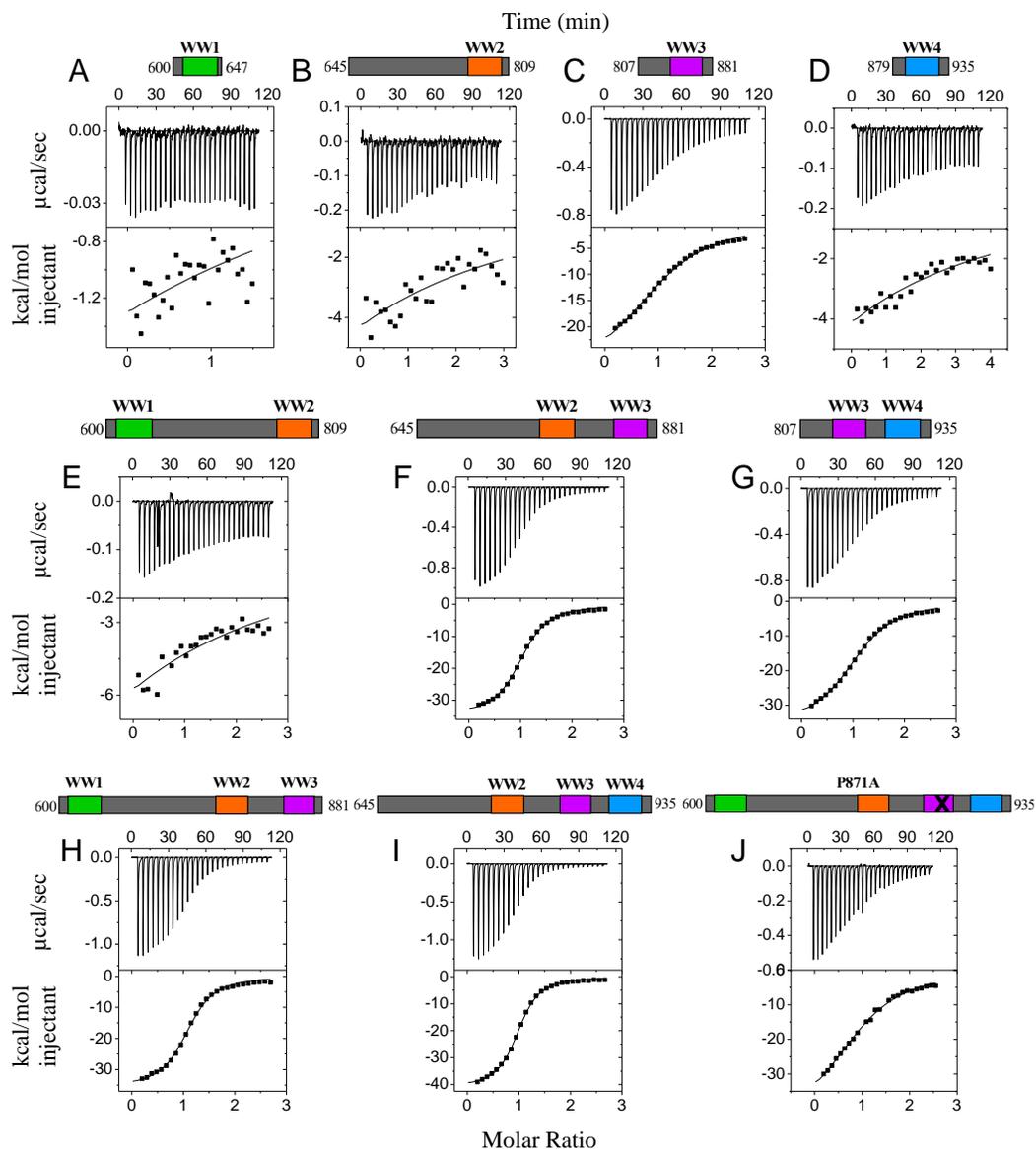


Figure 5.3: ITC of mono-, bi-, and trivalent NEDD4-1 WW domain constructs titrated into A₁₂₃. Representative isotherms of (A) WW1, (B) WW2, (C) WW3, (D) WW4, (E) WW1-2, (F) WW2-3, (G) WW3-4, (H) WW1-3, (I) WW2-4, and (J) P871A titrated into A₁₂₃. Data were collected at 25 °C in in buffer composed of 50 mM sodium phosphate, 50 mM sodium chloride, 5 mM β -mercaptoethanol, and 0.5 mM sodium azide at pH 7.5.

To determine the number and combination of domains needed to achieve full binding affinity, NEDD4-1 constructs containing sequential bivalent WW domains

were titrated into A₁₂₃ and monitored by ITC. Binding of WW1-2 and A₁₂₃ was too weak to accurately fit (Fig. 5.3E), indicating that WW1 does not improve the ability of WW2 to bind. Titration of WW2-3 into A₁₂₃ resulted in a K_d of $0.57 \pm 0.03 \mu\text{M}$, roughly equivalent to the WW1-4 binding affinity (Fig. 5.3F). Titration of WW3-4 into A₁₂₃ resulted in a K_d of $1.1 \pm 0.1 \mu\text{M}$, roughly 2.5-fold weaker than WW1-4 affinity (Fig. 5.3G).

Next, we titrated trivalent NEDD4-1 constructs into A₁₂₃. Binding of WW1-3 to A₁₂₃ produced a K_d of $0.4 \pm 0.1 \mu\text{M}$ (Fig. 5.3H). Relative to binding of WW2-3, there was almost no change in the enthalpy or entropy terms, indicating that WW1 is thermodynamically inert in the assembly process. Binding of WW2-4 to A₁₂₃ resulted in a K_d of $0.35 \pm 0.02 \mu\text{M}$ (Fig. 5.3I). Compared to WW2-3, this binding interaction was 5.5 kcal/mol more enthalpically favorable and 5.2 kcal/mol more entropically unfavorable. This indicates that WW4 contributes to binding through hydrogen bonds/Van der Waals interactions which are counterbalanced by unfavorable losses in entropy. To see if WW1, 2, and 4 could rescue binding in the absence of WW3, we generated a WW1-4 mutant, P871A, which abrogates binding of WW3, and performed an ITC titration into A₁₂₃ (Fig. 5.3J). Intriguingly, the P871A-A₁₂₃ interaction yielded a K_d of $1.7 \mu\text{M}$, reflecting a ~3.8-fold decrease in binding affinity relative to WT, but ≥ 35 -fold stronger than the isolated WW2 or WW4 affinity for A₁₂₃. This data indicates that while WW1, 2, and 4 cannot fully compensate for the loss of WW3, these domains do exhibit strong binding cooperativity. All bivalent and trivalent WW domain-A₁₂₃ interactions yielded stoichiometries close to one.

Altogether, these data show that (i) WW3 is the strongest binder and interacts predominantly with one PPXY motif of A₁₂₃, thus serving as a stable binding initiation site; (ii) both WW2 and WW4 cooperate with WW3 in binding to A₁₂₃; (iii) WW2 and WW4 bind A₁₂₃ cooperatively but cannot compensate for the loss of WW3; and (iv) the WW1 domain contributes no significant energetics to the NEDD4-1-AMOTL1 interaction.

All three AMOTL1 PPXY motifs participate in binding NEDD4-1. To identify which PPXY motifs are involved in binding NEDD4-1, we performed a series of ITC and NMR titration experiments (Fig. 5.4). First, we titrated WW1-4 into A₁₂₃ constructs

with each motif abrogated and monitored binding by ITC. Titration of WW1-4 into A $_{\Delta 1}$ (P2 and P3 active), A $_{\Delta 2}$ (P1 and P3 active), and A $_{\Delta 3}$ (P1 and P2 active) yielded stoichiometries close to 1 and K $_d$ values of $3.8 \pm 0.8 \mu\text{M}$, $1.2 \pm 0.01 \mu\text{M}$, and $0.5 \pm 0.04 \mu\text{M}$, respectively (Fig. 5.4A-C, Table 5.4). The favorable enthalpy contributions and unfavorable entropy contributions were reduced in binding of the A $_{\Delta 1}$ and A $_{\Delta 2}$ mutants. These data indicate that P1 and P2 contribute most of the binding energetics. Although the contributions of P3 were negligible, ITC measures global binding energetics, so minor interactions could be missed in the context of the more dominant interactions at P1 and P2.

To see if P3 interacts with WW1-4 in concert with P1 and P2, we performed an NMR titration of unlabeled WW1-4 into ^{15}N -labeled A $_{123}$ and collected ^1H - ^{15}N HSQC spectra at each titration point. Binding was evidenced by peak intensity changes. Figure 5.4D shows the normalized peak intensity of A $_{123}$ plotted versus residue number, with PPXY residues shaded in grey. Upon addition of 0.4 equivalents (eq.) WW1-4, peaks in the sequence vicinity of all three PPXY motifs disappeared. Upon addition of 0.7 eq. WW1-4, peaks remaining were predominantly localized to the P1-P2 linker, except for 3 residues each at the N-terminus and within the P2-P3 linker. Altogether, these data indicate that P1 is the tightest binder, but all three motifs interact with WW1-4.

Table 5.4: Thermodynamic parameters for the WW1-4 – A $_{123}$ /mutant complexes.
*Data from Table 5.3 for comparison.

Syringe	Cell	N	K $_d$ (μM)	ΔH (kcal/mol)	T ΔS (kcal/mol)	ΔG (kcal/mol)
WW1-4	A $_{123}$ *	0.9 ± 0.1	0.4 ± 0.1	-42.4 ± 0.5	-33.7 ± 0.6	-8.7 ± 0.1
WW1-4	A $_{\Delta 1}$	0.7 ± 0.1	3.8 ± 0.8	-29 ± 3	-22 ± 3	-7.4 ± 0.1
WW1-4	A $_{\Delta 2}$	0.7 ± 0.03	1.2 ± 0.01	-33 ± 1	-24 ± 1	-8.07 ± 0.01
WW1-4	A $_{\Delta 3}$	0.7 ± 0.1	0.5 ± 0.04	-39 ± 1	-30.8 ± 0.6	-8.63 ± 0.05

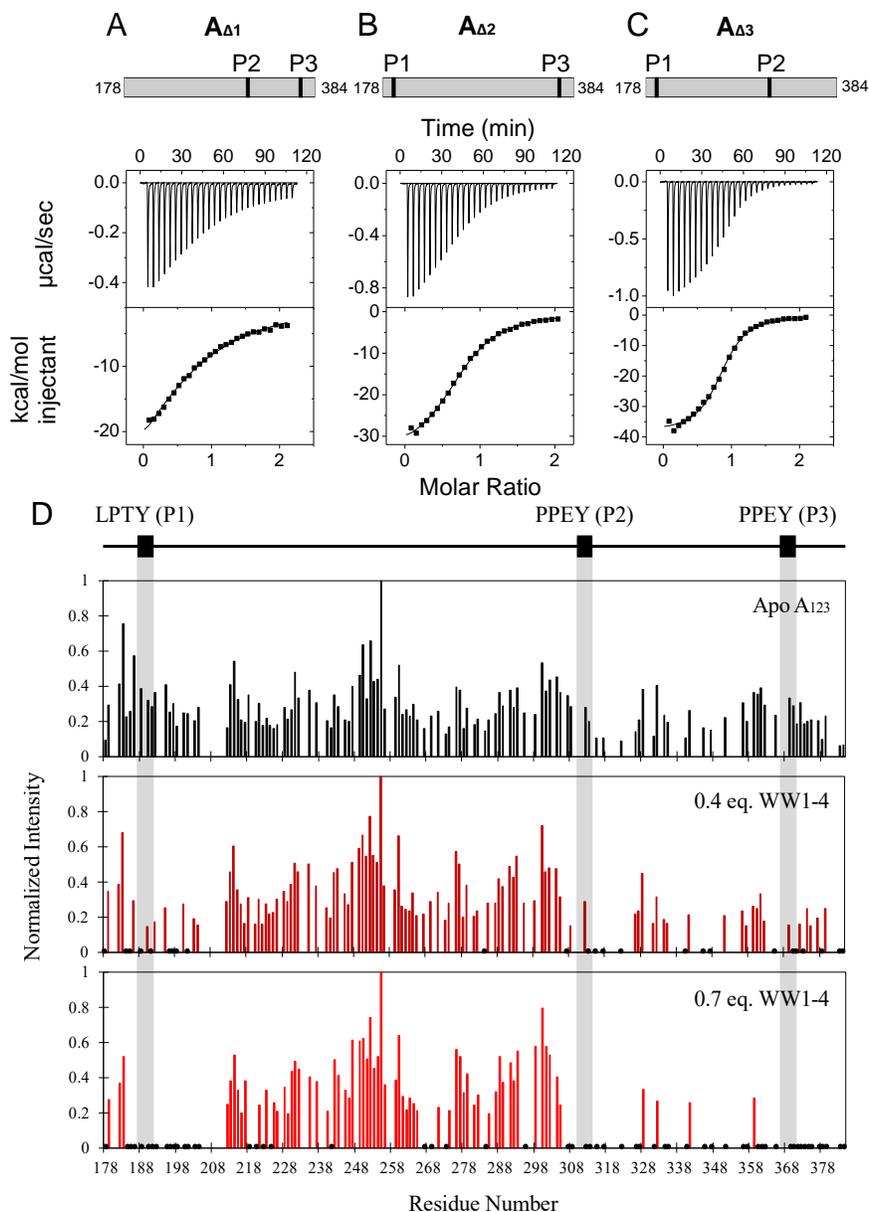


Figure 5.4: the WW1-4 binding interface of A₁₂₃ mapped by ITC and NMR. Representative isotherms of WW1-4 titrated into (A) A_{Δ1}, (B) A_{Δ2}, and (C) A_{Δ3} collected at 25 °C in buffer containing 50 mM sodium phosphate, 50 mM sodium chloride, 5 mM β-mercaptoethanol, and 0.5 mM sodium azide at pH 7.5. (D) Normalized peak intensity of ¹⁵N-labeled A₁₂₃ upon addition of 0 eq. (top panel, black), 0.4 eq. (middle panel, maroon), and 0.7 eq (bottom panel, red) unlabeled WW1-4 plotted versus residue number. Peak intensities are normalized to the most intense A₁₂₃ peak: Q255. Residues which disappear at each titration point are indicated with black dots along the X-axes. PPEY motif regions are shaded in grey. NMR data were collected at 10 °C.

Discussion

Regulation of AMOTL1 cellular concentrations is an important step in the control of cell growth. Understanding this regulatory process will provide insight into how cell growth regulation is maintained and how it could be manipulated. Degradation of AMOTL1 is mediated through direct interactions with NEDD4-1, but how the complex assembles in context of all WW domains and PPXY sites has remained unclear. In this study, we produced a construct of NEDD4-1 containing all WW domains and characterized its interaction with the trivalent PPXY motifs of AMOTL1. Our findings reveal that, while the WW3 domain and P1 motif are the primary binding sites, weak interactions of adjacent domains and motifs are required to achieve full binding affinity as discussed below.

Interdomain communication facilitates multivalent interactions with AMOTL1.

The presence of multiple WW domains and PPXY motifs in binding partners can result in numerous binding mechanisms[67]. Recent findings show that the WW domains of NEDD4-1 exhibit intramolecular communication (Chapter 4). Here, we show that WW domain interconnectivity plays an important role in complex assembly with AMOTL1. Figure 5.5 depicts a proposed mechanism of assembly in context of the complete binding domains of NEDD4-1 and AMOTL1. ITC and NMR data show that NEDD4-1 WW3 and AMOTL1 P1 form the tightest complex, acting as an assembly anchor. However, dynamic and cooperative interactions of WW2 and WW4 with P2 and P3 are required to achieve full binding affinity.

Several structural features help explain the enhanced binding ability of the WW3 domain. Recent work by Rheinmann et al. found that NEDD4L WW3 (WW3L) prefers P1 of AMOT due to favorable contacts with a helix adjacent to P1[163]. This helix is also present in AMOTL1 and is likely important for complex stability with NEDD4-1 WW3[4]. Additionally, NMR characterization of WW3 described in Chapter 4 shows that loop 1 and strand 2 are more dynamic and solvent exposed than the corresponding regions of WW1 or WW4. In binding of WW3L to AMOT P1, loops 1 and strand 2 undergo conformational changes necessary to prevent steric clash with

the LPTY motif[163]. The enhanced dynamics of WW3 likely aid in binding to AMOTL1 through conformational selection analogous to binding of WW3L to AMOT.

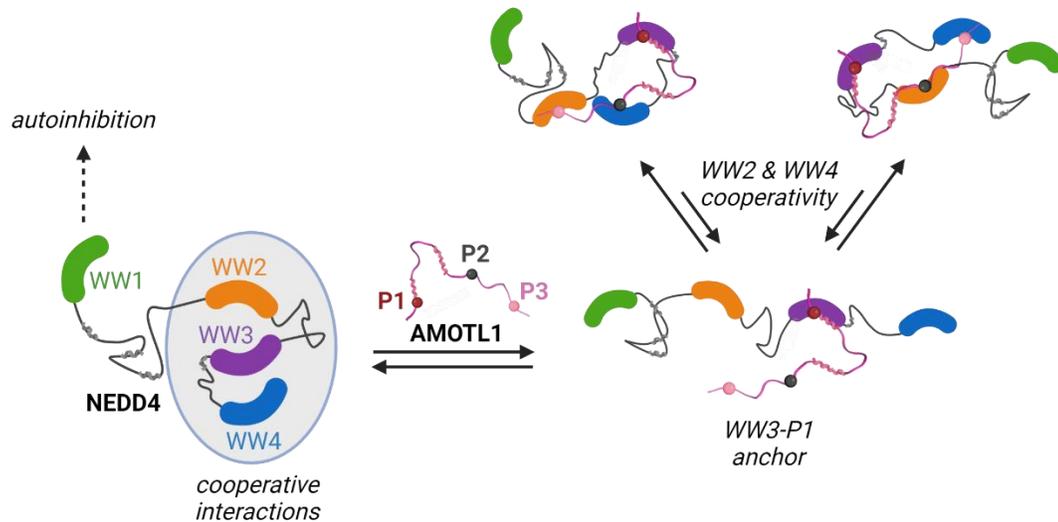


Figure 5.5: proposed model of NEDD4-1-AMOTL1 assembly. In the apo form, WW2, WW3, and WW4 exhibit intramolecular communication, while WW1 participates in autoinhibition. Binding to AMOTL1 is anchored by discrete interactions between WW3 and P1. However, full binding affinity requires dynamic and cooperative interactions of WW2 and WW4 with P2 and P3. The AMOTL1 cartoon was adapted from the AMOT polypeptide crystal structure (PDB: 6JJX). Created with Biorender.Com.

Although WW3 is the tightest binder, the affinity of WW3 by itself is ~10-fold weaker than that of the full binding segment of NEDD4-1. After WW3 and P1 anchor the complex, WW2 and WW4 form dynamic interactions with the second and third PPXY motifs, increasing the overall affinity by 10-fold. None of our data suggest that WW1 participates in binding. In a biological context, this could leave WW1 free to participate in autoinhibition even in the AMOTL1-bound state. Additionally, our data indicates that multivalent interactions of WW2 and WW4 occur even in the absence of WW3. This compensatory effect may allow NEDD4-1 to bind AMOTL1 while WW3 is occupied with another PPXY partner. This arrangement could allow NEDD4-1 to regulate multiple substrates simultaneously, allowing efficient cellular responsiveness.

Conclusions and impact. Transient, multivalent interactions are critical in regulating cell function. This work illustrates the importance of studying multivalent interactions in context of the complete set of binding-competent domains. Our findings

suggest a possible role for the multiple WW domains of the NEDD4-1 E3 ubiquitin ligase. While it was previously known that WW1 participates in autoinhibition, this study reveals that WW2-4 interact with the three PPXY motifs of AMOTL1, forming a dynamic three domain:three motif complex. Assembly occurs through formation of stable contacts between P1 and WW3, and dynamic interactions at neighboring sites. Our work illustrates the robust nature of multivalent WW domain-PPXY motif complex assembly and suggests functional importance for all four WW domains of NEDD4-1. Dynamic assembly may help NEDD4-1 sense the concentrations of AMOTL1 in the cell, leading to modulated ubiquitination levels in an AMOTL1 concentration-dependent manner. Further work is required to establish the significance of these findings in a cellular context and to determine the role of WW domain multivalency in assembly with other substrates.

Materials and Methods

Cloning of constructs and recombinant protein production are described in Appendix 1.

Multi-Angle Light Scattering. Absolute molecular weights were obtained by size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) using an AKTA FPLC (GE Healthcare), a DAWN multiple-angle light scattering, and Optilab refractive index system (Wyatt Technology). Data were collected at room temperature, on a Superdex²⁰⁰ 10/300 GL column (Cytiva life sciences) pre-equilibrated with 50 mM sodium phosphate, 50-400 mM NaCl, 1 mM NaN₃, and 5 mM β -mercaptoethanol at pH 7.5 buffer. Protein samples at concentrations of 25-75 μ M were injected at a flow rate of 0.7-0.8 mL/min. The data were analyzed with ASTRA software package, version 8 (Wyatt Technology).

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) data were recorded on a VP-ITC instrument (Malvern instruments Inc, MA) set to 25 °C. Protein samples were dialyzed against buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN₃, and 5 mM β -mercaptoethanol at pH 7.5. Twenty-seven or twenty-eight injections of 60-160 μ M WW domain proteins were titrated into 6-12 μ M A₁₂₃ or mutant proteins. Isotherms were analyzed by single-site fits of the

thermograms using the *Origin 7.0* software. The free energy of binding (ΔG) was calculated from the equation $\Delta G = -RT\ln(K_a)$, where R is the universal gas constant, T is temperature in Kelvin, and K_a is the association constant. Averages and standard deviations were calculated from triplicate titrations unless otherwise stated.

Nuclear Magnetic Resonance data collection and analysis. NMR experiments were performed on a Bruker Avance III, 800 MHz spectrometer (Bruker BioSpin) equipped with a triple resonance cryogenic probe. Isotopically labeled A_{123} at 60 μM was dialyzed against buffer composed of 50 mM sodium phosphate, 100 mM NaCl, 50 mM arginine, 50 mM glutamate, 1 mM NaN_3 , and 2 mM TCEP, pH 6.8, and data were collected at 10 °C. All samples contained 10% D_2O as the lock nucleus and 0.5% DSS for internal referencing. For NMR titration experiments, unlabeled WW1-4 was mixed with ^{15}N -labeled A_{123} at molar ratios of 0.1:1, 0.3:1, 0.4:1, 0.5:1, 0.7:1, and excess:1. To account for minor deviations in concentration, A_{123} peak intensities were normalized to the height of the most intense peak (Q255) within the spectrum of each titration point.

Chapter 6

Conclusions and Future Directions

Highlights of reported work

In the work presented in this dissertation, we successfully produce and provide the first residue-level biophysical characterization of the complete PPXY region of the multivalent scaffolding cell growth regulator – AMOTL1. This foundational study allowed us to unravel nuanced detail in the binding mechanisms of AMOTL1 and three multivalent partners: YAP, KIBRA, and NEDD4-1. Our studies reveal that all three PPXY motifs of AMOTL1 are active in complex assembly but serve different roles depending on the binding partner. In Chapter 2, we resolve discrepancies reported in the literature as to which PPXY motifs are important in complex assembly with YAP[124, 147, 149]. Our work shows that all three sites are important for binding, which results in a dynamic ensemble of complexes. The stability of this dynamic complex is modulated both by favorable and destabilizing interactions.

In Chapter 3, we expand on previously reported studies of the KIBRA-AMOT interactions which showed high affinity binding between KIBRA and a single motif peptide of AMOT[3]. Our work highlights the importance of multiple domains and motifs in complex formation. Notably, we reveal the importance of the weaker binding sites in assembly. Negative contributions of multivalent PPXY motifs results in attenuation of stability which may be important for complex reversibility and responsiveness to environmental cues.

Our work in Chapters 4 and 5 provides new insight into the purpose of WW domain multivalency for NEDD4-1 structure and binding. Biophysical studies characterizing NEDD4-substrate interactions are typically conducted with isolated domains and show WW3 to be the tightest binder[80, 88, 163, 193, 194, 210, 211]. However, the role of multiple domains is not addressed in isolated domain studies. Here, we use NEDD4-1 constructs of varying valency to show that multivalency can modulate WW domain chemical environment. Notably, NEDD4-1 WW3 structure/dynamics are modulated by its adjacent domains. This phenomenon translates to binding AMOTL1: neighboring domains modulate WW3 binding through affinity enhancement.

Transient protein-protein interactions allow intra- and extracellular signals to be rapidly spread throughout the cell[212]. Our findings point to multivalency as a tool to achieve and modulate the transient nature of signaling complexes.

Conclusions and Impact

Prior to this work, challenges in studying large, disordered, multivalent proteins such as AMOTL1 limited understanding of how multivalent binding sites regulate the assembly process. Therefore, most studies focused on smaller polypeptide segments. From these studies, only single domains or motifs which bind with high affinity were thought to participate in complex assembly. By using AMOTL1 and partner protein polypeptides containing all binding-competent motifs and domains, we show that these interactions do not form static, discrete complexes as previously reported. Instead, binding results in dynamic ensembles of complexes. Each motif or domain plays a unique role in the assembly process. For instance, one motif (the N-terminal LPTY motif in the AMOTL1-YAP interaction or the C-terminal PPEY motif in the AMOTL1-KIBRA interaction) or domain (The WW1 domain of YAP in the AMOTL1-YAP complex) may initiate the assembly process while neighboring domains and motifs finetune complex stability through weak/transient interactions. Alternatively, the domains contribute equally to the stability of the complex, as observed for the KIBRA WW domains in the AMOTL1-KIBRA interaction. The unique roles of motifs/domains in these assemblies could be functionally important.

Proper regulation of the cell requires sensitive, responsive, and tunable signaling processes[7]. Transient binding interactions are thought to confer the tunability and responsiveness needed for cell signaling, but it has been difficult to experimentally characterize these complexes[212]. Our work not only provides experimental evidence of transient binding interactions in signaling complexes, but the molecular level details provided opens the door to targeting weak interactions in multivalent systems to regulate signaling events. Several well-known drugs including ibuprofen, metformin, and alcohol bind multiple targets with weak affinity, which provides a more tunable cellular response and reduces side effects[212]. Development of drugs that target multiple weak Hippo signaling complexes could allow for

modulation of cell growth without completely disrupting fundamental processes mediated by WW domain-PPXY motif interactions. Finally, our findings underscore the significance of studying the assembly process of multivalent proteins in context of all potential binding sites – there is strength in numbers.

Future directions

How do additional domains regulate AMOTL1 protein complexes? Our work illustrates the importance of studying protein complexes in context of larger, more biologically relevant segments. However, the proteins studied here contain other domains in addition to the WW domains and PPXY motifs. The next logical step is to extend the studies beyond the multi-WW or PPXY segments. Ultimately, the goal is to study intact full-length proteins. The limitations of bacterial expression systems may require production of these large protein fragments in mammalian or insect cells. Experiments which may be useful in characterizing these large protein complexes include segmental isotopic labeling or methyl-labeling coupled with NMR, analytical ultracentrifugation, Förster resonance energy transfer (FRET), and cryo-electron microscopy.

How does AMOTL1 discriminate between its partners? In a biological context, WW domains and PPXY motifs exist in an environment rich in proteins with shared binding sites. Different combinations of WW domain-PPXY motif interactions result in different cellular outcomes. AMOTL1 binds numerous WW domain partners in addition to YAP, KIBRA, and NEDD4-1 including ITCH and NEDD4L[117]. These complexes result in different downstream effects[2, 141, 147]. Thus, it is important to understand how AMOTL1 discriminates between its partners and how complex stability is regulated in context of additional proteins. In Chapter 3, we show that YAP and KIBRA, which prefer alternate PPXY motifs, bind AMOTL1 simultaneously. However, it is unclear if/how NEDD4-1 competes with YAP and KIBRA for AMOTL1. In Appendix 2, we discuss preliminary findings which begin answering this question.

Kinetics of complex assembly. Multiple pieces of evidence indicate that the AMOTL1-YAP, -KIBRA, and -NEDD4-1 complexes are dynamic. However, the

kinetics of these interactions remain unknown. Future studies could include surface plasmon resonance experiments which measure k_{on} and k_{off} rates of binding. Differential binding kinetics may help explain how AMOTL1 discriminates between its partners and how each binding site contributes to complex dynamics.

How do transient interactions regulate cell growth? A primary finding of this work is that weak interaction sites contribute to complex stability through positive cooperativity or attenuation of affinity. However, whether these findings are functionally relevant in the cell remains to be determined. In principle, modulation of complex stability through multivalent interactions may serve as a mechanism for finetuning the subcellular localization of cell growth regulators such as YAP. In some multivalent systems, accumulation of proteins can result in phase transitions, where weak, multivalent interactions result in the formation of membraneless organelles or droplets[213-216]. Proteins partitioned into phase separated droplets are sequestered, which could attenuate their normal cell function. This process occurs in a protein concentration-dependent manner, which could allow retention of YAP in the cytoplasm when it accumulates at high levels. Similarly, the functions of other YAP regulators may be modulated through phase separation. In Appendix 5, we present preliminary findings which suggest that AMOTL1 and NEDD4-1 undergo phase separation.

In addition to phase separation, affinity modulation through multivalency may simply serve to balance the equilibrium between free and bound proteins. To test the importance of finetuned complex affinity, future experiments may include monitoring subcellular localization and activation of YAP in live cells where weak binding sites of the proteins discussed in this work are mutated. For example, mutating WW1, WW2, and WW4 of NEDD4-1, which reduces the interaction strength with AMOTL1 by 10-fold, would theoretically result in enhanced stability of AMOTL1. Downstream, increased concentrations of AMOTL1 would likely result in decreased YAP-mediated cell growth. Similarly, mutation of AMOTL1 P3, which increases the strength of the AMOTL1-YAP interaction, would likely result in the same outcome. Experiments which measure the cell growth activity of YAP include fluorescence microscopy to observe the subcellular localization of YAP, mRNA expression analysis of downstream cell growth activators, and quantification of cell size and number.

What factors drive motif specificity? In Chapters 2, 3, and 5, we determine the AMOTL1 PPXY motif preferences of YAP, KIBRA, and NEDD4-1. However, it remains unclear what molecular features govern motif specificity. Potential factors include properties of flanking residues, binding site structure, linker length, and distal allosteric changes. Preliminary data and potential future experiments to identify features which confer specificity are discussed in Appendix 6.

How do the multiple NEDD4-1 WW domains modulate each other? NMR experiments of isolated, bivalent, and tetravalent NEDD4-1 WW domains show that adjacent domains modulate the chemical environments of their neighbors. Altered chemical environment could result from direct domain-domain interactions, changes in dynamics, or altered domain structure. Future experiments to discern between these possibilities could include NMR titrations of unlabeled isolated domains into alternate ^{15}N -labeled domains, paramagnetic relaxation enhancement NMR measurements, FRET, and dynamics experiments using the multivalent WW domain constructs.

Improving A_{123} NMR data. The future of this work is in studying protein complexes of increasing size and complexity. However, using HSQC experiments to map the binding interfaces of large PPXY motif proteins has limitations. First, proline residues are undetectable in HSQC experiments due to the absence of an amide proton. Second, since multivalent PPXY motif regions are usually large and disordered, HSQC spectra are typically crowding with many overlapping peaks. These limitations restrict the information obtainable in NMR titration and dynamics experiments. Even in A_{123} , numerous HSQC peaks overlapped, and 10.7% of residues are proline. One method to overcome these challenges is through use of carbon-detect NMR experiments. Experiments such as CON provide a spectrum of the ^{15}N -amide nitrogen correlated to the ^{13}C of each residue[217]. These experiments are more dispersed than an HSQC, making them useful for large, disordered proteins. Another strategy is selective unlabeled, where unlabeled amino acids are added to the cell growth media and are preferentially incorporated into the recombinant protein of interest. In A_{123} , 11.2% of residues are Glu. Incorporation of unlabeled Glu into A_{123} could therefore increase the number of resolved peaks from which intensity data can be quantified.

Appendices

Appendix 1

Cloning, expression, and purification of proteins

Amber Vogel, Matthew McWhorter, and Alexandra Crawford

Cloning of constructs

The constructs used in this work and the template DNA used for gene amplification/assembly are listed in Table A1.1. The YAP cDNA was a gift from Kunliang Guan (Addgene plasmid # 24637). DNA for NEDD4-1 construct amplification was synthesized and purchased from Genescript, Piscataway, NJ. KIBRA and AMOTL1 cDNA were gifts from Joachim Kremerskothen (University of Munster, Germany).

AMOTL1, YAP, KIBRA, and NEDD4-1 constructs WW1-4 and P871A (Table A1.1) were cloned into NdeI/XhoI cut sites of pET24TM (Novagen, EMD Biosciences, San Diego, CA) containing an N-terminal 6x-histidine (His_{6x}) tag followed by a tobacco-etch virus (TEV) protease cleavage site using the Gibson AssemblyTM cloning protocol (New England Biolabs, MA). Cloning protocols for AMOTL1, YAP, and KIBRA polypeptides described in this work are published[4, 84].

NEDD4-1 constructs WW2_{short}, WW4, WW1-2, WW3-4, and WW2-4 were cloned into BamHI/XhoI cut sites of a modified pET vector gifted by Dr. Richard Cooley, pRBC[218] using the Gibson AssemblyTM cloning protocol (New England Biolabs, MA). The pRBC vector contains an N-terminal *Brachypodium distachyon* Small Ubiquitin-like Modifier (bdSUMO) tag followed by a *Brachypodium distachyon* Sentrin-specific protease 1 (bdSEN1) cleavage site[219], and an N-terminal His_{6x} tag. These constructs were cloned with non-native N-terminal GGG- and/or C-terminal LPETG/LAETG sequences for Sortase A reactions as detailed in Table A1.1.

NEDD4-1 constructs WW1, WW2, WW3, WW1-2* WW2-3, and WW1-3 were cloned as described above into a modified pRBC vector containing an N-terminal bdSUMO tag followed by a bdSEN1 cleavage site and a C-terminal His_{6x} tag.

The genes for LPXTG and LAXTG specificity eSrtA enzymes in pET29 expression vectors were gifted by Dr. Rina Rosenzweig (Department of Chemical and Structural biology, Weizmann Institute of Science, Israel). Cloning of eSrtA enzymes is described elsewhere[203].

Table A1.1: Construct information for genes used in Chapters 2-5. WW1-2 was used in Chapter 4. WW1-2* was used in Chapter 5. Tags in parenthesis are protease cleavage sites.

Protein	Residues	Uniprot ID	Construct	Expression vector	Tag (N-term)	Tag (C-term)
AMOTL1	178-384	Q8IY63	A ₁₂₃	pET24	His6x-(TEV)	-
			A _{Δ1}	pET24	His6x-(TEV)	-
			A _{Δ2}	pET24	His6x-(TEV)	-
			A _{Δ3}	pET24	His6x-(TEV)	-
			A ₃	pET24	His6x-(TEV)	-
			A _{RYΔAA}	pET24	His6x-(TEV)	-
YAP2	157-277	P46937-2	YWW _{TD}	pET24	His6x-(TEV)	-
			P202A _{TD}	pET24	His6x-(TEV)	-
			P261A _{TD}	pET24	His6x-(TEV)	-
KIBRA	1-91	Q8IX03	KWW _{TD}	pET24	His6x-(TEV)	-
			P37A _{TD}	pET24	His6x-(TEV)	-
			P84A _{TD}	pET24	His6x-(TEV)	-
NEDD4-1	600-647	P46934	WW1	pRBC	bdSUMO-(bdSEN1)	PETGG-His6x
	645-809		WW2	pRBC	bdSUMO-(bdSEN1)-GGG	PETGG-His6x
	700-809		WW2 _{short}	pRBC	His6x-bdSUMO-(bdSEN1)-GGG	LAETG
	807-881		WW3	pRBC	bdSUMO-(bdSEN1)-GGG	LPETGG-His6x
	879-935		WW4	pRBC	His6x-bdSUMO-(bdSEN1)-GGG	-
	600-809		WW1-2	pRBC	His6x-bdSUMO-(bdSEN1)	LAETG
	600-809		WW1-2*	pRBC	bdSUMO-(bdSEN1)	LPETGG-His6x
	645-881		WW2-3	pRBC	bdSUMO-(bdSEN1)-GGG	LPETGG-His6x
	807-935		WW3-4	pRBC	His6x-bdSUMO-(bdSEN1)-GGG	-
	600-881		WW1-3	pRBC	bdSUMO-(bdSEN1)	SLPETGG-His6x
	645-935		WW2-4	pRBC	His6x-bdSUMO-(bdSEN1)-GGG	-
	600-935		WW1-4	pET24	His6x-(TEV)	-
	600-935		P871A	pET24	His6x-(TEV)	-

Protein Expression

BL21-DE3 *E. coli* cells transformed with the expression vector containing the gene of interest were grown 37 °C to an OD₆₀₀ of 0.6-0.7 and induced with isopropyl β-D-1-thiogalactopyranoside (IPTG). The specific conditions used to express each

protein of interest are listed in Table A1.2. Following expression, cell cultures were harvested by centrifugation for 25 min at 4,000 rpm.

Table A1.2: Protein expression conditions. *MJ9 media was supplemented with ^{15}N -ammonium chloride/ ^{13}C -glucose. TB: terrific broth. LB: lysogeny broth.

Construct	Media	Temperature after induction (°C)	IPTG (mM)	Expression time (h)
A ₁₂₃ /mutants (AMOTL1)	TB or MJ9*	20	0.1	5
YWW _{TD} /mutants (YAP)	LB or MJ9*	20	0.1	16
KWW _{TD} /mutants (KIBRA)	LB or MJ9*	20	0.1	16
WW1-4/P871A (NEDD4-1)	LB or MJ9*	20	0.1	16
WW1, WW2, WW3, WW4, WW1-2, WW2-3, WW3-4, WW1-3, WW2-4 (NEDD4-1)	LB	37	1	3
	MJ9*	20	0.4	16
eSrtA (Sortase)	LB	30	0.4	3

Protein purification and storage

A₁₂₃/mutants. Polypeptides of AMOTL1 were highly prone to C-terminal degradation, both during and after expression. To reduce degradation, protein purifications were performed on ice as much as possible in a single day, and experiments were performed within two days after purification except for NMR assignments which were performed over the course of 5 days at 10 °C in the presence of NMR protease inhibitors. Purified protein was unamenable to long-term storage at room temperature, 4, -20, and -80 °C due to degradation.

Harvested *E. coli* cells grown in TB were resuspended in affinity column buffer (ACB) containing 20 mM sodium phosphate, 500 mM NaCl, 10 mM imidazole, 5% glycerol, and 1 mM NaN₃ at pH 8.0. To reduce degradation, several protease inhibitors were added: 2 mM phenylmethylsulfonyl fluoride (PMSF), benzamidine hydrochloride, and Roche protease inhibitor cocktail or Aprotinin, E64, AEBSF, and pepstatin A. Resuspended cells were lysed by sonication and cleared by centrifugation for 30 min at 15,000 rpm. His-tagged protein remaining in the cleared cell lysate was purified by Ni-NTA affinity chromatography, followed by addition of protease

inhibitors and 1 mM ethylenediaminetetraacetic acid (EDTA) and size-exclusion chromatography (SEC) on a Superdex⁷⁵ or Superdex²⁰⁰ (Cytiva life sciences, MA, USA) column in buffer composed of 50 mM sodium phosphate, 200 mM sodium sulfate, 1 mM NaN₃, 1 mM EDTA, and 5 mM β-mercaptoethanol (BME) at pH 7.5. Due to its minimal solubility, expression of > 1 L A₁₂₃/mutants required multiple SEC purification steps. To minimize degradation, the purification tag which is part of the expression vector was not removed from the N-terminal end of the polypeptide. This introduced 26 non-native residues to the N-terminus of the A₁₂₃ polypeptides.

A₁₂₃ expressed in MJ9 was insoluble and went into inclusion bodies. Thus, the purification protocol described above was modified as follows. Harvested *E. coli* cells were resuspended in ACB containing 8 M urea. Following sonication, lysed cells were incubated at 37 °C, 100 rpm to solubilize inclusion bodies before centrifugation to remove cell debris. Ni-NTA affinity chromatography was performed using ACB with decreasing urea concentrations from 8 M to 2 M. The final elution containing 2 M urea and was injected directly onto the SEC column.

YWW_{TD}/KWW_{TD}/mutants. Harvested *E. coli* cells were resuspended in ACB, lysed by sonication, and cleared by centrifugation for 30 min at 15,000 rpm. His-tagged protein remaining in the cleared cell lysate was purified by Ni-NTA affinity chromatography and cleaved with 0.1 μM TEV protease overnight at room temperature in pH 7.5 buffer composed of 50 mM Tris Base, 100 mM NaCl. After TEV cleavage, Ni-NTA affinity chromatography was used to remove His-tagged/uncleaved proteins, followed by size-exclusion chromatography (SEC) on a Superdex⁷⁵ or Superdex²⁰⁰ (Cytiva life sciences, MA, USA) column in buffer composed of 50 mM sodium phosphate, 400 mM NaCl, 1 mM NaN₃, and 5 mM BME at pH 7.5. Proteins were stable at 4 °C up to two weeks after purification and were amenable to storage at -20 °C.

WW1-4/P871A. These proteins were purified as described for YWW_{TD}/KWW_{TD} with the following modifications. The solubility of WW1-4/P871A was salt sensitive. Thus, Ni-NTA was performed using ACB containing less (100 mM) NaCl. SEC purification was performed in buffer composed of 50 mM sodium phosphate, 100 mM NaCl, 1 mM NaN₃, and 5 mM BME at pH 7.5.

bdSUMO-NEDD4 constructs. NEDD4-1 constructs containing bdSUMO tags had a propensity to stick poorly to Ni-NTA resin. Thus, lysis and Ni-NTA affinity chromatography were performed in ACB containing only 5 mM imidazole. The bdSUMO tag was removed by incubation with bdSENPI1 protease for 1 h at 4 °C. Following cleavage, the bdSUMO tag were removed by Ni-NTA affinity chromatography. Proteins were further purified by SEC as described for YWW_{TD}/KWW_{TD}.

eSrtA. *E. coli* cells from 1 L expressions were lysed by sonication in 30 mL buffer composed 50 mM Tris, 0.3 M NaCl, 10 mM imidazole, 1 mM MgCl₂, 1 mM NaN₃ at pH 8.0 with 2 µL DNase, 1 mM PMSF, pepstatin A, E64, and AEBSF protease inhibitors, followed by incubation on ice for 10 min and centrifugation at 15,000 rpm for 30 min. Sortase proteins were purified by Ni-NTA affinity chromatography followed by SEC as described above in buffer composed of 50 mM Tris, 0.3 M NaCl, 1 mM MgCl₂, and 1 mM NaN₃ at pH 8.0. Purified enzyme was stored at -80 °C in buffer composed of 40 mM Tris, 110 mM NaCl, 2.2 mM KCl, 20% glycerol, and 1 mM NaN₃ at pH 7.5.

Purified protein concentrations were determined using absorbance measurements at 280 nm and extinctions coefficients calculated using ProtParam (<https://web.expasy.org/protparam/>).

Appendix 2**Preliminary data on competition between YAP, KIBRA, and NEDD4-1 for
AMOTL1**

Amber Vogel and Matthew McWhorter

Overview

Angiomotin-Like 1 (AMOTL1) is implicated in both tumor suppression and cell growth, and regulation of its cellular concentrations is an important regulatory step in these processes[119]. Three proteins known to regulate the stability of AMOTL1 are Yes-associated protein (YAP), Kidney and Brain-expressed protein (KIBRA), and neural precursor cell-expressed developmentally downregulated 4-1 (NEDD4-1). YAP and KIBRA each contain two WW domains, and NEDD4-1 contains four WW domains, which bind the three PPXY motifs of AMOTL1. Interactions with the bivalent WW domains of YAP and KIBRA stabilize AMOTL1, which has been seen to downregulate cell growth and tumorigenesis in certain contexts[2, 126]. In contrast, the interaction of AMOTL1 with NEDD4-1 results in ubiquitination and proteasomal degradation of AMOTL1, leading to enhanced cell growth[126, 141, 209]. The preceding chapters (Chapters 2, 3, and 5) focused on the mechanisms of assembly of AMOTL1 with YAP, KIBRA, and NEDD4-1. Our work shows that P1 is the predominant binding site for both YAP and NEDD4-1, with minor contributions from P2, while KIBRA prefers P3. These differences in motif preference allow YAP and KIBRA to bind AMOTL1 simultaneously. However, questions remain as to how AMOTL1 discriminates between NEDD4-1, YAP, and KIBRA. In this Appendix, we describe preliminary experiments to ascertain which partners AMOTL1 binds to when multiple WW domain proteins are present in the same environment.

Results

In this study, we used polypeptide sequences containing all binding-competent domains and motifs for AMOTL1 (residues 178-384, A₁₂₃), YAP (residues 157-277, YWW_{TD}), KIBRA (residues 1-91, KWW_{TD}), and NEDD4-1 (residues 600-935, WW1-4). To determine if AMOTL1 can form ternary complexes with NEDD4-1 and YAP or KIBRA, we performed SEC-MALS experiments with samples containing 1:1:1 or 1:1:2 mixtures of WW1-4:A₁₂₃:KWW_{TD} and WW1-4:A₁₂₃:YWW_{TD}. Figure A2.1A-B show the SEC-MALS elution profiles of these ternary mixtures plotted alongside each set of A₁₂₃-WW domain binary complexes for reference. Addition of both WW1-4 and

KWW_{TD} to A₁₂₃ resulted in an earlier elution volume and higher molar mass relative to either binary complex (Fig. A2.1A). The 1:1:1 and 1:1:2 WW1-4:A₁₂₃:KWW_{TD} complexes yielded molar masses of 63.0 and 63.8 kDa, respectively (Table A2.1). These values are below the theoretical ternary complex mass of 75.5 kDa, but above the A₁₂₃-WW1-4 and A₁₂₃-KWW_{TD} experimental masses of 55.9 and 35.2 kDa, respectively (Chapters 3 and 5). This data suggests that the three proteins form a mixture of species that may include a ternary complex. In contrast, addition of both WW1-4 and YWW_{TD} to A₁₂₃ resulted in an elution volume between the A₁₂₃-WW1-4 and A₁₂₃-YWW_{TD} elution volumes (Fig. A2.1B). Furthermore, the 1:1:1 and 1:1:2 WW1-4:A₁₂₃:YWW_{TD} mixtures yielded molar masses of 56.7 and 56.9 kDa, values between the theoretical masses of the A₁₂₃-WW1-4 and A₁₂₃-YWW_{TD} binary complexes, suggesting that mixing all three species results in an equilibrium of binary complexes.

Table A2.1: SEC-MALS of A₁₂₃ bound to WW1-4, YWW_{TD}, and KWW_{TD}.
Experimental masses are the average of duplicate experiments.

Complex	Theoretical MW, kDa	Experimental MW, kDa
1:1:1 WW1-4:A ₁₂₃ :KWW _{TD}	75.5	63.0
1:1:2 WW1-4:A ₁₂₃ :KWW _{TD}	86.8	63.8
1:1:1 WW1-4:A ₁₂₃ :YWW _{TD}	78.4	56.7
1:1:2 WW1-4:A ₁₂₃ :YWW _{TD}	92.6	56.9

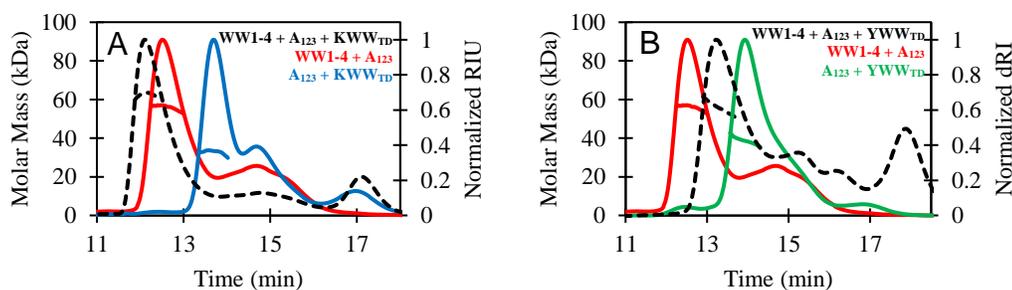


Figure A2.1: Competition SEC-MALS. SEC-MALS elution profiles of A₁₂₃ combined with equimolar ratios of (A) KWW_{TD} (blue), WW1-4 (red), and KWW_{TD} and WW1-4 (black, dashed); and (B) YWW_{TD} (green), WW1-4 (red), and YWW_{TD} and WW1-4 (black, dashed).

Next, we performed competition ITC experiments in which A_{123} was pre-saturated with WW1-4, YWW_{TD} , or KWW_{TD} , and titrated with a competing WW domain protein (Fig. A2.2). Titration of KWW_{TD} into A_{123} -WW1-4 and WW1-4 into A_{123} - KWW_{TD} resulted in K_d values of 0.7 and 1.1 μM , respectively (Fig. A2.2A-B, Table A2.2). These values are similar or equivalent to the K_d values obtained for either set of binary complexes (Chapters 3 and 5). However, the favorable enthalpy and unfavorable entropy contributions are much weaker in context of all three proteins relative to the binary complexes. Furthermore, titration of KWW_{TD} into A_{123} -WW1-4 produced a stoichiometry of 0.5. These data suggest that initial binding of A_{123} to KWW_{TD} or WW1-4 reduces the binding interactions of the subsequent WW domain partner, resulting in a mixture of binary and ternary complexes. By contrast, titration of YWW_{TD} into A_{123} -WW1-4 and WW1-4 into A_{123} - YWW_{TD} resulted in weak binding (K_d of 14 μM) and no binding, respectively. Like the WW1-4- A_{123} - KWW_{TD} competition experiment, the enthalpy and entropy contributions were both greatly reduced. These data suggest that A_{123} cannot bind WW1-4 and YWW_{TD} simultaneously. Instead, the two WW domain proteins compete for binding to A_{123} .

Table A2.2: Thermodynamic parameters for WW1-4/KWW_{TD}/YWW_{TD} titrated into preformed A₁₂₃-WW domain complexes. Values reported for titration of KWW_{TD} into A₁₂₃-WW1-4 are the average of duplicate experiments. All other reported values were obtained from single experiments. No binding: n.b. * Data from Chapter 5 for comparison. ^aData from Chapter 3 for comparison. ^bData from Chapter 2, reference [4] for comparison.

Syringe	Cell	N	Kd (μM)	ΔH (kcal/mol)	TΔS (kcal/mol)	ΔG (kcal/mol)
WW1-4	A ₁₂₃ *	0.9 ± 0.1	0.4 ± 0.1	-42.4 ± 0.5	-33.7 ± 0.6	-8.7 ± 0.1
KWW _{TD}	A ₁₂₃ ^a	0.9 ± 0.1	0.7 ± 0.2	-37 ± 6	-29 ± 6	-8.4 ± 0.1
YWW _{TD}	A ₁₂₃ ^b	1.1	0.26 ± 0.01	39.6 ± 0.9	30.3 ± 0.9	8.98 ± 0.03
WW1-4	A ₁₂₃ -KWW _{TD}	1	1.1	-27	-19	-8.1
KWW _{TD}	A ₁₂₃ -WW1-4	0.5	0.7	-21	-13	-8.4
WW1-4	A ₁₂₃ -YWW _{TD}			n.b.		
YWW _{TD}	A ₁₂₃ -WW1-4	0.5	14	-28	-21	-6.6

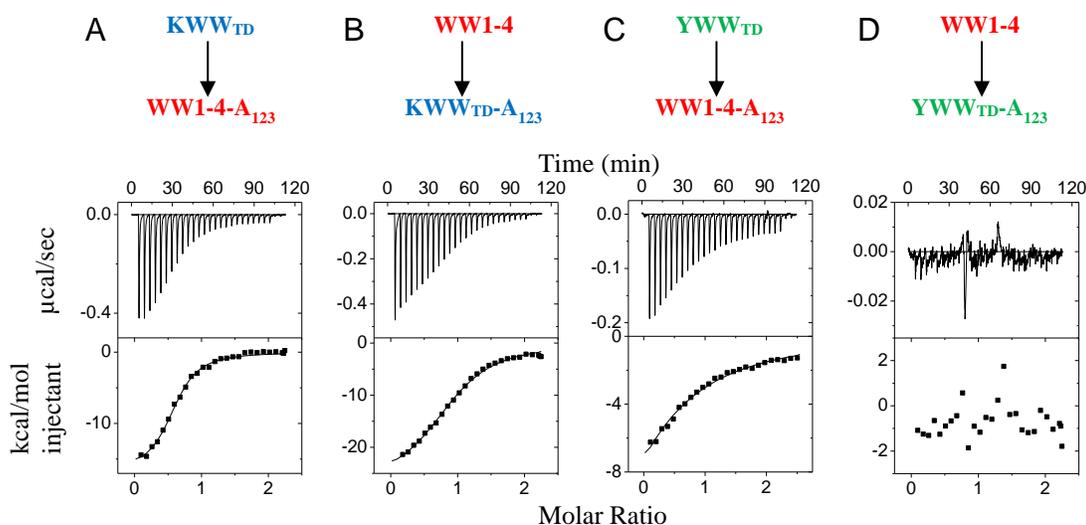


Figure A2.2: Competition ITC. Competition ITC experiments of (A) KWW_{TD} titrated into a preformed WW1-4-A₁₂₃ complex, (B) WW1-4 titrated into a preformed KWW_{TD}-A₁₂₃ complex, (C) YWW_{TD} titrated into a preformed WW1-4-A₁₂₃ complex, and (D) WW1-4 titrated into a preformed YWW_{TD}-A₁₂₃ complex collected at 25 °C.

Conclusions

These preliminary data suggest that WW domain proteins that bind different sites of AMOTL1 can bind simultaneously, while partners that prefer the same PPXY motifs compete for binding. For WW1-4 and KWW_{TD}, which prefer P1 and P3, respectively, our preliminary findings show that a ternary complex may form. However, these WW domain proteins diminish the interactions of the alternate partner, resulting in enthalpy-entropy compensation that does not result in large changes in binding affinities. In contrast, WW1-4 and YWW_{TD}, which both prefer P1 of AMOTL1, compete for binding to A₁₂₃. To validate these findings, these experiments must be replicated with corroborating SDS PAGE gel electrophoresis following SEC-MALS to see which sets of proteins elute together. Additionally, further work is needed to determine the biological relevance of these findings.

Materials and methods

Cloning of constructs and recombinant protein production are described in Appendix 1.

SEC-MALS. Absolute molecular weights were obtained by size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) using an AKTA FPLC (GE Healthcare), a DAWN multiple-angle light scattering, and an Optilab refractive index system (Wyatt Technology). Data were collected at room temperature, on a Superdex²⁰⁰ 10/300 GL column (Cytiva life sciences) pre-equilibrated with 50 mM sodium phosphate, 400 mM NaCl, 1 mM NaN₃, and 5 mM β-mercaptoethanol at pH 7.5. Protein samples at concentrations of 20-50 μM were injected at a flow rate of 0.7-0.8 mL/min. The data were analyzed with ASTRA software package, version 8 (Wyatt Technology).

Isothermal Titration Calorimetry. Isothermal titration calorimetry (ITC) data were recorded on a VP-ITC instrument (Malvern instruments Inc, MA) set to 25 °C. Prior to data collection, proteins were dialyzed against buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN₃, and 5 mM β-mercaptoethanol at pH 7.5. Twenty-seven or twenty-eight injections of 68-135 μM WW domain proteins were

titrated into 6-10 μM A_{123} mixed with two-fold excess competing WW domain partner. Isotherms were analyzed by single-site fits of the thermograms using the *Origin 7.0* software. The free energy of binding (ΔG) was calculated from the equation $\Delta G = -RT\ln(K_a)$, where R is the universal gas constant, T is temperature in Kelvin, and K_a is the association constant.

Appendix 3**NMR analysis of the tandem WW domains of Yes-associated protein**

Amber Vogel and Miranda Leek

Overview

Mapping the binding interface of the Yes-associated protein (YAP) tandem WW domains in detail is critical for understanding its binding mechanisms to partner proteins such as Angiomotin-Like 1 (AMOTL1). NMR is a useful technique in obtaining residue-level information on structure and binding. In this appendix, we describe preliminary NMR experiments conducted on the tandem WW domains of YAP with the goal of identifying optimum conditions for solution NMR studies of the YAP WW domains.

Results

Initial NMR characterization shows heterogeneity in the YAP WW domains. In this work, we used a construct of YAP containing both tandem WW domains (residues 157-277, referred to as YWW_{TD}). NMR data are usually collected in low pH buffer to reduce the rate of amide proton exchange and increase sensitivity. However, since YWW_{TD} has a pI of 6.04, we initially chose to use a buffer at pH 7.3 composed of 10 mM sodium phosphate, 10 mM NaCl, and 1 mM NaN₃ (buffer A) to improve solubility. The YWW_{TD} construct contains no cysteine residues, so no reducing agent was included. YWW_{TD} contains 110 native, non-proline residues. To ascertain the data quality for our initial conditions, we collected BEST-TROSY HSQC spectra at a range of temperatures (Fig. A3.1A). At the lowest temperature (10 °C), we observed roughly 102 peaks (93%) with a range of chemical shifts mapping to both structured and disordered/helical regions of the spectrum. However, the dispersed peaks showed weak intensity, and subsequent attempts to assign these peaks resulting in weak or no intensity in the 3D experiments. Increasing the temperature caused peak disappearance in the disordered region of the spectrum but increased the sensitivity of structured peaks.

Using 3D experiments collected at both 10 °C and 35 °C allowed us to assign 62 (56%) residues (Fig. A3.1B). Of these assignments, 37 belonged to the linker and terminal residues, 4 belonged to residues in the WW1 domain, and 21 belonged to residues within the WW2 domain. At 10 °C, domain residues showed much weaker

Attempts to optimize NMR conditions for YWW_{TD}. To obtain spectra with more WW domain peaks, we next performed a series of HSQC experiments with varied conditions including construct mutagenesis, sample concentration, buffer composition, temperature, and pH. First, we collected HSQCs at 10 °C and 35 °C of YWW_{TD} with a highly conserved Pro mutated to Ala: P202A_{TD} (Fig. A3.2). While there were a few changes in the spectra of P202A_{TD} relative to YWW_{TD}, the spectra did not show many new peaks. Therefore, we did not pursue this mutant further. We next used the same buffer conditions (buffer A) but increased the sample concentration to 1.4 mM and collected HSQCs at different temperatures (Fig. A3.3A-C). Several new peaks were observed at 25 and 35 °C, but most of these were weak and showed up in the disordered region of the spectrum between 8-8.5 ppm in the proton dimension. Furthermore, at 10 °C, only the disordered peaks were present, suggesting that the sample may be partly aggregated.

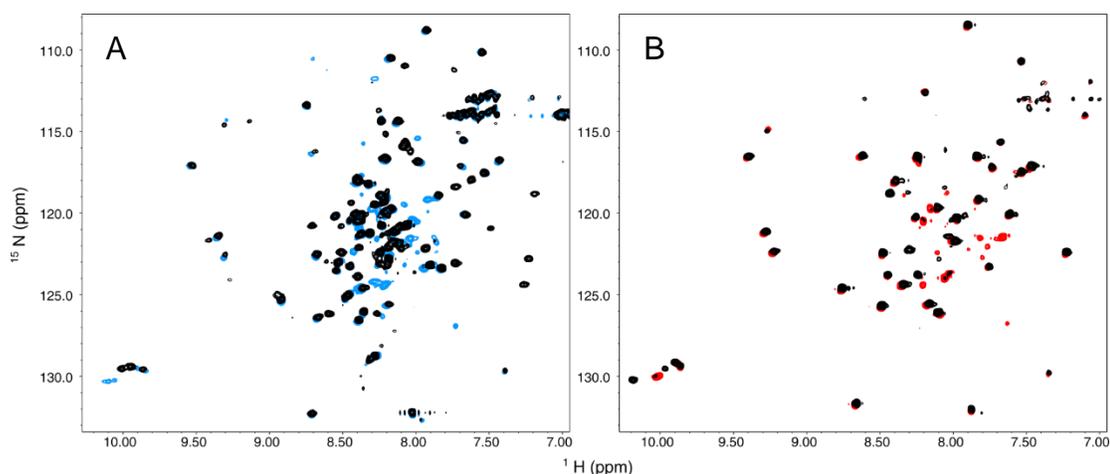


Figure A3.2: HSQCs of P202A_{TD}. ¹H-¹⁵N HSQC spectra YWW_{TD} (black) overlaid on P202A_{TD} collect at (A) 10 °C (blue), and (B) 35 °C (red). Samples were prepared in buffer composed of 10 mM sodium phosphate, and 10 mM NaCl, 1 mM NaN₃ at pH 7.3.

Since the buffer used for previous NMR characterization of the YAP tandem WW domains contained potassium phosphate and ethylenediaminetetraacetic acid (EDTA) at pH 5.5[220], we next tried changing the buffer composition to 50 mM potassium phosphate, 50 mM KCl, 0.5 mM EDTA, and 1 mM NaN₃ (buffer B). Due to the pI of our construct, we first tried changing the pH to 6.8. We collected HSQCs of samples at 132 μM at 10, 25, and 35 °C (Fig. A3.3D-F). A few new peaks appeared in

the disordered region of the spectrum, but no new dispersed peaks were present. Next, we tried reducing the pH to 5.0 using buffers composed of 50 mM citrate phosphate, 50 mM NaCl or KCl, 0.5 mM EDTA, and 1 mM NaN₃ (buffers C and D) and collected HSQCs of samples ranging from 230-305 μ M at 10, 25, and 35 °C (Fig. A3.3G-I). Several new peaks were present in these HSQCs, and many of the peaks that disappeared at higher temperature in previous conditions remained at 35 °C. No spectral differences were observed for samples containing NaCl versus KCl.

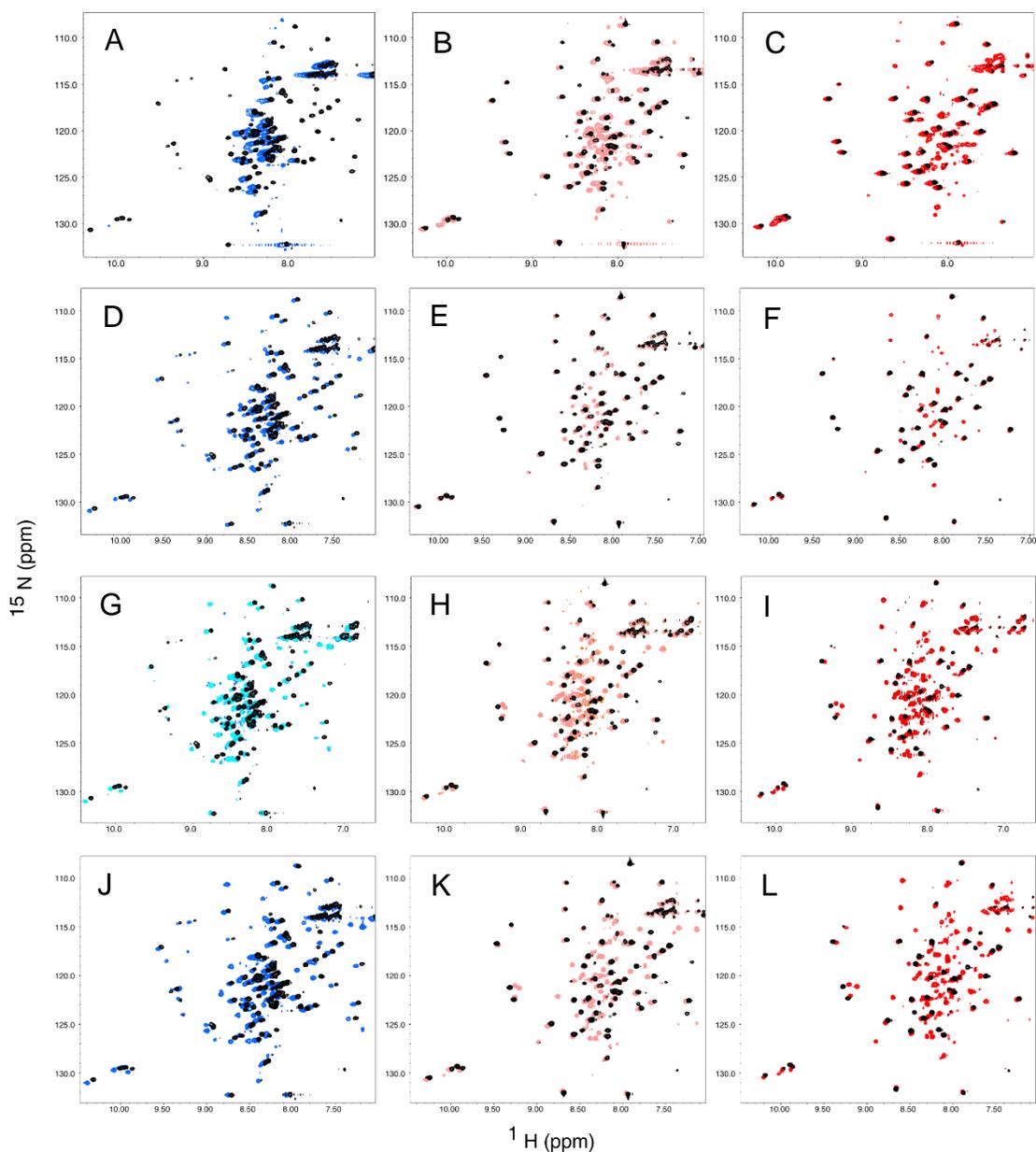


Figure A3.3: Optimization of YWW_{TD} NMR conditions. ^1H - ^{15}N HSQC spectra of 345 μM YWW_{TD} in buffer composed of 10 mM sodium phosphate, 10 mM NaCl, and 1 mM NaN₃ at pH 7.3 overlaid on (A-C) 1.4 mM YWW_{TD} in the same buffer; (D-F) 132 μM YWW_{TD} in buffer composed of 50 mM potassium phosphate, 50 mM KCl, 0.5 mM EDTA, and 1 mM NaN₃ at pH 6.8; (G-I) 230-305 μM YWW_{TD} in buffer composed of 50 mM citrate phosphate, 50 mM NaCl or KCl, 0.5 mM EDTA, and 1 mM NaN₃ at pH 5.0; and (J-L) 160 μM YWW_{TD} in buffer composed of 50 mM MES, 50 mM KCl, 0.5 mM EDTA, and 1 mM NaN₃ at pH 5.0. Spectra were collected at 10 °C (left, blue), 25 °C (middle, salmon), and 35 °C (right, red).

Encouraged by the improved HSQCs at low pH, we tested an additional buffer

YWW_{TD}. Small to moderate exchange terms were observed for a handful of residues, but due to large error, the data was not analyzed further.

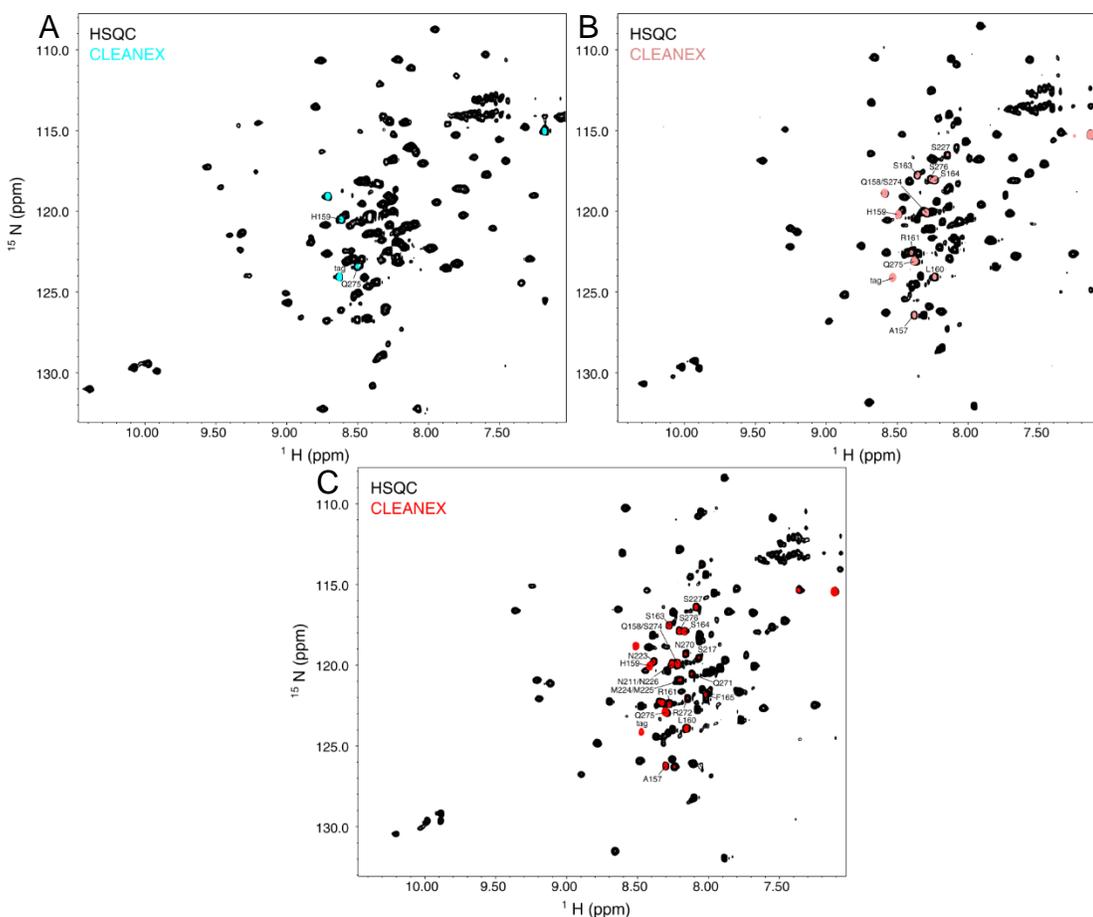


Figure A3.5: Amide solvent exchange measured by CLEANEX NMR. ^1H - ^{15}N HSQC spectra of YWW_{TD} (black) overlaid with CLEANEX spectra collected at (A) 10 °C (blue), (B) 25 °C (salmon), and (C) 35 °C (red). Peaks observable in CLEANEX experiments are labeled in each spectrum and correspond to H159 and Q275 at 10 °C; A157, Q158, H159, L160, R161, S163, S164, S227, S274, Q275, and S276 at 25 °C; A157, Q158, H159, L160, R161, S163, S164, F165, N211, S217, N223, M224, M225, N226, S227, N270, Q271, R272, S274, Q275, and S276 at 35 °C.

To determine the degree of solvent accessibility for assigned residues, we performed CLEANEX NMR experiments in buffer E at 10, 25, and 35 °C (Fig. A3.5). Intriguingly, very few peaks were detectable in the CLEANEX experiments, indicating that the amide protons of most residues are protected from solvent. Furthermore, the number of peaks increased with increasing temperature, indicating that the exchange

rate of amide protons is temperature dependent. At 35 °C, CLEANEX peaks for 21 assigned residues corresponding to the linker and termini were present.

Conclusions

Obtaining residue-level insight into the structure, dynamics, and binding of the tandem WW domains of YAP will illuminate how these domains function. In this work, we attempted to obtain quality NMR data for a segment of YAP containing both WW domains. In our initial conditions, only 62 of the 110 native, non-proline residues were assigned. After sample optimization, we were able to assign a total of 78 residues. Unfortunately, missing peaks and low signal intensity in 3D experiments precluded assignment of 20 peaks in the WW1 domain. The poor data quality for WW1 is likely an artifact of intermediate chemical exchange. Additional work is needed to further optimize the NMR conditions for YWW_{TD}. Future work could include adding a WW1-binding PPXY peptide to potentially stabilize the domain, modifying buffer conditions and pH, and changing NMR parameters.

Materials and Methods

Protein preparation. Cloning protocols and recombinant protein expression and purification are described in Appendix 1. Protein samples for NMR experiments ranged from 132-1400 μ M. All NMR samples contained 10% D₂O as the lock nucleus and 0.5% DSS for internal referencing.

NMR data collection and analysis. NMR experiments were conducted using a Bruker Avance III, 800 MHz spectrometer (Bruker Biospin) equipped with a triple resonance cryogenic probe. Sample conditions were compared using 2D BEST-TROSY ¹H-¹⁵N HSQC experiments[169]. Backbone assignments were performed using a suit of 3D experiments collected with non-uniform sampling (NUS): HNCACB, HNCOCACB, HNCO, HNCACO, and HNCA. NUS data reconstruction was performed using the iterative shrinkage threshold approach in NMRPipe[170].

¹⁵N BEST-TROSY Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments were performed at 25 °C with a constant delay time of 40 ms using the

BEST modified pulse sequence of Franco et al. [221]. Data were collected with 12 different interleaved refocusing pulse delay values of 0, 25, 50 (x2), 100, 150, 200, 250, 400, 500, 600, 750, and 1000 Hz. Experimental error was calculated from experimental duplicates (x2) using Monte Carlo analysis. CPMG data analysis was performed in Relax[222]. TROSY-based CLEAN chemical exchange (CLEANEX) experiments were performed with mixing times of 100 ms and recycle delays of 1-1.5 s. All NMR data were processed in NMRPipe[170] and visualized in Sparky[171] or NMRView[205].

Appendix 4

Interactions of YAP, KIBRA, and NEDD4-1 WW domains with viral PPXY sequences

Amber Vogel and Matthew McWhorter

Overview

Viral late domains (L domains) are conserved regions of viral proteins important for the “late” stages of the viral life cycle. L domains contain one of three classes of short linear motifs: PT/SAP, YP(X)_nL/LXXLF, and PPXY motifs. These motifs bind host proteins involved in endosomal sorting complexes required for transport (ESCRT) pathway. These interactions then cause membrane fission and escape of newly formed viral particles from the cell, a process termed viral egress[223]. Many viruses express proteins with one or more L domains including human immunodeficiency virus (HIV), Hepatitis B virus (HBV), T-lymphotropic virus 1 (HTLV-1), SARS-CoV-2, Ebola virus, Marburg virus, and herpesviruses to name a few[138, 142, 223-225]. Viral L domains containing PPXY motifs can bind the WW domains of E3 ubiquitin ligase protein from the NEDD4 family which link viral proteins to the ESCRT pathway[138]. A prominent example is the PPAY motif at the N-terminus of SARS-CoV-2 Spike (S) protein, which supposedly interacts with NEDD4, leading to ubiquitination and enhanced viral egress and infectivity[142].

In addition to ubiquitin ligase proteins, recent findings implicate Hippo signaling pathway proteins containing WW domains or PPXY motifs in the regulation of viral egress. In some cases, viral proteins lack a PPXY motif and require an adaptor to link them to NEDD4 proteins. For example, HIV Gag protein contains PTAP and YPXL motifs, but lacks the PPXY motif. Therefore, HIV Gag protein binds AMOT which bridges Gag and NEDD4L, promoting viral egress[163]. A contrasting example is the VP40 matrix protein from Marburg virus, which contains a PPXY motif. While NEDD4 enhances Marburg egress, the VP40 PPXY motif can also interact with YAP and TAZ transcriptional coactivators, which reduce viral egress[226]. Intriguingly, AMOT competes with VP40 for binding to YAP/TAZ, which results in enhanced viral egress. Given the many WW domains and PPXY motifs in the proteome, these findings open investigation into numerous potential virus-host interactions.

The goal of this study was to identify novel host WW domain protein-viral PPXY peptide interactions. PPXY-containing sequences of viral proteins, HTLV-1 Gag, HBV Core, SARS-CoV S, and SARS-CoV-2 S were used to design 20-residue peptides (Table A4.1, Fig. A4.1). In our initial binding screen, we identified a novel

interaction between HBV Core PPAY peptide and the WW domains of Kidney and Brain-expressed protein (KIBRA). However, a longer HBV Core polypeptide (residues 1-149) failed to bind the KIBRA WW domains nor its previously identified WW domain partner, NEDD4. Furthermore, 20- and 70-residue polypeptide-based studies showed no interaction between SARS-CoV-2 S protein and NEDD4, as previously suggested[142].

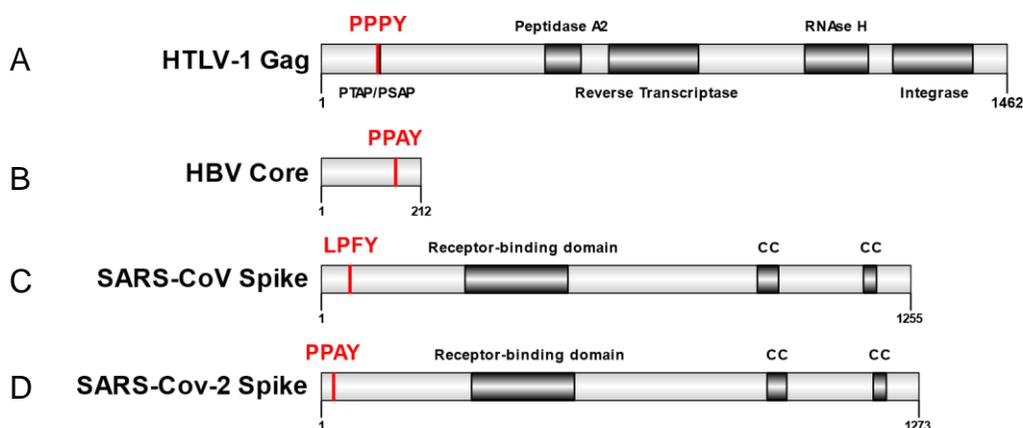


Figure A4.1: Viral protein domain architecture[1]. (A) Domain architecture of full-length (FL)-HTLV-1 Gag protein containing an N-terminal PPPY motif, PTAP/PSAP motif, peptidase A2 domain, reverse transcriptase domain, RNase H domain, and C-terminal Integrase domain. (B) Domain architecture of FL-Hepatitis B Core protein containing a PPAY motif. (C) Domain architecture of FL-SARS-CoV Spike protein, containing an N-terminal LPFY motif, Receptor-binding domain, and two C-terminal coiled-coil (CC) domains. (D) Domain architecture of FL-SARS-CoV-2 Spike protein, containing an N-terminal PPAY motif, Receptor-binding domain, and two C-terminal coiled-coil (CC) domains.

Table A4.1: Sequences of viral PPXY peptides.

Virus	Protein	Residues	Sequence
HTLV-1	Gag	110-129	PPDSDPQI <u>PPPY</u> VEPTAPQV
HBV	Core	150-169	SFGVWIRT <u>PPAY</u> RPPNAPIL
SARS-CoV	Spike	52-71	LYLTQDLF <u>LPFY</u> SNVTGFHT
SARS-CoV-2	Spike	17-36	NLTTRTQL <u>PPAY</u> TNSFTRGV

Results

ITC and NMR studies reveal novel virus-host binding interactions. To identify novel WW domain-viral L domain interactions, we performed an ITC-based binding screen using three host WW domain proteins and four 20-residue viral PPXY peptides.

The host proteins included the tandem WW domains of YAP (residues 157-277, YWW_{TD}), the tandem WW domains of KIBRA (residues 1-91, KWW_{TD}), the tetravalent WW domains of NEDD4-1 (residues 600-935, WW1-4), and the WW3 domain of NEDD4-1 (residues 835-878, NWW3). The viral proteins used in our screen included: (1) HTLV-1 Gag protein, an identified substrate of NEDD4 required for viral egress involved in YAP-mediated oncogenesis[227, 228]; (2) original SARS-CoV Spike (S) protein, which contains an undocumented LPFY motif; (3) SARS-CoV-2 S protein, a recently identified substrate of NEDD4 family proteins; and (4) HBV Core (HBC) protein, another NEDD4 substrate and viral activator of YAP activity[224, 229].

Figure A4.2 shows representative isotherms of each peptide titrated into host WW domain proteins. Titration of HTLV-1 Gag PPPY motif into YWW_{TD}, KWW_{TD}, and WW1-4 yielded weak binding affinities (K_d values $>100 \mu\text{M}$) (Fig. A4.2A-C, Table A4.2). In a previous study, a 12-residue peptide of HTLV-1 Gag PPPY motif interacted with the WW3 domain of NEDD4 with an affinity of $61 \mu\text{M}$ [88]. Surprised by the weak binding observed for our HTLV-1 Gag peptide and WW1-4, we next titrated the peptide into a construct of NEDD4-1 containing only WW3 (NWW3) which yielded a binding affinity of $42 \mu\text{M}$, consistent with previously reports (Fig. A4.2D, Table A4.2). ITC experiments measure average binding affinities of all interactions present. Thus, the reduced binding affinity of HTLV-1 Gag PPPY for WW1-4 versus NWW3 suggests that interactions with other domains are weak, resulting in an average weaker binding affinity.

Next, we titrated the LPFY motif from SARS-CoV S protein into YWW_{TD}, KWW_{TD}, WW1-4, and NWW3 and saw weak or no binding (Fig. A4.2E-H, Table A4.2). This is unsurprising, given that this motif is unreported in the literature and is likely not a biologically relevant interaction motif. We next titrated a peptide of the PPAY motif from SARS-CoV-2 S into our WW domain proteins and found that the peptide binds YWW_{TD}, KWW_{TD}, and WW1-4 with affinities of 79, 188, and $143 \mu\text{M}$, respectively (Fig. A4.2I-K, Table A4.2). These K_d values are fairly weak but indicate potentially relevant interactions of SARS-CoV-2 S with these WW domain proteins. Like the HTLV-1 Gag PPPY peptide, binding of SARS-CoV-2 S PPAY peptide to

NWW3 was much tighter than WW1-4, having a K_d of 46 μM (Fig. A4.2L), indicating weak contributions from the other domains.

Table A4.2: Binding parameters for viral peptides titrated into WW domain proteins monitored by ITC. ^a Average calculated from experimental repeats. ^b Average and error calculated from triplicate experiments. All other reported values were obtained from single experiments. No binding: n.b.

Syringe (virus/protein)	Cell	N	K_d (μM)	ΔH (kcal/mol)	TAS (kcal/mol)	ΔG (kcal/mol)
HTLV-1/ Gag	YWW _{TD}	3.9	156	-4	1	-5.2
	KWW _{TD}	1.7	228	-18	-13	-5.0
	WW1-4	2.5	182	-16	-11	-5.1
	NWW3 ^a	0.8	42	-23	-17	-6.0
SARS-CoV/ Spike	YWW _{TD}	<i>n.b.</i>				
	KWW _{TD}	<i>n.b.</i>				
	WW1-4	<i>weak binding</i>				
	NWW3	<i>weak binding</i>				
SARS-CoV2/ Spike	YWW _{TD}	2.5	79	-13	-7	-5.6
	KWW _{TD}	0.6	188	-39	-34	-5.1
	WW1-4	2.7	143	-12	-6	-4.4
	NWW3 ^a	1.1	46	-15	-10	-5.9
HBV/ Core	YWW _{TD} ^b	2.1 ± 0.2	38 ± 6	-19 ± 2	-13 ± 2	-6.0 ± 0.08
	KWW _{TD} ^b	1.0 ± 0.1	9.9 ± 0.5	-35 ± 1	-28 ± 1	-6.8 ± 0.03
	KWW1 ^b	1.2 ± 0.1	21 ± 1	-17 ± 2	-10 ± 2	-6.4 ± 0.04
	KWW2	1.8	231	-9	-4	-5.0
	WW1-4	3.6	49	-4	2	-5.9

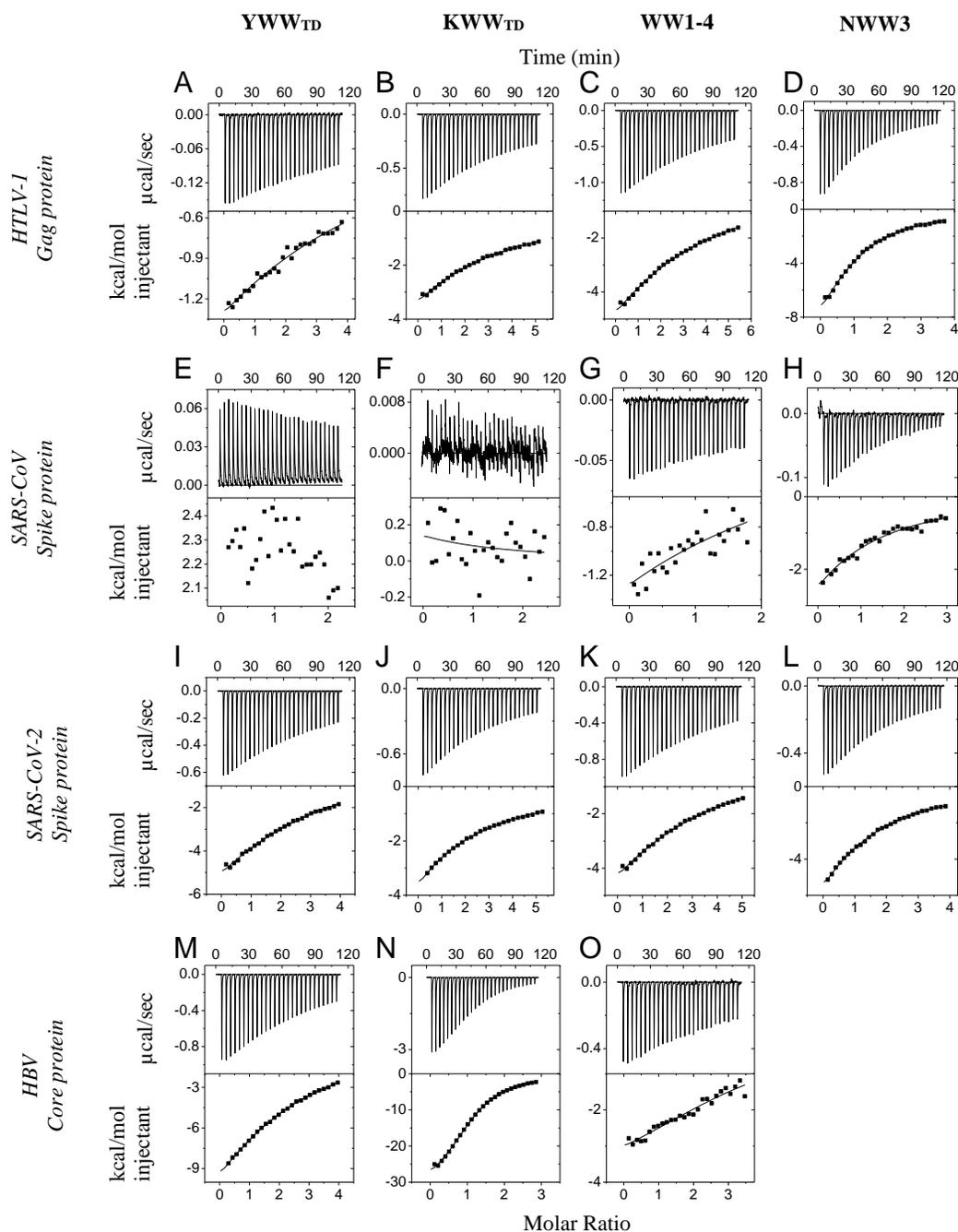


Figure A4.2: ITC of viral peptides titrated into YWW_{TD} (YAP), KWW_{TD} (KIBRA), WW1-4 and NWW3 (NEDD4-1). Representative isotherms of HTLV-1 Gag PPPY peptide titrated into (A) YWW_{TD}, (B) KWW_{TD}, (C) WW1-4, and (D) NWW3; SARS-CoV Spike LPFY peptide titrated into (E) YWW_{TD}, (F) KWW_{TD}, (G) WW1-4, and (H) NWW3; SARS-CoV-2 Spike PPAY peptide titrated into (I) YWW_{TD}, (J) KWW_{TD}, (K) WW1-4, and (L) NWW3; and HBV Core PPAY peptide titrated into (M) YWW_{TD}, (N) KWW_{TD}, and (O) WW1-4 at 25 °C.

Finally, we titrated the HBC PPAY peptide into our WW domain proteins. To our amazement, the HBC peptide bound YWW_{TD} , KWW_{TD} , and WW1-4 with affinities much higher than observed for the other interactions, having K_d values of 38 ± 6 , 9.9 ± 0.5 , and $49 \mu\text{M}$, respectively (Fig. A4.2J-L, Table A4.2). Binding to bivalent YWW_{TD} and tetravalent WW1-4 occurred with stoichiometries of close to 2 and 4, respectively, indicating that one peptide binds each WW domain. However, the HBC peptide bound KWW_{TD} with a stoichiometry of 1, indicating that the peptide binds only one WW domain of KIBRA. Thus, we titrated the HBC peptide into constructs of KIBRA containing only WW1 (KWW1) or WW2 (KWW2). Surprisingly, we obtained K_d values of 21 ± 1 and $231 \mu\text{M}$ for KWW1 and KWW2, respectively. These data indicate that WW1 is the predominant binder but requires WW2 to achieve full binding affinity.

To investigate the HBC peptide- KWW_{TD} interaction on the residue-level, we next performed a titration of HBC peptide into ^{15}N -labeled KWW_{TD} and monitored binding by NMR (Fig. A4.3). Changes in peak intensity and chemical shift are indicators of binding. Upon addition of HBC peptide, two types of spectral changes were observed: peak disappearance and chemical shift perturbations (CSPs). As expected, these changes were more pronounced in the WW1 domain. Since the peptide is small, spectral changes were attributed to chemical exchange between free and bound states as opposed to fast T2 relaxation caused by increased molecular size. At sub-stoichiometric ratios of peptide, many peaks in WW1 disappeared, indicating intermediate chemical exchange on the NMR timescale. However, several peaks shifted, and a few peaks disappeared and then reappeared upon addition of excess peptide, indicating fast and slow chemical exchange, respectively. Most peaks in WW2 were not perturbed until addition of excess peptide. These data suggest that binding may be anchored by WW1, which undergoes chemical exchange on multiple timescales. These data are consistent with ITC data that show WW1 is the predominant binding site but is stabilized by WW2.

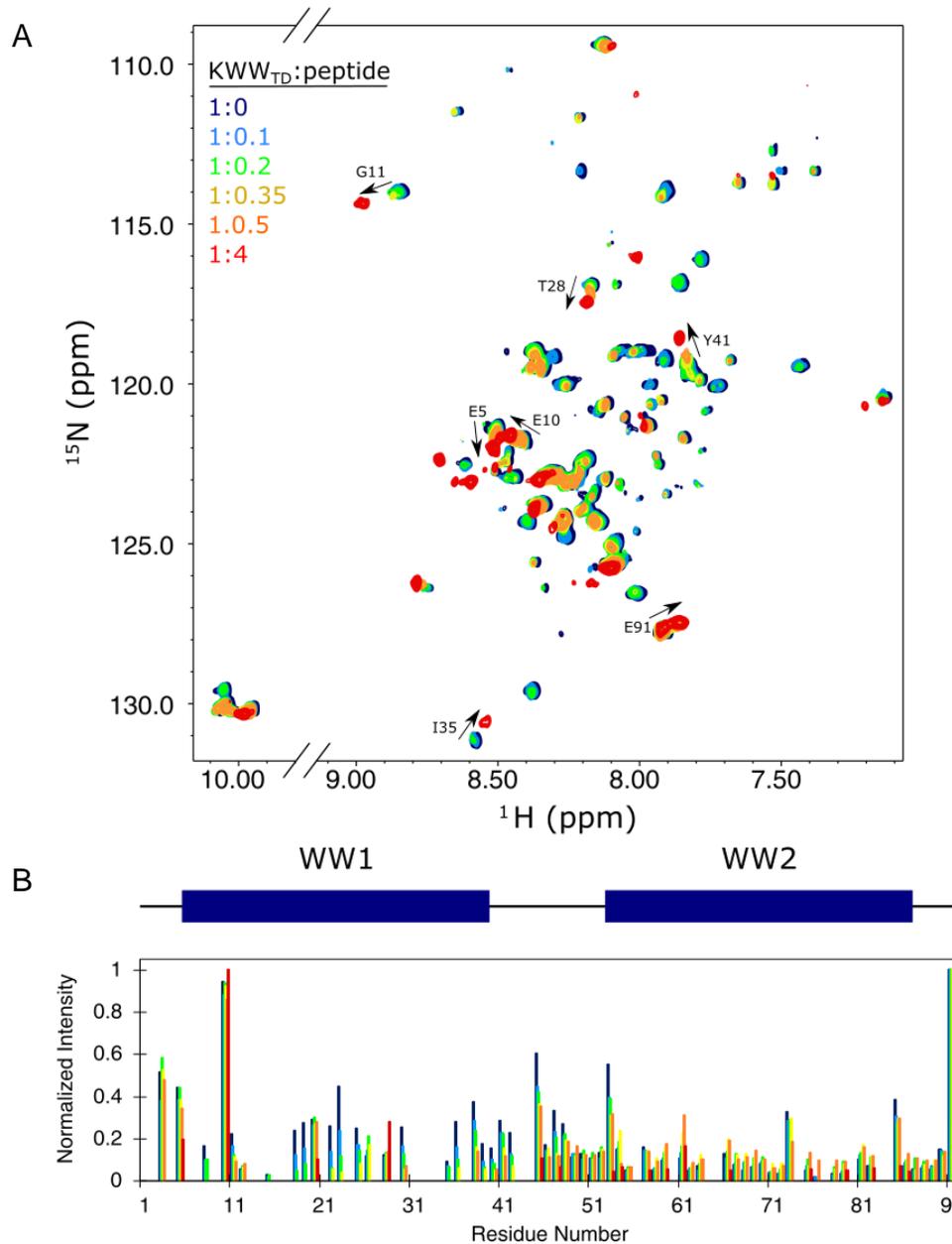


Figure A4.3: NMR titration of HBC PPAY peptide into ^{15}N -labeled KWW_{TD} . (A) ^1H - ^{15}N -HSQC overlay of unlabeled HBC PPAY peptide titrated into ^{15}N - KWW_{TD} at molar ratios of 1:0 (dark blue), 1:0.1 (light blue), 1:0.2 (neon green), 1:0.35 (yellow), 1:0.5 (orange), and 1:4 (red) KWW_{TD} :HBC peptide. Assigned peaks with notable chemical shifts are labeled: E5, E10, G11, T28, I35, Y41, and E91. (B) Normalized ^1H - ^{15}N -HSQC peak intensities of ^{15}N - KWW_{TD} plotted versus residue number. Each titration point is colored as described above.

To quantify changes in chemical exchange on the intermediate timescale, we next performed Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion NMR experiments. Due to peak broadening at increasing peptide concentrations, CPMG experiments were performed with free KWW_{TD} and KWW_{TD} bound to 0.1 equivalents HBC peptide. CPMG analysis revealed 20 residues in apo KWW_{TD} with k_{ex} terms: 11 in WW1, 5 in the linker, 3 in WW2, and one at the C-terminus. Upon addition of 0.1 equivalents peptide, changes were seen in the magnitude and identity of residues experiencing chemical exchange. Unfortunately, low signal to noise in the NMR spectra impeded reliable chemical exchange quantification. Therefore, the CPMG data was not analyzed further.

In summary, we used a low-throughput ITC-based binding screen to detect novel or confirm three host-virus interactions: (1) SARS-CoV-2 S PPAY motif peptide and WW3 of NEDD4-1 (and potentially YAP and KIBRA), (2) HBC PPAY motif peptide and the tandem WW domains of YAP, and (3) HBC PPAY motif peptide and the tandem WW domains of KIBRA. We then used ITC and NMR experiments to show that HBC peptide binds WW1 of KIBRA with minor contributions from WW2. Peak broadening and preliminary CPMG data suggest that numerous residues of KWW_{TD} may experience chemical exchange on the intermediate timescale, both in the apo and peptide-bound states. These chemical exchange processes may influence the binding behavior of KWW_{TD} for partners such as HBC peptide. Having identified novel viral peptide-host WW domain complexes, we next sought to confirm these interactions in context of larger polypeptide chains of the viral L domains.

Binding Assays of nCoV-S and WW1-4 show no complex formation. In the previous section, we saw weak binding between SARS-CoV-2 PPAY peptide and WW1-4 (Table A4.2). To see if binding was tighter in context of a larger, more native piece of SARS-CoV-2 S, we generated a construct spanning residues 1-70 (nCoV-S). We next performed an ITC titration of nCoV-S into WW1-4 and saw no binding (Fig. A4.4), which may be explained by several experimental shortcomings. Due to its poor expression, we had to add a bdSUMO tag to the N-terminus of nCoV-S. The tag may have buried the PPAY motif, causing it to be inaccessible for WW domain binding. Additionally, overexpressed bdSUMO-nCoV-S was insoluble and required

purification under denaturing conditions. Therefore, it is possible the protein was aggregated or not properly folded. Finally, while previously reported pull-down experiments showed an interaction between nCoV Spike protein and NEDD4[142], it is possible that the interaction is indirect and requires an adaptor molecule. Further work is needed to resolve these discrepancies.

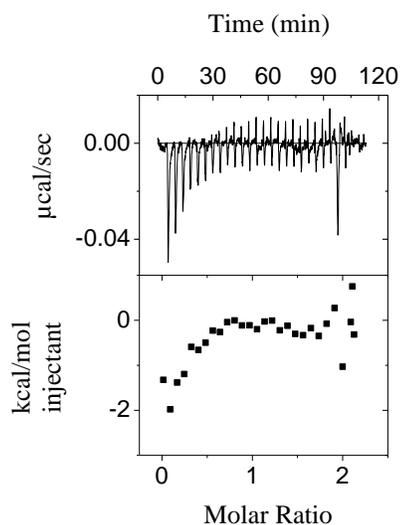


Figure A4.4: ITC titration of bdSUMO-nCoV-S into WW1-4. ITC titration of 147 μM nCoV-S into 13.6 μM WW1-4 was performed at 25 $^{\circ}\text{C}$ in buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol.

Binding Assays of HBC149 and WW1-4 show no complex formation. Having shown that a 20-residue PPAY peptide of Hepatitis B Core (HBC) binds WW1-4 with a K_d of 49 μM (Table A4.1), we next wanted to characterize the interaction in context of a larger construct of HBC, as the peptide may not include all residues involved in binding. We generated a construct of HBC spanning residues 1-149, which contains the PPAY motif and is sufficient to form native capsid particles of 240 monomeric units used for viral particle assembly and genome encapsulation[230]. HBC149 is an obligate dimer, but mid-low pH, mid-high temperature, and high salt concentration favor capsid assembly[231, 232]. To ensure our construct behaves as previously described in the literature, we performed SEC-MALS to determine the oligomerization state. Under dimerization conditions (pH 9.5, no salt), we observed a single peak corresponding to the molecular mass of a dimer (Fig. A4.5A, Table A4.3). Following capsid formation, we observed an elution profile showing two peaks corresponding to the dimer and capsid, respectively (Fig. A4.5B, Table A4.3), indicating partial capsid assembly. The peak corresponding to capsid had an average molar mass of 4,761.9

kDa, close to the expected mass of 4,776 kDa. Temperature appeared to have no effect on the elution profiles, and subsequent experiments were performed at 25 °C.

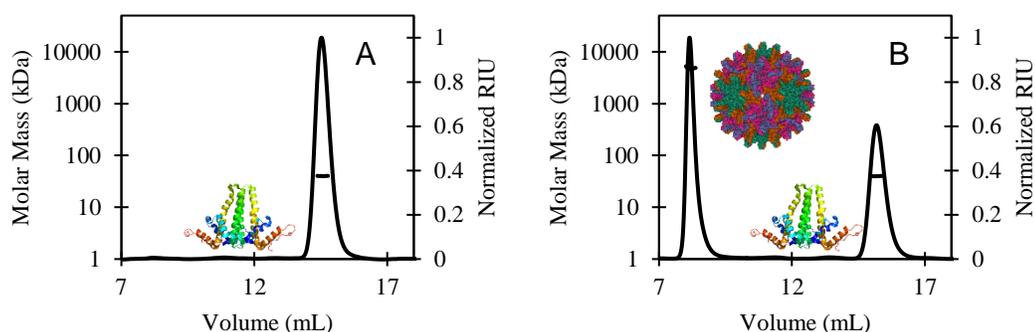


Figure A4.5: Oligomerization state of HBC149. SEC-MALS elution profiles of (A) HBC149 dimer, and (B) HBC149 capsid/dimer mixture after capsid assembly. Dimer (PDB: 3KXS) and capsid (PDB: 6VZP) structures are shown.

Table A4.3: HBC149 oligomerization state determined by SEC-MALS. ^a Value reported was obtained from a single experiment. All other values were obtained from duplicate experiments. Room temperature: r.t.

pH	Temperature (°C)	Experimental MW (kDa)	Theoretical Dimer MW (kDa)
7.5	r.t.	36.7 ^a	39.8
7.5	r.t.	34.4	
7.5	iced	37.4	
9.5	r.t.	40.7	
9.5	iced	38.9	

Having successfully validated dimer and capsid formation of HBC149, we performed SEC-MALS of both HBC149 alone and with excess WW1-4 (Fig. A4.6A). Apo HBC149 ran as a dimer with a molar mass of 39.2 kDa. The HBC149 and WW1-4 mixture eluted as two distinct peaks with molar masses of 37.5 and 38.6 kDa, which correspond to the HBC149 dimer and WW1-4 monomer, respectively. To see if the interaction was detectable at higher concentrations, next performed an ITC titration of WW1-4 into HBC149 dimer in a no salt buffer (Fig. A4.6B). To our surprise, only very weak heats were observed, suggesting that the proteins do not form a complex under these experimental conditions. To see if WW1-4 only binds HBC149 in the capsid form, we performed an ITC titration of HBC149 capsids into WW1-4 (Fig. A4.6C). Again, only weak heats were observed, indicating that WW1-4 does not interact with

HBC149 capsids. Altogether, this data indicates that the multivalent WW domains of NEDD4-1 do not interact with HBC149 dimers or capsids.

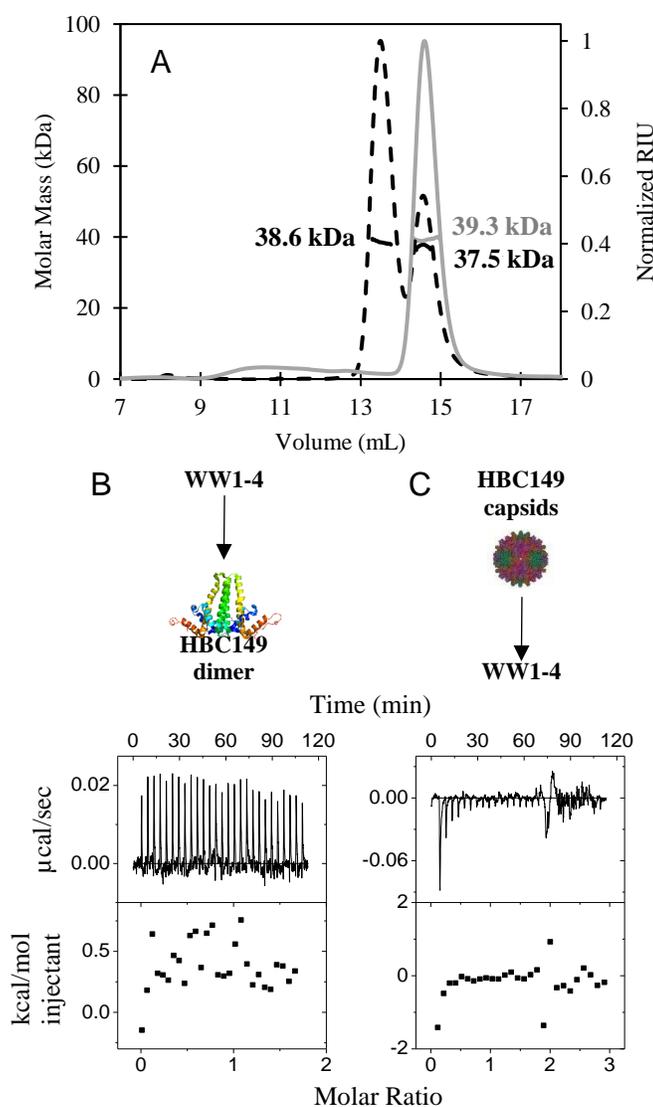


Figure A4.6: SEC-MALS and ITC studies show no binding of HBC149 and WW1-4. (A) SEC-MALS of apo HBC149 dimer (grey, solid line) and HBC149 mixed with WW1-4 (black, dashed line) in buffer composed of 50 mM sodium phosphate, 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol at pH 7.5. (B) Isotherm of 108 μM WW1-4 titrated into HBC149 dimer at 13.6 μM (concentration of monomer) in buffer composed of 50 mM sodium phosphate, 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol at pH 7.5. (C) Isotherm of HBC149 capsids at 343 μM (concentration of monomer) titrated into 25 μM WW1-4 in buffer composed of 50 mM sodium phosphate, 150 mM NaCl , 0.5 mM NaN_3 , and 5 mM β -mercaptoethanol at pH 7.5.

Binding Assays of HBC149 and KWW_{TD} show no complex formation. Having characterized the KWW_{TD} (tandem WW domains of KIBRA) interaction with HBC peptide, we next wanted to validate binding in context of the more biologically relevant HBC dimer. Thus, we first performed an SEC-MALS experiment with a sample of HBC149 and KWW_{TD} combined at a molar ratio of 1:2. To our surprise, we saw two distinct peaks in the elution profile, with molar masses of 37.1 and 11.0 kDa, which are close to the masses of HBC149 dimer and KWW_{TD}, respectively (Fig. A4.7A). Only a

small peak corresponding to capsid was observed under these conditions. This data shows that HBC149 dimer and KWW_{TD} do not bind when diluted over SEC-MALS. Since it is common for weakly associated complexes to dissociate over SEC, we also ran a basic native gel with samples of apo HBC149, apo KWW_{TD}, and the combined proteins under different conditions (Fig. A4.7B). The bands corresponding to HBC149 and KWW_{TD} ran the same when combined as when prepared separately, indicating no complex formation.

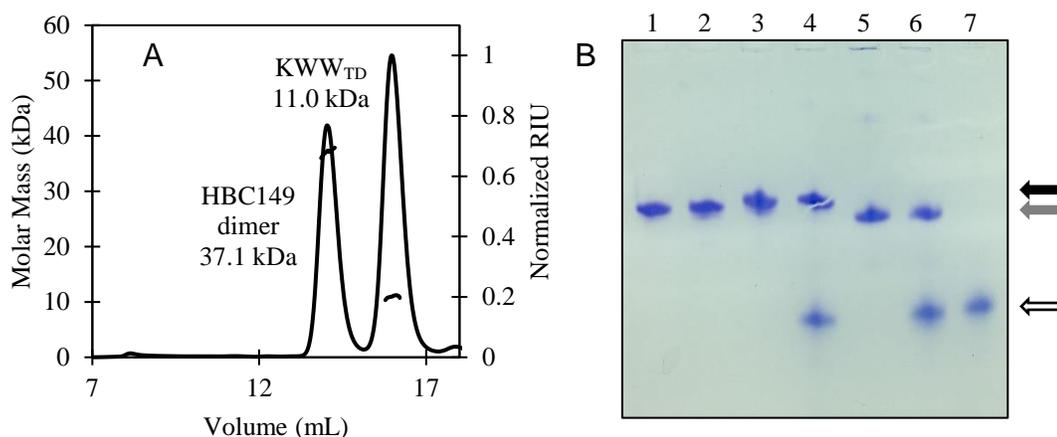


Figure A4.7: SEC-MALS and native gel electrophoresis show no binding of HBC149 and KWW_{TD}. (A) SEC-MALS elution profile of 30 μ M HBC149 and 60 μ M KWW_{TD} in buffer composed of 50 mM HEPES, 1 mM NaN₃, and 5 mM BME at pH 7.5 at room temperature. (B) 12% basic native gel run overnight at 5 mAmps showing (lane 1) His_{6x}-HBC149 in 100 mM sodium carbonate, pH 9.5; (lane 2) His_{6x}-HBC149 in 50 mM HEPES, pH 7.5; (lane 3) His_{6x}-HBC149 in 50 mM sodium phosphate, pH 7.5; (lane 4) His_{6x}-HBC149 + KWW_{TD} in 50 mM sodium phosphate, pH 7.5; (lane 5) HBC149 in 50 mM sodium phosphate, pH 7.5; (lane 6) HBC149 + KWW_{TD} in 50 mM sodium phosphate, pH 7.5; and (lane 7) KWW_{TD} in 50 mM sodium phosphate, pH 7.5. Protein bands corresponding to His_{6x}-HBC149, HBC149, and KWW_{TD} are indicated with black, grey, and white arrows, respectively. All samples were prepared at concentrations of 20-40 μ M.

To see if HBC149 and KWW_{TD} bind at higher concentrations, we performed ITC titrations of KWW_{TD} into HBC149 dimer in different buffers and sample preparation conditions (Fig. A4.8A-C). These titration experiments showed little heat and no indication of binding. To see if KWW_{TD} binds HBC149 capsid, we performed the inverse ITC titration of HBC149 capsids into KWW_{TD}, which also resulted in no

apparent binding (Fig. A4.8D). To rule out a binding event with zero enthalpic contributions, we also added unlabeled HBC149 to ^{15}N -KWW_{TD} and monitored peak changes (Fig. A4.8E). A few weak peaks showed small chemical shift perturbations, which are likely the result of slight changes in buffer composition or sample concentration and not due to binding. In summary, SEC-MALS, native gel electrophoresis, ITC, and NMR experiments all show no binding between HBC149 and KWW_{TD}.

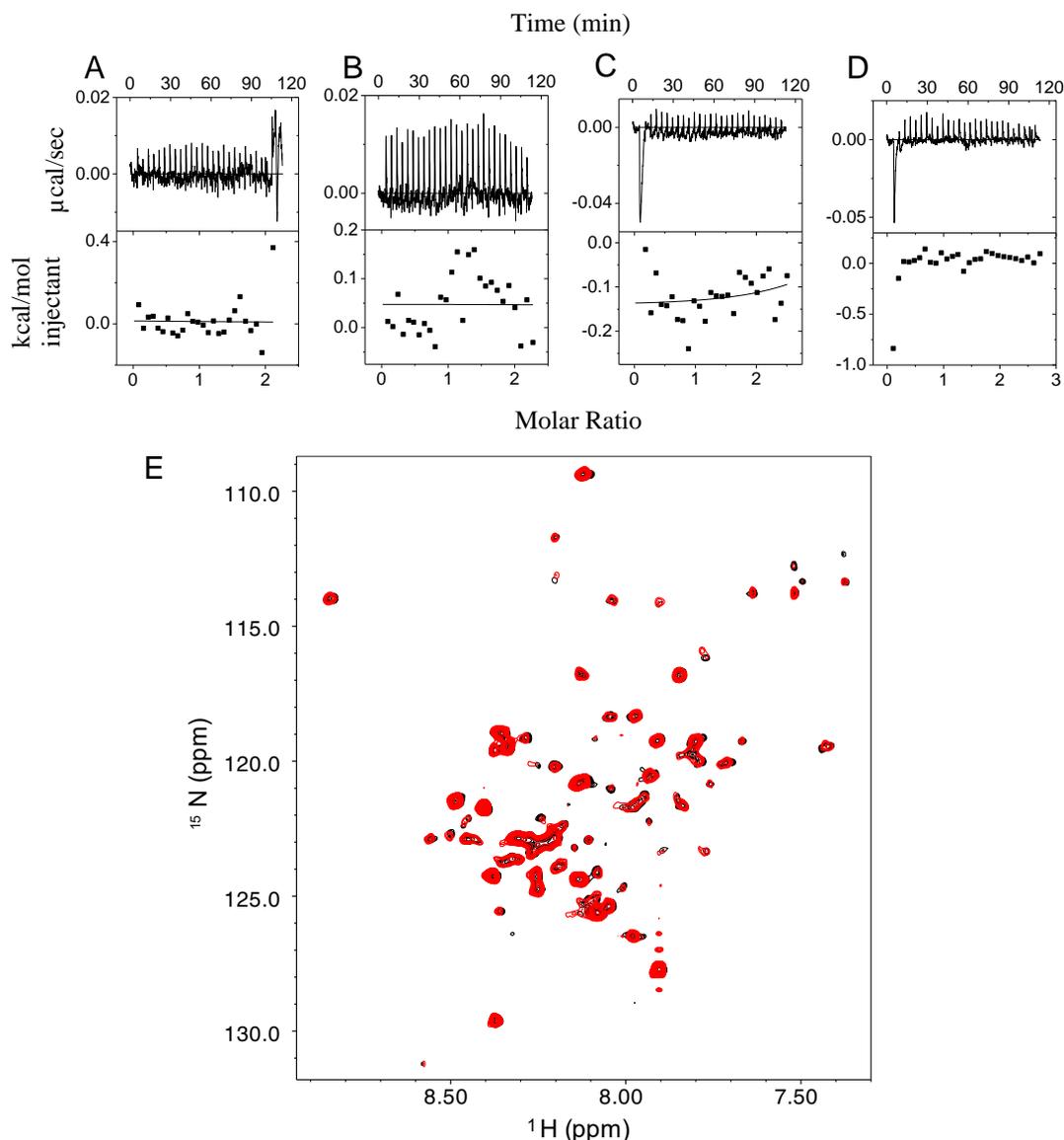


Figure A4.8: ITC and NMR studies of HBC149 and KWW_{TD} show no binding. Isotherms of (A) $302 \mu\text{M}$ KWW_{TD} titrated into $30 \mu\text{M}$ His_{6x}-HBC149 (uncut) dimer in buffer composed of 50 mM HEPES, pH 7.5, 0.5 mM NaN₃, 5 mM β -mercaptoethanol (BME); (B) $329 \mu\text{M}$ KWW_{TD} titrated into $30 \mu\text{M}$ His_{6x}-HBC149 (uncut) dimer in buffer composed of 50 mM sodium phosphate, pH 7.5, 0.5 mM NaN₃, 5 mM BME; (C) $341 \mu\text{M}$ KWW_{TD} titrated into $29 \mu\text{M}$ HBC149 (cut) dimer in buffer composed of 50 mM sodium phosphate, pH 7.5, 0.5 mM NaN₃, 5 mM BME; and (D) $343 \mu\text{M}$ HBC149 capsids titrated into $25 \mu\text{M}$ KWW_{TD} in buffer composed of 50 mM sodium phosphate, pH 7.5, 150 mM NaCl, 0.5 mM NaN₃, 5 mM BME. HBC149 concentrations are based on the extinction coefficient of the monomer. (E) ^1H - ^{15}N -HSQC of apo ^{15}N -KWW_{TD} (black) overlaid with ^{15}N -KWW_{TD} with added unlabeled HBC149 (red). All experiments were performed at 25 °C.

Conclusions

Viral L domain-host protein interactions regulate viral particle escape from the cell[138]. Elucidating the viral L domain interactome and the mechanisms by which host-virus complexes assemble will uncover potential avenues for antiviral drugs. In this study, we performed a low-throughput ITC binding screen with four viral L domain peptides and three human WW domain proteins to uncover novel interactions. Using 20-residue peptides, we validated the proposed interaction of SARS-CoV-2 with NEDD4-1 and showed potential interactions with YAP and KIBRA. We also identified novel interactions between an HBV Core peptide and KIBRA. However, when we generated larger polypeptides containing these L domains, no evidence of binding was observed. This work emphasizes the importance of sequence context in characterizing WW domain-PPXY motif interactions.

Future work

nCoV-S and NEDD4-1. Recent work by Novelli. G. et al showed that the PPAY motif of nCoV Spike protein interacts with the WW domains of NEDD4, leading to its ubiquitination[142]. However, this interaction remains to be validated with biophysical experiments. In this study, we performed binding assays using 20- and 70-residue polypeptides of the nCoV-S protein containing the PPAY motif. Binding of the shorter peptide to NEDD4-1 WW3 produced the tightest binding affinity (K_d of 46 μ M). However, binding was not recapitulated with the larger construct. Future studies may consider testing binding of full-length nCoV-S to NEDD4-1 to see if binding occurs in a native sequence context. Additional factors which may impact binding include post-translational modifications and adaptor proteins.

HBC and NEDD4-1. Several previous studies have shown that HBC interacts through its PPAY motif with the WW domains of NEDD4, leading to HBC ubiquitination, which is likely an important step in viral egress[224, 225, 233]. However, the direct interaction between HBC and NEDD4 has never been characterized biophysically. In this study, we show that a 20-residue peptide of the HBC PPAY motif interacts weakly with NEDD4-1, but when put in context of the native HBC dimer or capsid, no binding was observed. Structural examination revealed

that the Tyr of the PPAY motif is buried in the capsid structure, which would inhibit binding to WW domains[234]. However, in the dimer, the PPAY motif remains solvent accessible (Fig. A4.9). Thus, it remains unclear why the motif is inactive in the dimer form. Future work should consider the use of full-length HBC, which includes a C-terminal Arg-rich domain, and investigation into potential adaptor proteins that may link HBC with NEDD4.

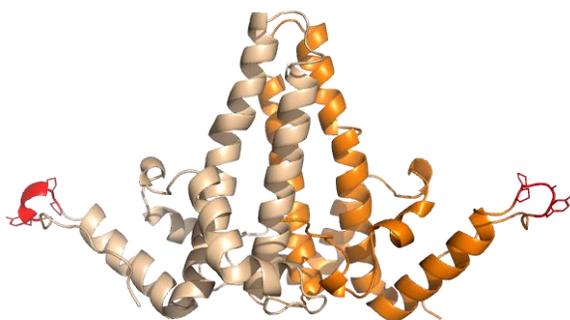


Figure A4.9: Structure of HBC149 dimer. Monomeric units are shown in wheat and orange. PPAY motifs are shown in red. Structure rendered from PDB: 1QGT.

KIBRA WW domain dynamics. In this study, we examined the interaction between the tandem WW domains of KIBRA and a PPAY peptide of HBC. Although the interaction was not recapitulated in context of the HBC149 dimer, preliminary dynamics studies revealed that KWW_{TD} likely undergoes intermediate chemical exchange. These exchange processes appeared to be modulated by PPXY interactions. In Chapter 3, we show cooperative binding of the WW domains of KIBRA, even when interactions are mediated by only one PPXY motif. These findings leave open questions as to how the WW domains cooperate. Perhaps WW domain dynamics play a role in partner recognition and cooperative binding. Future studies may include a thorough investigation into the dynamics of KWW_{TD} with high quality data in the apo form and in the presence of biologically relevant binding partners.

Materials and Methods

Viral peptide synthesis. Twenty-residue synthetic peptides containing the PPXY motif and eight flanking residues from four HTLV-1 Gag protein (Uniprot ID: P14078, residues 110-129), SARS-CoV Spike protein (Uniprot ID: P59594, residues 52-71), SARS-CoV-2 Spike protein (Uniprot ID: P0DTC2, residues 17-36), and HBV Core protein (Uniprot ID: Q89446, residues 150-169) were purchased from Synpeptide

Co., Ltd (Shanghai, China) at > 90% purity. Peptide concentrations were determined from the absorbance at 280 nm calculated from the protein sequence (<http://web.expasy.org/protparam/>).

Cloning of construct. Cloning of residues 157-277 of YAP, (YWW_{TD}), residues 1-91 of KIBRA (KWW_{TD}), and residues 600-935 of NEDD4 (WW1-4) into pET24 expression vectors is described in Appendix 1. For this work, three additional constructs were generated. Residues 835-878 of NEDD4 (Uniprot ID: P46934, Genscript) spanning the WW3 domain, and residues 1-70 of SARS-CoV2 Spike protein (nCoV-S) were cloned into BamHI and XhoI cut sites of a modified pET vector, pRBC[218]. The pRBC vector contained an N-terminal 6x-Histidine tag (His_{6x}), followed by a *Brachypodium distachyon* Small Ubiquitin-like Modifier (bdSUMO) tag and a *Brachypodium distachyon* Sentrin-specific protease 1 (bdSEN1) cleavage site[219]. SARS-CoV2 Spike DNA was amplified from DTwist-EF1 α -nCoV-S-2xStrep gifted by the Dr. Elisar Barbar (Department of Biochemistry and Biophysics, OSU). A construct of Hepatitis B Core protein (Uniprot ID: Q89446, Genscript) spanning the capsid assembly domain, (residues 1-149, HBC149) was cloned into NdeI and XhoI cut sites of pET24 with an N-terminal His_{6x} tag followed by a TEV protease cut site. All constructs were cloned into their respective expression vectors using the Gibson AssemblyTM cloning protocol (New England Biolabs, MA).

KWW1 and KWW2 production. Cloning and purification of KIBRA constructs containing WW1 (residues 1– 48, KWW1), and WW2 (residues 45–91, KWW2) are described elsewhere[84].

Production of YWW_{TD}, KWW_{TD}, and WW1-4. Expression and purification are described in Appendix 1.

Production and purification of NWW3. BL21-DE3 *E. coli* cells transformed with bdSUMO-NWW3 in pRBC were grown in LB supplemented with 0.2% glucose at 37 °C. Upon reaching an OD₆₀₀ of 0.6-0.7, cultures were induced with 1 mM IPTG for 3 h at 37 °C and harvested by centrifugation at 4 krpm for 25 min. Pelleted cells were resuspended in buffer composed of 20 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole, 5% v/v glycerol, and 1 mM NaN₃ at pH 8.0 (Buffer B) and lysed by sonication, followed by centrifugation at 15 krpm for 30 min to remove cell debris.

Cleared lysate was incubated with Ni-NTA resin for 30 min, 4 °C, 100 rpm on an Innova 200 platform shaker (New Brunswick™) to allow His_{6x}-tagged NWW3 to bind, followed by washes of Buffer B containing 10-20 mM imidazole. Washed Ni-NTA resin was resuspended in 5 mL Buffer B and incubated with 30 nM bdSENP1 for 1 h at 4 °C, 100 rpm to cleave off the His_{6x}-bdSUMO tag. Cleaved NWW3 was eluted in Buffer B and further purified by SEC on a Superdex⁷⁵ or Superdex²⁰⁰ (Cytiva life sciences) column in buffer composed of 50 mM sodium phosphate, 100 mM NaCl, 1 mM NaN₃, and 5 mM β-mercaptoethanol (BME).

Production and purification of nCoV-S. BL21-DE3 *E. coli* cells transformed with bdSUMO-nCoV-S in pRBC were grown in LB supplemented with 0.2% glucose at 37 °C. Upon reaching an OD₆₀₀ of 0.6-0.7, cultures were induced with 0.1 mM IPTG at 20 °C overnight and harvested by centrifugation at 4 krpm for 25 min. Overexpressed bdSUMO-nCoV-S went into inclusion bodies, which were isolated as follows. Pelleted cells from a 1 L culture were resuspended and incubated at room temperature for 30 in 30 mL buffer containing 50 mM Tris, 500 mM NaCl, 2 mg/mL lysozyme, and 2 mM EDTA at pH 8.0, followed by lysis by sonication and centrifugation at 15 krpm for 30 min. The supernatant was discarded. The following step was repeated twice: the pellet was resuspended in 40 mL buffer containing 50 mM Tris, 500 mM NaCl, 1% Triton X-100 at pH 8.0, followed by centrifugation at 15 krpm for 10 min. The supernatant was discarded after each wash. Next, the inclusion body pellet was resuspended in 6 mL buffer containing 50 mM Tris, and 500 mM NaCl at pH 8.0, transferred to 1.6 mL Eppendorf Tubes®, and centrifuged at 16 krpm for 5 min. The supernatant was aspirated off. The pellet in each Eppendorf Tube® was resuspended in 1 mL buffer composed of 50 mM Tris, and 500 mM NaCl at pH 8.0, followed by centrifugation and aspiration as described in the previous step. Inclusion bodies were stored at -20 °C.

Insoluble bdSUMO-nCoV-S in isolated inclusion bodies was purified as follows. Each 1.6 mL aliquot (described above) was resuspended in 5 mL buffer composed of 50 mM Tris, 6 M guanidine-HCl, 200 mM NaCl, 10 mM BME, and 1 mM NaN₃ at pH 8.0 and rocked for 1 h at 4 °C, followed by centrifugation at 10 krpm for 10 min. Denatured protein in the supernatant was added dropwise (1 drop/5-10 s) to 200 mL refolding buffer composed of 100 mM Tris, 0.4 M arginine, 5 mM BME,

200 mM NaCl, 5% glycerol, and 1 mM NaN₃ at pH 8.0 at 4 °C while gently stirring. After overnight stirring at 4 °C, 3 mL Ni-NTA resin pre-equilibrated with refolding buffer was added to the refolded protein and stirred for 1 h at 4 °C. Ni-NTA resin was washed with 35 mL buffer containing 5 mM Tris, 500 mM NaCl, 10 mM imidazole, 1 mM NaN₃, and 5 mM BME at pH 8.0, followed by elution in 30 mL of the same buffer, but with 350 mM imidazole. For improved solubility, the bdSUMO tag was not removed. nCoV-S was further purified by SEC on a Superdex⁷⁵ or Superdex²⁰⁰ (Cytiva life sciences) column in buffer composed of 50 mM sodium phosphate, 100 mM NaCl, 1 mM NaN₃, and 5 mM BME.

Production and purification of HBC149. BL21-DE3 *E. coli* cells transformed with HBC149 in pET24 were grown in TY media supplemented with 0.2% glucose at 37 °C. Upon reaching an OD₆₀₀ of 0.6-0.7, cultures were induced with 0.1 mM IPTG at 37 °C overnight and harvested by centrifugation at 4 krpm for 25 min. Pelleted cells were resuspended in buffer composed of 50 mM sodium carbonate, 50 mM NaCl, 2 M urea, 2 mM dithiothreitol (DTT), 5 mM imidazole, 5% glycerol, 1 mM NaN₃, and 1 mM PMSF at pH 9.5 (Buffer C), lysed by sonication, and centrifuged at 15 krpm for 30 min. Cleared lysate was incubated with Ni-NTA resin for 30 min, 4 °C, 100 rpm to allow His_{6x}-tagged HBC149 to bind, followed by washes of Buffer C containing 10-20 mM imidazole and elution in buffer composed of 50 mM Tris, 50 mM NaCl, 2 mM DTT, 5% glycerol, and 350 mM imidazole at pH 8.0. To remove the His_{6x} affinity tag, HBC149 was incubated overnight with 0.5 μM TEV protease and dialyzed against 2 L buffer composed of 50 mM Tris, 10 mM NaCl, 5 mM BME, and 1 mM EDTA at pH 8.0. To removed uncut HBC149, protein was bound to Ni-NTA resin and cleaved protein was eluted with 50 mM Tris, 50 mM NaCl, 10 mM imidazole, 5% glycerol, and 1 mM NaN₃ at pH 8.0. HBC149 was further purified by SEC on a Superdex²⁰⁰ (Cytiva life sciences) column in buffer composed 100 mM sodium carbonate, 1 mM NaN₃, 5 mM BME at pH 9.5.

To assemble capsids, HBC149 was dialyzed against buffer composed of 50 mM sodium phosphate, 250 mM NaCl, 1 mM NaN₃, and 5 mM BME overnight at room temperature. Following dialysis, unassembled HBC149 dimers were removed by SEC

on a Superdex²⁰⁰ (Cytiva life sciences) column in buffer composed of 50 mM sodium phosphate, 400 mM NaCl, 1 mM NaN₃, and 5 mM BME.

SEC-MALS. Absolute molar masses and elution profiles were obtained from size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) experiments using an AKTA FPLC (GE Healthcare) a DAWN multiple-angle light scattering, and Optilab refractive index system (Wyatt Technology). Data were collected on a Superdex²⁰⁰ 10/300 GL column (Cytiva life sciences) pre-equilibrated with ice cold or room temperature 50 mM sodium phosphate, 250 mM NaCl, 1 mM NaN₃, and 5 mM BME at pH 7.5 buffer or 100 mM sodium carbonate, 1 mM NaN₃, and 5 mM BME at pH 9.5. Volumes of 100 μ L HBC149 samples at 20-30 μ M were injected 0.8 mL/min. The data were analyzed with ASTRA software package, version 8 (Wyatt Technology).

Isothermal titration calorimetry (ITC). ITC experimental data were collected on a VP-ITC instrument (Malvern instruments Inc, MA) at 25 °C. For peptide binding experiments, 28 injections of 124-770 μ M peptide was titrated into 12-30 μ M WW domain proteins in buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 5 mM BME, and 0.1 mM NaN₃ at pH 7.5. Conditions and sample concentrations for ITC titrations are described in the main text and figure captions. Isotherms were fit using Origin 7.0 software to a single-site binding model which produces the average dissociation constant (K_d), a measure of complex stability, and the binding stoichiometry (N) which gives the average molar ratio of the complex. The free energy of binding (ΔG) was calculated from the equation $\Delta G = -RT\ln(K_a)$, where R is the universal gas constant, T is temperature in Kelvin, and K_a is the association constant.

NMR spectroscopy titration experiments. NMR experiments were conducted using a Bruker Avance III, 800 MHz spectrometer (Bruker Biospin) equipped with a triple resonance (HCN) cryogenic probe. Titration of unlabeled HBC peptide into ¹⁵N-KWW_{TD} was performed in buffer composed of 10 mM sodium phosphate, 10 mM NaCl, 5 mM tris(2-carboxyethyl)phosphine (TCEP), and 1 mM NaN₃ at pH 6.8 (NMR buffer 1). KWW_{TD} at 120 μ M and HBC peptide were combined at molar ratios of 1:0, 1:0.1, 1:0.2, 1:0.35, 1:0.5, and 1:4 (¹⁵N-labeled KWW_{TD}:HBC peptide). Titration of unlabeled HBC149 dimers into ¹⁵N-labeled KWW_{TD} were performed in buffer

composed of 10 mM sodium phosphate, 3 mM TCEP, and 1 mM NaN_3 at pH 6.8. KWW_{TD} at 100 μM and HBC149 were mixed at molar ratios of 1:0 and 1:0.3 (^{15}N -labeled $\text{KWW}_{\text{TD}}:\text{HBC149}$). All samples contained 10% D_2O as the lock nucleus and 0.5% DSS for internal referencing. NMR titrations were performed using 2D BEST-TROSY ^1H - ^{15}N HSQC experiments collected at 25 $^\circ\text{C}$ [169].

^{15}N -CPMG relaxation dispersion NMR experiments. ^{15}N BEST-TROSY Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion experiments were performed using Bruker Avance III, 800 and 500 MHz spectrometers (Bruker Biospin) equipped with triple resonance (HCN) cryogenic and conventional probes, respectively. Samples were prepared in NMR buffer 1 as described above using 120 μM ^{15}N - KWW_{TD} with and without 0.1 equivalents HBC peptide. Experiments were performed at 25 $^\circ\text{C}$ with a constant delay time of 40 ms using a BEST modified the pulse sequence of Franco et al. [221]. Data were collected with 12 different interleaved refocusing pulse delay values of 0, 25, 50 (x2), 100, 150, 200, 250, 400, 500, 600, 750, and 1000 Hz. Experimental error was calculated from experimental duplicates (x2) using Monte Carlo analysis. CPMG data analysis was performed in Relax using the CR72 two-state exchange equations of Carver and Richards[222, 235]. All NMR data were processed in NMRPipe[170] and visualized in Sparky[171] or NMRView[205].

Appendix 5

Liquid-liquid phase separation studies of the NEDD4-1-AMOTL1 complex

Amber Vogel and Matthew McWhorter

Overview

Liquid-liquid phase separation (LLPS, otherwise referred to as phase transitions or condensate formation) is a complex, reversible biophysical process observed in protein-protein, protein-RNA, and RNA-RNA interaction systems[236]. Upon reaching a critical concentration where the energetics of transient weak interactions outweigh entropic costs, solutes spontaneously demix into a “dilute” phase and a concentrated, gel-like “dense” phase, often referred to as membraneless organelles, biomolecular condensates, or droplets[213-216]. Phase separated droplets provide spatially segregated cellular compartments. Outcomes of LLPS include enhanced oligomerization/complex assembly, sequestration, enhanced enzymatic activity, genomic packaging, buffering, and generation of mechanical force[236, 237]. In some cases, phase separated droplets can age and solidify, forming pathological aggregates, a process enhanced by mutation[213, 236, 238]. Disorders caused/enhanced by aberrant LLPS include cancers[215], infectious disease[216], and neurodegenerative diseases such as Alzheimer’s disease[239], dementia[240], Parkinson’s disease[241], amyotrophic lateral sclerosis (ALS)[242], and Huntington’s disease[243].

Disorder and multivalency are two features commonly found in systems that undergo LLPS (Fig. A5.1)[244]. Although these features are found in many proteins within the Hippo signaling pathway (described in Chapter 1), there are only a few examples of proteins from this pathway participating in LLPS[245-248], and no examples where LLPS is mediated by WW domains or PPXY motifs. Phase separation of Hippo signaling components could serve as a tunable biophysical mechanism of cell growth regulation. Here, we describe preliminary experiments which suggest phase separation occurs within the semi-disordered, multivalent WW domain/PPXY motif regions of NEDD4-1 and AMOTL1. Our findings implicate LLPS in the possible regulation of protein homeostasis and cell growth.

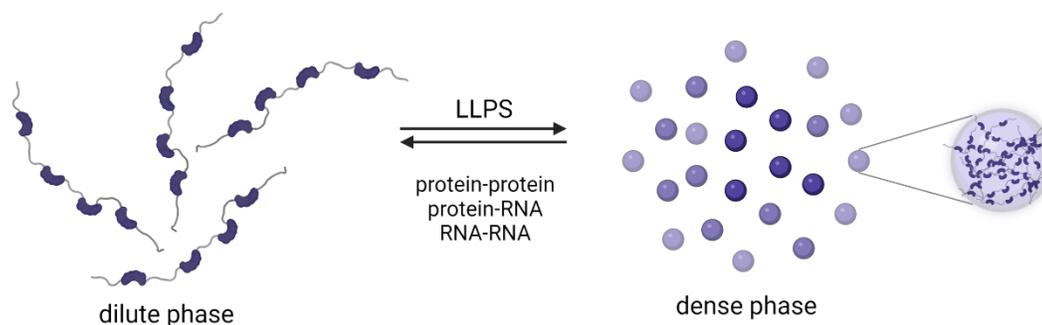


Figure A5.1: Liquid-liquid phase separation diagram. Proteins and nucleic acids, typically with multivalent binding sites, can undergo reversible phase transitions from a dilute phase to a dense phase. Created with BioRender.com

Results

NEDD4-1 and AMOTL1 LLPS prediction. Phase separating proteins are often composed of low complexity regions (LCRs) which are enriched in specific amino acids or repeating sequences, although enriched amino acids are protein-specific[167]. Several algorithms have been developed that predict the ability of polypeptide sequences to undergo LLPS including DeePhase prediction (<https://deepphase.ch.cam.ac.uk/>) and Potential phase separation score (PSPredictor) (<http://www.pkumdl.cn:8000/PSPredictor/>)[249]. Prior to attempting LLPS experiments, we used both predictive algorithms to see if our WW domain construct of NEDD4-1 (residues 600-935, WW1-4) and PPXY motif construct of AMOTL1 (residues 178-384, A₁₂₃) were likely to phase separate based on their sequences. Both DeePhase prediction and PSPredictor gave WW1-4 scores above 0.5, which indicates that WW1-4 is likely to undergo LLPS (Table A5.1). Of these algorithms, only DeePhase prediction gave A₁₂₃ a score above 0.5, indicating that it is less likely to phase separate.

Table A5.1: Phase separation prediction scores.

Protein	DeePhase prediction	PSPredictor
WW1-4	0.82	0.797
A ₁₂₃	0.67	0.401

SEC-MALS shows concentration-dependent elution behavior of WW1-4. LLPS

occurs when solutes reach a critical concentration. To see if WW1-4 exhibits differences in size at different concentrations, we performed SEC-MALS on WW1-4 samples ranging from 20-165 μM . Elution volume decreased linearly with increasing concentration of WW1-4 (Fig. A5.2). The experimental molar masses of all samples ranged from 38.1-42.1 kDa, close to the molecular weight of a WW1-4 monomer (Table A5.2). This data suggests that WW1-4 forms higher order oligomerization states that dissociates over SEC-MALS, resulting in a monomeric molar mass but earlier elution time.

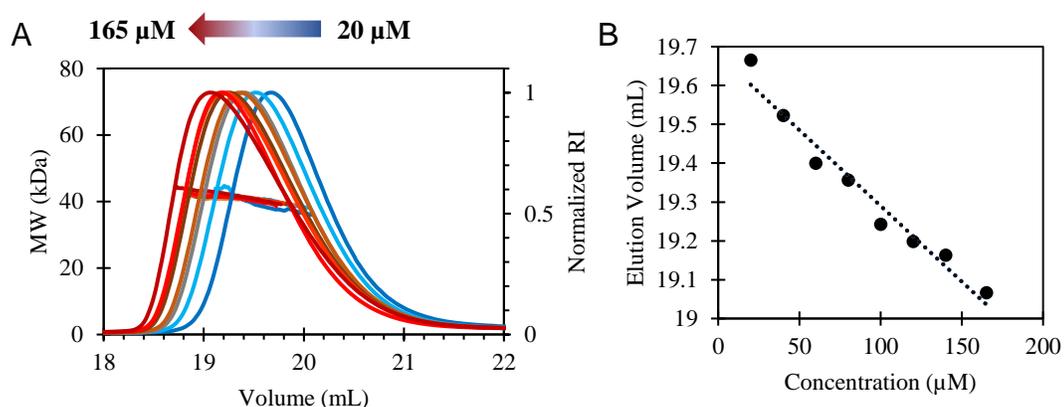


Figure A5.2: SEC-MALS of WW1-4. (A) SEC-MALS of WW1-4 at concentrations of 20 (dark blue), 40 (light blue), 60 (grey), 80 (dark orange), 100 (dark brown), 120 (light orange), 140 (light red), and 165 (dark red) μM . (B) Elution volume at maximum refractive index plotted versus WW1-4 concentration fit with a linear curve. R^2 value = 0.96.

Table A5.2: SEC-MALS of WW1-4 at increasing concentrations. Reported values were obtained from single experiments.

Concentration (μM)	Empirical Molar Mass (kDa, monomer)	Experimental Molar Mass (kDa)
20	38.3	38.1
40		40.7
60		40.2
80		40.4
100		40.8
120		41.3
140		41
165		42.1

Turbidity of WW1-4. Qualitative observation during various purifications and experiments revealed that WW1-4 becomes turbid when at high concentrations or in high salt buffers. Turbidity can be quantified using optical density measurements at 340 or 600 nm[250, 251]. Thus, we measured the A_{340} of WW1-4 at concentrations ranging from 10-192 μM in buffers containing 50 mM or 400 mM NaCl (Fig. A5.3A). At lower concentrations, the A_{340} values were low in both conditions. However, upon reaching concentration thresholds of 70 μM (400 mM NaCl) or 170 μM (50 mM NaCl), the A_{340} began to increase nonlinearly. These data suggest that WW1-4 undergoes LLPS in a concentration and salt-dependent manner. Furthermore, visual inspection of a WW1-4 sample mixed with A_{123} and the tandem WW domains of KIBRA (KWW_{TD} , described in Chapter 2 and Appendix 1) under a Nikon microscope showed the presence of droplets (Fig. A5.3B). This prompted us to examine the phase separation behavior of WW1-4 in the presence of A_{123} . Since KWW_{TD} is not a binding partner of WW1-4, this protein was not included in subsequent experiments.

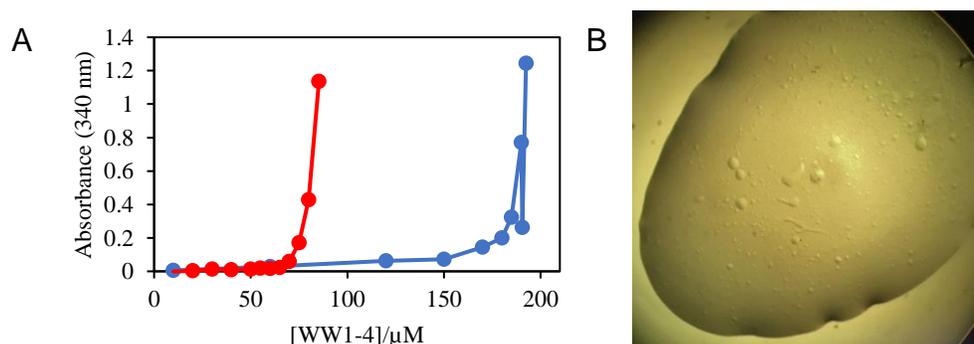


Figure A5.3: Turbidity of WW1-4. (A) Absorbance measured at 340 nm plotted versus WW1-4 concentration. Data were collected in pH 7.5 buffer composed of 50 mM sodium phosphate, 1 mM NaN_3 , and 5 mM β -mercaptoethanol containing 50 mM NaCl (blue scatter plot) or 400 mM NaCl (red scatter plot). Connecting lines are included to guide the eye. Absorbances were measured in a quartz cuvette using an Agilent Technologies Cary 60 UV-Vis Spectrophotometer. (B) Image of WW1-4 mixed with A_{123} and KWW_{TD} taken under a Nikon SMZ1500 microscope.

Preliminary fluorescence confocal microscopy shows the A_{123} -WW1-4 complex forms droplets. To see if WW1-4 and A_{123} undergo colocalized phase separation when mixed, we used the Zeiss LSM 780 NLO confocal microscope system to image fluorescently labeled mixtures of WW1-4 and A_{123} . Conveniently, both constructs

contain one native cysteine residue. This allowed us to fluorescently label each protein with either Alexa488 or Alexa350 maleimide dye using cysteine crosslinking. Figure A5.4 shows confocal microscopy images of fluorescently labeled WW1-4, A₁₂₃, and a mixture of the two proteins. Images of apo WW1-4 and A₁₂₃ at ~10 μM in buffer containing 50 mM NaCl showed diffuse fluorescence expected for soluble proteins in solution (Fig. A5.4A-B). Next, WW1-4 and A₁₂₃ were mixed, immediately transferred to cover glass slides, and imaged. Small droplets were observed (Fig. 5.4C). Next, we prepared a ~1:1 mixture of fluorescently labeled A₁₂₃:WW1-4 in buffer containing 150 mM NaCl. Approximately 1 hour after mixing, samples were imaged (Fig. A5.4D). In this experiment, large spherical droplets composed of both A₁₂₃ and WW1-4 were observed. In summary, mixtures of fluorescently labeled A₁₂₃ and WW1-4 appeared to form phase separated droplets, and droplet size appeared to be dependent on salt concentration and/or incubation time after sample preparation.

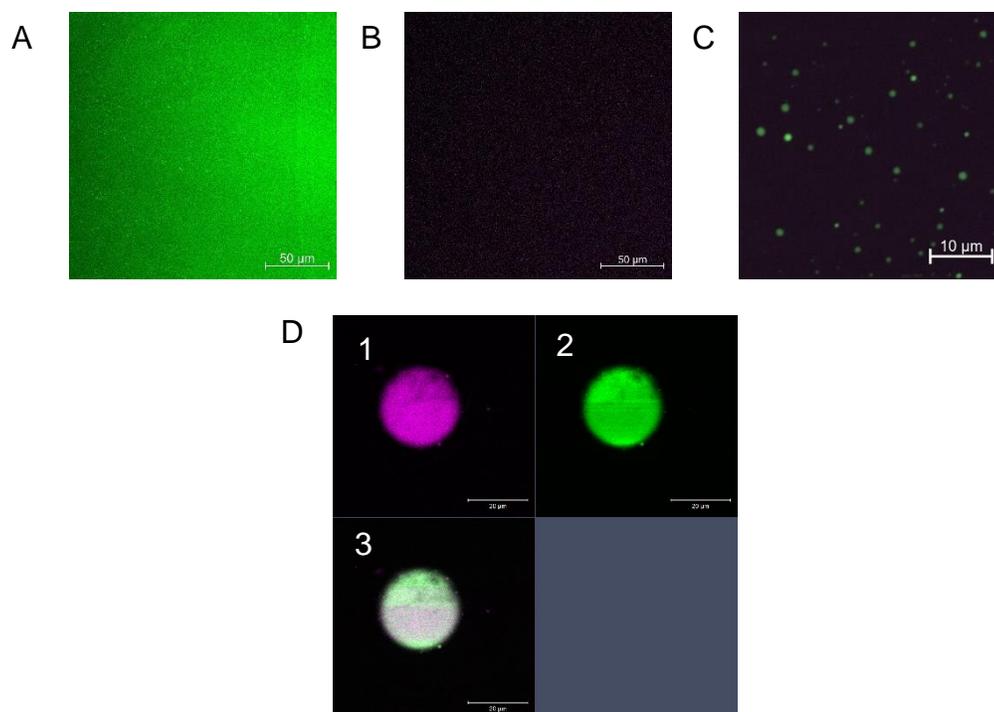


Figure A5.4: Confocal microscopy shows droplet formation of WW1-4 mixed with A123. Confocal microscopy images of (A) Alexa488-WW1-4, (B) Alexa350-A₁₂₃, and (C) a mixture of Alexa488-WW1-4 and Alexa350-A₁₂₃. Samples were prepared in buffer composed of 50 mM sodium phosphate, 50 mM NaCl, and 0.5 mM NaN₃ at pH 7.5. (C) A mixture of Alexa488-WW1-4 and Alexa350-A₁₂₃ showing 1) the Alexa350 channel (A₁₂₃), 2) the Alexa488 channel (WW1-4), and 3) both Alexa350 and Alexa488 channels overlayed. Samples were prepared in buffer composed of 50 mM sodium phosphate, 150 mM NaCl, and 0.5 mM NaN₃ at pH 7.5. Protein concentrations were estimated at ~10 μM.

Conclusions

LLPS is a spatiotemporal compartmentalization strategy used by proteins and nucleic acids to accomplish numerous cell functions include enhancing enzymatic reactions[252]. Phase separation of ubiquitin ligases such as NEDD4-1 could provide a concentration and/or condition-dependent mechanism for substrate ubiquitination and protein homeostasis. To our knowledge, NEDD4-1 has not been shown to phase separate. In this work, we conduct preliminary SEC-MALS, turbidity, and confocal microscopy experiments that show the WW domains of NEDD4-1 may undergo LLPS. This process appears to depend on protein concentration, ionic strength, and the presence of the NEDD4-1 substrate, AMOTL1. Future experiments such as

fluorescence recovery after photobleaching and droplet fission/fusion assays are needed to validate LLPS of these proteins[253].

Materials and methods

Cloning of constructs and recombinant protein production are described in Appendix 1.

Size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). Absolute molecular weights and elution volumes were obtained by size exclusion chromatography-coupled multi-angle light scattering (SEC-MALS) using an AKTA FPLC (GE Healthcare), a DAWN multiple-angle light scattering, and Optilab refractive index system (Wyatt Technology). Data were collected on a Superdex²⁰⁰ 10/300 GL column (Cytiva life sciences) pre-equilibrated with 50 mM sodium phosphate, 50 mM NaCl, 1 mM NaN₃, and 5 mM β-mercaptoethanol at pH 7.5 buffer at room temperature. Protein samples at concentrations of 20-165 μM were injected at a flow rate of 0.7 mL/min. The data were analyzed with ASTRA software package, version 8 (Wyatt Technology).

Turbidity Assays. WW1-4 in buffer composed of 50 mM sodium phosphate, 1 mM NaN₃, and 5 mM β-mercaptoethanol containing 50 or 400 mM NaCl was concentrated to 85 or 190 μM, respectively. Serial dilutions were performed to obtain a range of protein concentrations down to 10 μM. Absorbances were measured at 340 nm in a quartz cuvette using an Agilent Technologies Cary 60 UV-Vis Spectrophotometer.

Alexa Maleimide Dye Labeling Reactions. Alexa350 or Alexa488 (ThermoFisher Scientific, Waltham, MA) stock solutions were prepared at 10 mM in DMSO. Protein-labeled reactions were performed at a 100 μL scale. Proteins were dialyzed against 50 mM sodium phosphate, 100 mM NaCl, and 5 mM tris(2-carboxyethyl)phosphine at pH 7.5. Fifty μM A₁₂₃ and 100 μM WW1-4 were mixed with 10-fold molar excess Alexa350 or Alexa488 maleimide dye, respectively for 2 hr at room temperature in the dark. Reactions were quenched with addition of 5 mM β-mercaptoethanol, and excess dye was removed using a HiTrapTM Desalting column (GE Healthcare) equilibrated with buffer containing 50 mM sodium phosphate, 0.5 mM

NaN₃, and 50 or 150 mM NaCl at pH 7.5. Fluorescently labeled protein concentrations were estimated at ~10 μM based on gel intensities and absorbance measured at 280 nm.

Confocal Microscopy. Protein sample volumes of 10 μL were transferred to 18 x 18 mm Corning Cover Glass slides for imaging. Confocal fluorescence images were obtained using a Zeiss LSM 780 NLO inverted confocal microscope using a 40x water objective and an Argon laser. Laser excitation wavelengths of 405 nm and 488 nm were used for Alexa350 and Alexa488-labeled proteins, respectively. This work was made possible by the Center for Quantitative Life Sciences (CQLS), the Confocal Microscope Facility, and the National Science Foundation (award number 1337774).

Appendix 6

What factors govern AMOTL1 motif specificity?

Amber Vogel and Alexandra Crawford

Overview

In previous chapters of this thesis, we show that all PPXY motifs of AMOTL1 are important for complex assembly with YAP, KIBRA, and NEDD4-1. In each interaction, certain motifs bind with higher affinity than others (Fig. A6.1). However, it remains unclear what drives these motif preferences. In this appendix, we discuss potential features which may govern the differential binding affinities of YAP, KIBRA, and NEDD4-1 for the different PPXY motifs of AMOTL1.

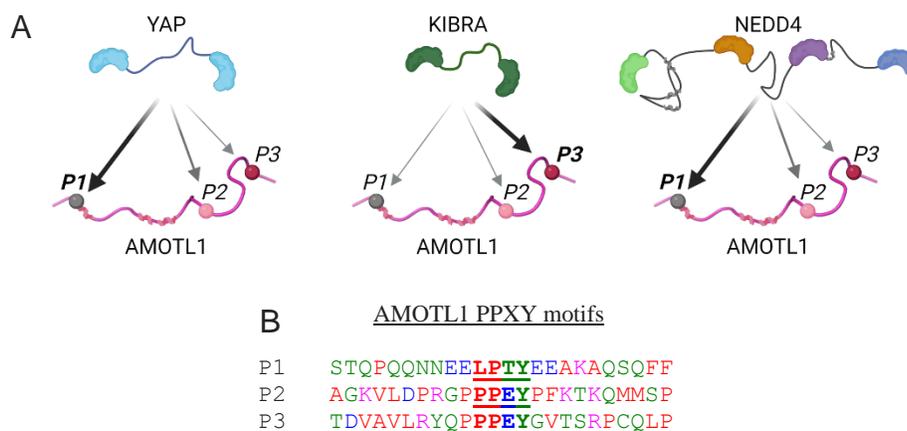


Figure A6.1: PPXY motif preferences of YAP, KIBRA, and NEDD4-1. (A) Cartoon representations of YAP YWW_{TD} (blue), KIBRA KWW_{TD} (green) and NEDD4 WW1-4 (mixed colors) WW domains binding AMOTL1 A₁₂₃ (pink) PPXY motifs (P1, P2, and P3). Thicker and darker arrows indicate more preferred PPXY motifs. The AMOTL1 cartoon was adapted from the AMOT polypeptide crystal structure (PDB: 6JJX). Created with Biorender.com. (B) Sequence alignment of AMOTL1 PPXY motifs and 10 flanking residues on either side. Amino acid coloring of red, blue, magenta, and green correspond to properties of small/hydrophobic, acidic, basic, hydroxyl/sulfidryl/amine/glycine residues, respectively.

Proposed rationale for motif specificity and future experiments

YAP and NEDD4-1 both have the highest affinity for P1, which may be explained by its unique properties relative to P2 and P3. P1 contains the sequence LPTY, whereas P2 and P3 contain the sequence PPEY. To see if the motif sequence provides specificity, P1 could be replaced with a PPEY sequence, followed by binding experiments with YAP and NEDD4-1 monitored by ITC. A second unique feature of

P1 is the four flanking Glu residues (Fig. 6.1B). Perhaps electrostatic interactions between these flanking residues and the WW domains of YAP and NEDD4-1 provide the enthalpy needed for motif specificity. These residues could be mutated to Ala and or Arg/Lys to test the effect of no charge and positive charges on binding affinity. Finally, P1 (residues 188-191) is adjacent to helix 1 (H1, residues 193-197). In the crystal structure of NEDD4L WW3 bound to a P1 peptide of AMOT, helical contacts were found to stabilize the interaction[163]. To test the importance of H1, residues 193-197 could be swapped with a disordered sequence and a sequence with high helical probability, followed by ITC binding studies.

Both YAP and NEDD4-1 have higher affinities for P2 than P3, even though both motifs contain PPEY sequences. No obvious differences in the properties of flanking residues are evident (Fig. A6.1B). However, since the flanking regions are not identical, sequence features could explain the different affinities for P2 and P3. This could be tested by swapping P2 and P3 including the ~10-residues on either side, followed by mutation of each motif to abolish binding. If YAP/NEDD4-1 show higher affinity for the new P3, it would be clear that flanking residues account for the different affinities of P2 versus P3. An alternate explanation is that linker length confers specificity. After YAP/NEDD4-1 bind P1, the next nearest PPXY motif is P2. The P1-P2 linker (L1) of AMOTL1 is already much longer than the WW domain linkers of YAP and NEDD4-1 (Fig. A6.2A). Thus, P3 might simply be too far away. To test this, P2 could be mutated, and the length of L1 could be reduced so that mutated L1+L2 is the same length as original L1 (Fig. A6.2B). Titration of YAP/NEDD4-1 into this mutated construct would reveal if linker length confers motif specificity.

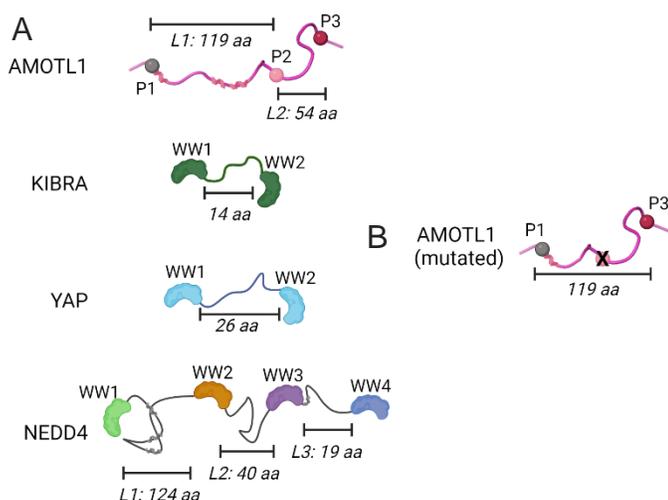


Figure A6.2: Motif and domain linker length.

(A) Cartoon representations of AMOTL1 A₁₂₃ (pink), YAP YWW_{TD} (blue), KIBRA KWW_{TD} (green) and NEDD4-1 WW1-4 (mixed colors) with the amino acid (aa) length (L) of each domain/motif linker indicated. (B) Mutated AMOTL1 construct with P2 abolished and a shorter P1-P3 linker. The AMOTL1 cartoon was adapted from the AMOT polypeptide crystal structure (PDB: 6JJX). Created with Biorender.com.

For KIBRA, it is less clear what features confer specificity for P3, since P2 and P3 contain PPEY sequences, and yet P2 has much weaker affinity. Previous work by Lin. et al. showed that KIBRA has the same motif preference for the parent isoform, AMOT[3], albeit with tighter binding affinity. In this study, the crystal structure of the KIBRA tandem WW domains in complex with a peptide of AMOT containing P3 and a noncanonical AMOT LMRY motif was solved. Here, P3 and the LMRY motif form a discrete 2:2 complex with WW1 and WW2 of KIBRA, respectively. AMOTL1 contains an analogous VLRY sequence two residues upstream of P3. To determine the importance of P3-flanking residues on binding, we generated an A₁₂₃ variant in which R363 and Y364, separated from P3 by two residues, were mutated to Ala (referred to as A_{RYΔAA}) and performed an ITC titration with KWW_{TD} (Fig. A6.3). The KWW_{TD}-A_{RYΔAA} did not reach sufficient saturation to compute accurate thermodynamic parameters. However, by fixing the stoichiometry, N to 1 in the data analysis process, we obtained a K_d of 11.8 μM, which represents an almost 17-fold weaker affinity relative to WT. This data provides strong evidence that P3-flanking residues play an important role in binding KWW_{TD}. Flanking residues may be important for established the full binding interface required for WW domain recognition, or the VLRY motif

may form a discrete complex with WW2 as seen in the crystal structure of AMOT and KIBRA. Further experiments are needed to differentiate between these possibilities.

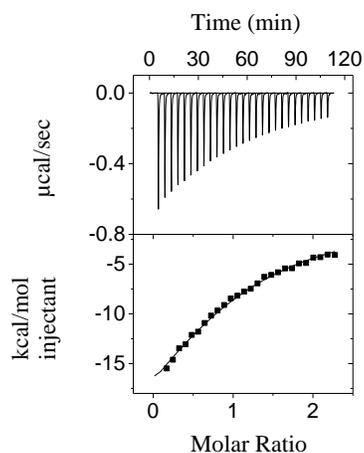


Figure A6.3: Binding of ARY Δ AA and KWW_{TD}. Representative ITC isotherm of 116-123 μ M KWW_{TD} titrated into 8-10 μ M ARY Δ AA in buffer composed of 50 mM sodium phosphate, 50 mM NaCl, 0.5 mM NaN₃, and 5 mM β -mercaptoethanol at pH 7.5. Data were collected at 25 °C. Instrument parameters and data analysis are described in Chapters 1, 2, and 5.

In NMR titration experiments of YWW_{TD} and KWW_{TD} into ¹⁵N-labeled A₁₂₃ (Chapter 2-3), we observed that, in addition to PPXY motif residues, the intensities of helix 3 (H3) residues were attenuated. These data suggest that binding may induce structural changes in H3. Unfortunately, since peaks disappear in the bound state, it is impossible to obtain NMR data for H3 after addition of YWW_{TD} or KWW_{TD}. To test the importance of H3 in complex stability, future experiments may replace this region with a random disordered segment and measure binding by ITC. Additionally, to see if helical structure is important for binding, H3 could be replaced with a sequence predicted to form a strong helix, followed by ITC binding experiments. Distal residue perturbations have been described for other PPXY systems and may be important for the regulation of complex assembly[90].

References

1. Liu, W., et al., *IBS: an illustrator for the presentation and visualization of biological sequences*. Bioinformatics, 2015. **31**(20): p. 3359-61.
2. Hermann, A., et al., *WW and C2 domain-containing proteins regulate hepatic cell differentiation and tumorigenesis through the hippo signaling pathway*. Hepatology, 2018. **67**(4): p. 1546-1559.
3. Lin, Z., et al., *Decoding WW domain tandem-mediated target recognitions in tissue growth and cell polarity*. Elife, 2019. **8**.
4. Vogel, A., A. Crawford, and A. Nyarko, *Multivalent Angiomin-like 1 and Yes-associated protein form a dynamic complex*. Protein Sci, 2022. **31**(5): p. e4295.
5. Madeira, F., et al., *Search and sequence analysis tools services from EMBL-EBI in 2022*. Nucleic Acids Res, 2022.
6. Radhakrishnan, K., et al., *Quantitative understanding of cell signaling: the importance of membrane organization*. Curr Opin Biotechnol, 2010. **21**(5): p. 677-82.
7. Bondos, S.E., A.K. Dunker, and V.N. Uversky, *Intrinsically disordered proteins play diverse roles in cell signaling*. Cell Commun Signal, 2022. **20**(1): p. 20.
8. Uversky, V.N., *Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics*. Frontiers in Physics, 2019. **7**.
9. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins: Structure, Function, and Genetics, 2000. **41**(3): p. 415-427.
10. Deiana, A., et al., *Intrinsically disordered proteins and structured proteins with intrinsically disordered regions have different functional roles in the cell*. PLoS One, 2019. **14**(8): p. e0217889.
11. Dill, K.A., *Dominant forces in protein folding*. Biochemistry, 1990. **29**(31): p. 7133-55.
12. Camilloni, C., et al., *Towards a structural biology of the hydrophobic effect in protein folding*. Sci Rep, 2016. **6**: p. 28285.
13. Dyson, H.J., *Making Sense of Intrinsically Disordered Proteins*. Biophys J, 2016. **110**(5): p. 1013-6.
14. Mateos, B., et al., *The Ambivalent Role of Proline Residues in an Intrinsically Disordered Protein: From Disorder Promoters to Compaction Facilitators*. J Mol Biol, 2020. **432**(9): p. 3093-3111.
15. Cheng, S., M. Cetinkaya, and F. Grater, *How sequence determines elasticity of disordered proteins*. Biophys J, 2010. **99**(12): p. 3863-9.
16. Khan, T., et al., *Polymorphism Analysis Reveals Reduced Negative Selection and Elevated Rate of Insertions and Deletions in Intrinsically Disordered Protein Regions*. Genome Biol Evol, 2015. **7**(6): p. 1815-26.
17. Uversky, V.N., *Chapter Four - Intrinsic Disorder, Protein-Protein Interactions, and Disease*. Advances in Protein Chemistry and Structural Biology. Vol. 110. 2018.

18. Babu, M.M., *The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease*. Biochem Soc Trans, 2016. **44**(5): p. 1185-1200.
19. Zarin, T., et al., *Proteome-wide signatures of function in highly diverged intrinsically disordered regions*. Elife, 2019. **8**.
20. Zarin, T., et al., *Identifying molecular features that are associated with biological function of intrinsically disordered protein regions*. Elife, 2021. **10**.
21. Bianchi, G., et al., *Distribution of Charged Residues Affects the Average Size and Shape of Intrinsically Disordered Proteins*. Biomolecules, 2022. **12**(4).
22. Tedeschi, G., et al., *Conformational response to charge clustering in synthetic intrinsically disordered proteins*. Biochim Biophys Acta Gen Subj, 2018. **1862**(10): p. 2204-2214.
23. Beveridge, R., et al., *Ion Mobility Mass Spectrometry Uncovers the Impact of the Patterning of Oppositely Charged Residues on the Conformational Distributions of Intrinsically Disordered Proteins*. J Am Chem Soc, 2019. **141**(12): p. 4908-4918.
24. Sherry, K.P., et al., *Control of transcriptional activity by design of charge patterning in the intrinsically disordered RAM region of the Notch receptor*. Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9243-E9252.
25. Das, R.K., et al., *Cryptic sequence features within the disordered protein p27Kip1 regulate cell cycle signaling*. Proc Natl Acad Sci U S A, 2016. **113**(20): p. 5616-21.
26. Muller-Spath, S., et al., *From the Cover: Charge interactions can dominate the dimensions of intrinsically disordered proteins*. Proc Natl Acad Sci U S A, 2010. **107**(33): p. 14609-14.
27. Zhou, H.X., *Intrinsic disorder: signaling via highly specific but short-lived association*. Trends Biochem Sci, 2012. **37**(2): p. 43-8.
28. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm*. J Mol Biol, 1999. **293**(2): p. 321-31.
29. Cortese, M.S., V.N. Uversky, and A.K. Dunker, *Intrinsic disorder in scaffold proteins: getting more from less*. Prog Biophys Mol Biol, 2008. **98**(1): p. 85-106.
30. van der Lee, R., et al., *Classification of intrinsically disordered regions and proteins*. Chem Rev, 2014. **114**(13): p. 6589-631.
31. Sugase, K., H.J. Dyson, and P.E. Wright, *Mechanism of coupled folding and binding of an intrinsically disordered protein*. Nature, 2007. **447**(7147): p. 1021-5.
32. Mugabo, Y. and G.E. Lim, *Scaffold Proteins: From Coordinating Signaling Pathways to Metabolic Regulation*. Endocrinology, 2018. **159**(11): p. 3615-3630.
33. Mana-Capelli, S. and D. McCollum, *Angiomotins stimulate LATS kinase autophosphorylation and act as scaffolds that promote Hippo signaling*. J Biol Chem, 2018. **293**(47): p. 18230-18241.
34. Ivarsson, Y. and P. Jemth, *Affinity and specificity of motif-based protein-protein interactions*. Curr Opin Struct Biol, 2019. **54**: p. 26-33.

35. Bah, A. and J.D. Forman-Kay, *Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications*. J Biol Chem, 2016. **291**(13): p. 6696-705.
36. Sorensen, C.S., A. Jendroszek, and M. Kjaergaard, *Linker Dependence of Avidity in Multivalent Interactions Between Disordered Proteins*. J Mol Biol, 2019. **431**(24): p. 4784-4795.
37. Teilum, K., J.G. Olsen, and B.B. Kragelund, *On the specificity of protein-protein interactions in the context of disorder*. Biochem J, 2021. **478**(11): p. 2035-2050.
38. Errington, W.J., B. Bruncsics, and C.A. Sarkar, *Mechanisms of noncanonical binding dynamics in multivalent protein-protein interactions*. Proc Natl Acad Sci U S A, 2019. **116**(51): p. 25659-25667.
39. Erlendsson, S. and K. Teilum, *Binding Revisited-Avidity in Cellular Function and Signaling*. Front Mol Biosci, 2020. **7**: p. 615565.
40. Uyar, B., et al., *Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer?* Mol Biosyst, 2014. **10**(10): p. 2626-42.
41. Stein, A. and P. Aloy, *Contextual specificity in peptide-mediated protein interactions*. PLoS One, 2008. **3**(7): p. e2524.
42. Flock, T., et al., *Controlling entropy to tune the functions of intrinsically disordered regions*. Curr Opin Struct Biol, 2014. **26**: p. 62-72.
43. Iesmantavicius, V., et al., *Helical propensity in an intrinsically disordered protein accelerates ligand binding*. Angew Chem Int Ed Engl, 2014. **53**(6): p. 1548-51.
44. Bugge, K., et al., *Interactions by Disorder - A Matter of Context*. Front Mol Biosci, 2020. **7**: p. 110.
45. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions*. Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
46. Kumar, M., et al., *The Eukaryotic Linear Motif resource: 2022 release*. Nucleic Acids Res, 2022. **50**(D1): p. D497-D508.
47. Sriram, S.M., et al., *Multivalency-assisted control of intracellular signaling pathways: application for ubiquitin- dependent N-end rule pathway*. Chem Biol, 2009. **16**(2): p. 121-31.
48. Tan, Z.C. and A.S. Meyer, *A general model of multivalent binding with ligands of heterotypic subunits and multiple surface receptors*. Math Biosci, 2021. **342**: p. 108714.
49. Slifka, M.K. and I.J. Amanna, *Role of Multivalency and Antigenic Threshold in Generating Protective Antibody Responses*. Front Immunol, 2019. **10**: p. 956.
50. Perez-Vale, K.Z., et al., *Multivalent interactions make adherens junction-cytoskeletal linkage robust during morphogenesis*. J Cell Biol, 2021. **220**(12).
51. He, X., et al., *From Monovalent to Multivalent Vaccines, the Exploration for Potential Preventive Strategies Against Hand, Foot, and Mouth Disease (HFMD)*. Virol Sin, 2021. **36**(2): p. 167-175.

52. Arsiwala, A., et al., *Designing Multivalent Ligands to Control Biological Interactions: From Vaccines and Cellular Effectors to Targeted Drug Delivery*. Chem Asian J, 2019. **14**(2): p. 244-255.
53. Li, Q., L. Yi, and P. Marek, *Antibodies and their multivalent constructs for cancer therapy*. Protein Pept Lett, 2014. **21**(10): p. 1017-24.
54. Deyev, S.M. and E.N. Lebedenko, *Multivalency: the hallmark of antibodies used for optimization of tumor targeting by design*. Bioessays, 2008. **30**(9): p. 904-18.
55. Zelikin, A.N. and F. Stellacci, *Broad-Spectrum Antiviral Agents Based on Multivalent Inhibitors of Viral Infectivity*. Adv Healthc Mater, 2021. **10**(6): p. e2001433.
56. Xiang, Y., et al., *Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2*. Science, 2020. **370**(6523): p. 1479-1484.
57. Long, D.D., et al., *A multivalent approach to drug discovery for novel antibiotics*. J Antibiot (Tokyo), 2008. **61**(10): p. 595-602.
58. Liu, Z., J. Zhang, and A. Zhang, *Design of multivalent ligand targeting G-protein-coupled receptors*. Curr Pharm Des, 2009. **15**(6): p. 682-718.
59. Kopinathan, A., et al., *Multivalent approaches and beyond: novel tools for the investigation of dopamine D2 receptor pharmacology*. Future Med Chem, 2016. **8**(11): p. 1349-72.
60. Bork, P. and M. Sudol, *The WW domain: a signalling site in dystrophin?* Trends Biochem Sci, 1994. **19**(12): p. 531-3.
61. Pucheta-Martinez, E., et al., *Changes in the folding landscape of the WW domain provide a molecular mechanism for an inherited genetic syndrome*. Sci Rep, 2016. **6**: p. 30293.
62. Hofmann, K. and P. Bucher, *The Rsp5-Domain Is Shared by Proteins of Diverse Functions*. Febs Letters, 1995. **358**(2): p. 153-157.
63. Andre, B. and J.Y. Springael, *Wwp, a New Amino-Acid Motif Present in Single or Multiple Copies in Various Proteins Including Dystrophin and the Sh3-Binding Yes-Associated Protein Yap65*. Biochemical and Biophysical Research Communications, 1994. **205**(2): p. 1201-1205.
64. Chang, N.S., et al., *Editorial: WW Domain Proteins in Signaling, Cancer Growth, Neural Diseases, and Metabolic Disorders*. Front Oncol, 2019. **9**: p. 719.
65. Salah, Z., A. Alian, and R.I. Aqeilan, *WW domain-containing proteins: retrospectives and the future*. Front Biosci (Landmark Ed), 2012. **17**(1): p. 331-48.
66. Huang, S.S., L.J. Hsu, and N.S. Chang, *Functional role of WW domain-containing proteins in tumor biology and diseases: Insight into the role in ubiquitin-proteasome system*. FASEB Bioadv, 2020. **2**(4): p. 234-253.
67. Dodson, E.J., et al., *Versatile communication strategies among tandem WW domain repeats*. Exp Biol Med (Maywood), 2015. **240**(3): p. 351-60.
68. Chen, H.I. and M. Sudol, *The Ww Domain of Yes-Associated Protein Binds a Proline-Rich Ligand That Differs from the Consensus Established for Src Homology 3-Binding Modules*. Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(17): p. 7819-7823.

69. Chen, H.I., et al., *Characterization of the WW domain of human yes-associated protein and its polyproline-containing ligands*. Journal of Biological Chemistry, 1997. **272**(27): p. 17070-17077.
70. Kay, B.K., M.P. Williamson, and M. Sudol, *The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains*. The FASEB Journal, 2000. **14**(2): p. 231-241.
71. Ermekova, K.S., et al., *The WW domain of neural protein FE65 interacts with proline-rich motifs in Mena, the mammalian homolog of Drosophila enabled*. J Biol Chem, 1997. **272**(52): p. 32869-77.
72. Bedford, M.T., R. Reed, and P. Leder, *WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: The proline glycine and methionine-rich motif*. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(18): p. 10602-10607.
73. Lu, P.J., et al., *Function of WW domains as phosphoserine- or phosphothreonine-binding modules*. Science, 1999. **283**(5406): p. 1325-8.
74. Bedford, M.T., et al., *A novel pro-Arg motif recognized by WW domains*. J Biol Chem, 2000. **275**(14): p. 10359-69.
75. Tapia, V.E., et al., *Y65C missense mutation in the WW domain of the Golabi-Ito-Hall syndrome protein PQBP1 affects its binding activity and deregulates pre-mRNA splicing*. J Biol Chem, 2010. **285**(25): p. 19391-401.
76. Zarrinpar, A. and W.A. Lim, *Converging on proline: the mechanism of WW domain peptide recognition*. Nature Structural Biology, 2000. **7**(8): p. 611-613.
77. Macias, M.J., et al., *Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide*. Nature, 1996. **382**(6592): p. 646-9.
78. Iglesias-Bexiga, M., et al., *WW domains of the yes-kinase-associated-protein (YAP) transcriptional regulator behave as independent units with different binding preferences for PPxY motif-containing ligands*. PLoS One, 2015. **10**(1): p. e0113828.
79. Kasanov, J., et al., *Characterizing Class I WW domains defines key specificity determinants and generates mutant domains with novel specificities*. Chem Biol, 2001. **8**(3): p. 231-41.
80. Castillo, F., et al., *Phage display identification of nanomolar ligands for human NEDD4-WW3: Energetic and dynamic implications for the development of broad-spectrum antivirals*. Int J Biol Macromol, 2022. **207**: p. 308-323.
81. Hu, H., et al., *A map of WW domain family interactions*. Proteomics, 2004. **4**(3): p. 643-55.
82. Abu-Odeh, M., et al., *Characterizing WW domain interactions of tumor suppressor WWOX reveals its association with multiprotein networks*. J Biol Chem, 2014. **289**(13): p. 8865-80.
83. Schuchardt, B.J., et al., *Ligand binding to WW tandem domains of YAP2 transcriptional regulator is under negative cooperativity*. FEBS J, 2014. **281**(24): p. 5532-51.

84. Kwok, E., et al., *Intrinsic disorder and amino acid specificity modulate binding of the WW2 domain in kidney and brain protein (KIBRA) to synaptodin*. J Biol Chem, 2019. **294**(46): p. 17383-17394.
85. Ji, Z., et al., *Kibra Modulates Learning and Memory via Binding to Dendrin*. Cell Rep, 2019. **26**(8): p. 2064-2077 e7.
86. Wiesner, S., et al., *Solution structure and ligand recognition of the WW domain pair of the yeast splicing factor Prp40*. J Mol Biol, 2002. **324**(4): p. 807-22.
87. Huang, X., et al., *Structure and function of the two tandem WW domains of the pre-mRNA splicing factor FBP21 (formin-binding protein 21)*. J Biol Chem, 2009. **284**(37): p. 25375-87.
88. Iglesias-Bexiga, M., et al., *Binding site plasticity in viral PPxY Late domain recognition by the third WW domain of human NEDD4*. Sci Rep, 2019. **9**(1): p. 15076.
89. Rentschler, S., et al., *The WW domain of dystrophin requires EF-hands region to interact with beta-dystroglycan*. Biol Chem, 1999. **380**(4): p. 431-42.
90. Nyarko, A., *Differential Binding Affinities and Allosteric Conformational Changes Underlie Interactions of Yorkie and a Multivalent PPxY Partner*. Biochemistry, 2018. **57**(5): p. 547-556.
91. Salah, Z. and R.I. Aqeilan, *WW domain interactions regulate the Hippo tumor suppressor pathway*. Cell Death Dis, 2011. **2**: p. e172.
92. Sudol, M., *Newcomers to the WW Domain-Mediated Network of the Hippo Tumor Suppressor Pathway*. Genes Cancer, 2010. **1**(11): p. 1115-8.
93. Sudol, M., *Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product*. Oncogene, 1994. **9**(8): p. 2145-2152.
94. Sudol, M., et al., *Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain*. J Biol Chem, 1995. **270**(24): p. 14733-41.
95. Sudol, M., *YAP1 oncogene and its eight isoforms*. Oncogene, 2013. **32**(33): p. 3922.
96. LeBlanc, L., N. Ramirez, and J. Kim, *Context-dependent roles of YAP/TAZ in stem cell fates and cancer*. Cell Mol Life Sci, 2021. **78**(9): p. 4201-4219.
97. Li, Z., et al., *Structural insights into the YAP and TEAD complex*. Genes Dev, 2010. **24**(3): p. 235-40.
98. Callus, B.A., et al., *YAPping about and not forgetting TAZ*. FEBS Lett, 2019. **593**(3): p. 253-276.
99. Kim, M.K., J.W. Jang, and S.C. Bae, *DNA binding partners of YAP/TAZ*. BMB Rep, 2018. **51**(3): p. 126-133.
100. Totaro, A., T. Panciera, and S. Piccolo, *YAP/TAZ upstream signals and downstream responses*. Nat Cell Biol, 2018. **20**(8): p. 888-899.
101. Luo, J. and P. Li, *Context-dependent transcriptional regulations of YAP/TAZ in stem cell and differentiation*. Stem Cell Res Ther, 2022. **13**(1): p. 10.
102. Yu, F.X., B. Zhao, and K.L. Guan, *Hippo Pathway in Organ Size Control, Tissue Homeostasis, and Cancer*. Cell, 2015. **163**(4): p. 811-28.

103. Dey, A., X. Varelas, and K.L. Guan, *Targeting the Hippo pathway in cancer, fibrosis, wound healing and regenerative medicine*. Nat Rev Drug Discov, 2020. **19**(7): p. 480-494.
104. Jin, J., et al., *The Effects of YAP and Its Related Mechanisms in Central Nervous System Diseases*. Front Neurosci, 2020. **14**: p. 595.
105. Pan, D., *The hippo signaling pathway in development and cancer*. Dev Cell, 2010. **19**(4): p. 491-505.
106. Jiang, L., et al., *YAP-mediated crosstalk between the Wnt and Hippo signaling pathways (Review)*. Mol Med Rep, 2020. **22**(5): p. 4101-4106.
107. Tang, F. and G. Christofori, *The cross-talk between the Hippo signaling pathway and autophagy: implications on physiology and cancer*. Cell Cycle, 2020. **19**(20): p. 2563-2572.
108. Chen, X., et al., *Role of Hippo-YAP1/TAZ pathway and its crosstalk in cardiac biology*. Int J Biol Sci, 2020. **16**(13): p. 2454-2463.
109. Nishio, M., et al., *The Hippo Signaling Pathway: A Candidate New Drug Target for Malignant Tumors*, in *Innovative Medicine: Basic Research and Development*, K. Nakao, N. Minato, and S. Uemoto, Editors. 2015: Tokyo. p. 79-94.
110. Meng, Z., T. Moroishi, and K.L. Guan, *Mechanisms of Hippo pathway regulation*. Genes Dev, 2016. **30**(1): p. 1-17.
111. Zheng, Y. and D. Pan, *The Hippo Signaling Pathway in Development and Disease*. Dev Cell, 2019. **50**(3): p. 264-282.
112. Zhang, L., et al., *KIBRA: In the brain and beyond*. Cell Signal, 2014. **26**(7): p. 1392-9.
113. Schneider, A., et al., *KIBRA: A New Gateway to Learning and Memory?* Front Aging Neurosci, 2010. **2**: p. 4.
114. Yu, J., et al., *Kibra functions as a tumor suppressor protein that regulates Hippo signaling in conjunction with Merlin and Expanded*. Dev Cell, 2010. **18**(2): p. 288-99.
115. Nishio, M., et al., *Capturing the mammalian Hippo: elucidating its role in cancer*. Cancer Sci, 2013. **104**(10): p. 1271-7.
116. Troyanovsky, B., et al., *Angiomotin: an angiostatin binding protein that regulates endothelial cell migration and tube formation*. J Cell Biol, 2001. **152**(6): p. 1247-54.
117. Moleirinho, S., W. Guarrant, and J.L. Kissil, *The Angiomotins--from discovery to function*. FEBS Lett, 2014. **588**(16): p. 2693-703.
118. Bratt, A., et al., *Angiomotin belongs to a novel protein family with conserved coiled-coil and PDZ binding domains*. Gene, 2002. **298**(1): p. 69-77.
119. Lv, M., et al., *Angiomotin Family Members: Oncogenes or Tumor Suppressors?* Int J Biol Sci, 2017. **13**(6): p. 772-781.
120. Ernkvist, M., et al., *p130-angiomotin associates to actin and controls endothelial cell shape*. FEBS J, 2006. **273**(9): p. 2000-11.
121. Matsumoto, H., et al., *Differential expression of the motin family in the peri-implantation mouse uterus and their hormonal regulation*. J Reprod Dev, 2012. **58**(6): p. 649-53.

122. Leung, C.Y. and M. Zernicka-Goetz, *Angiomotin prevents pluripotent lineage differentiation in mouse embryos via Hippo pathway-dependent and -independent mechanisms*. Nat Commun, 2013. **4**: p. 2251.
123. Hirate, Y., et al., *Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos*. Curr Biol, 2013. **23**(13): p. 1181-94.
124. Zhao, B., et al., *Angiomotin is a novel Hippo pathway component that inhibits YAP oncoprotein*. Genes Dev, 2011. **25**(1): p. 51-63.
125. Thanh Nguyen, H., et al., *Deubiquitylating enzyme USP9x regulates hippo pathway activity by controlling angiomotin protein turnover*. Cell Discov, 2016. **2**: p. 16001.
126. Skouloudaki, K. and G. Walz, *YAP1 recruits c-Abl to protect angiomotin-like 1 from Nedd4-mediated degradation*. PLoS One, 2012. **7**(4): p. e35735.
127. Choi, K.S., et al., *The endothelial E3 ligase HECW2 promotes endothelial cell junctions by increasing AMOTL1 protein stability via K63-linked ubiquitination*. Cell Signal, 2016. **28**(11): p. 1642-51.
128. Kumar, S., Y. Tomooka, and M. Noda, *Identification of a set of genes with developmentally down-regulated expression in the mouse brain*. Biochem Biophys Res Commun, 1992. **185**(3): p. 1155-61.
129. Boase, N.A. and S. Kumar, *NEDD4: The founding member of a family of ubiquitin-protein ligases*. Gene, 2015. **557**(2): p. 113-22.
130. Persaud, A., et al., *Comparison of substrate specificity of the ubiquitin ligases Nedd4 and Nedd4-2 using proteome arrays*. Mol Syst Biol, 2009. **5**: p. 333.
131. Huang, X., et al., *The many substrates and functions of NEDD4-1*. Cell Death Dis, 2019. **10**(12): p. 904.
132. Weber, J., S. Polo, and E. Maspero, *HECT E3 Ligases: A Tale With Multiple Facets*. Front Physiol, 2019. **10**: p. 370.
133. Delvecchio, V.S., et al., *Emerging roles of the HECT-type E3 ubiquitin ligases in hematological malignancies*. Discov Oncol, 2021. **12**(1): p. 39.
134. Zhang, Y., et al., *E3 Ubiquitin ligase NEDD4 family-regulatory network in cardiovascular disease*. Int J Biol Sci, 2020. **16**(14): p. 2727-2740.
135. Conway, J.A., G. Kinsman, and E.R. Kramer, *The Role of NEDD4 E3 Ubiquitin-Protein Ligases in Parkinson's Disease*. Genes (Basel), 2022. **13**(3).
136. Schmidt, M.F., et al., *Ubiquitin signalling in neurodegeneration: mechanisms and therapeutic opportunities*. Cell Death Differ, 2021. **28**(2): p. 570-590.
137. Ye, X., et al., *NEDD4: a promising target for cancer therapy*. Curr Cancer Drug Targets, 2014. **14**(6): p. 549-56.
138. Shepley-McTaggart, A., et al., *Viruses go modular*. J Biol Chem, 2020. **295**(14): p. 4604-4616.
139. Salah, Z., et al., *NEDD4 E3 ligase inhibits the activity of the Hippo pathway by targeting LATS1 for degradation*. Cell Cycle, 2013. **12**(24): p. 3817-23.
140. Bae, S.J., et al., *NEDD4 controls intestinal stem cell homeostasis by regulating the Hippo signalling pathway*. Nat Commun, 2015. **6**: p. 6314.
141. Wang, C., et al., *The Nedd4-like ubiquitin E3 ligases target angiomotin/p130 to ubiquitin-dependent degradation*. Biochem J, 2012. **444**(2): p. 279-89.

142. Novelli, G., et al., *Inhibition of HECT E3 ligases as potential therapy for COVID-19*. Cell Death Dis, 2021. **12**(4): p. 310.
143. Valerdi, K.M., et al., *The Role of the Host Ubiquitin System in Promoting Replication of Emergent Viruses*. Viruses, 2021. **13**(3).
144. Dong, J., et al., *Elucidation of a universal size-control mechanism in Drosophila and mammals*. Cell, 2007. **130**(6): p. 1120-33.
145. Abylkassov, R. and Y. Xie, *Role of Yes-associated protein in cancer: An update*. Oncol Lett, 2016. **12**(4): p. 2277-2282.
146. Huang, T., et al., *The physiological role of Motin family and its dysregulation in tumorigenesis*. J Transl Med, 2018. **16**(1): p. 98.
147. Chan, S.W., et al., *Hippo pathway-independent restriction of TAZ and YAP by angiominin*. J Biol Chem, 2011. **286**(9): p. 7018-26.
148. Ernkvist, M., et al., *Differential roles of p80- and p130-angiominin in the switch between migration and stabilization of endothelial cells*. Biochim Biophys Acta, 2008. **1783**(3): p. 429-37.
149. Yi, C., et al., *The p130 isoform of angiominin is required for Yap-mediated hepatic epithelial cell proliferation and tumorigenesis*. Sci Signal, 2013. **6**(291): p. ra77.
150. Wang, W., J. Huang, and J. Chen, *Angiominin-like proteins associate with and negatively regulate YAP1*. J Biol Chem, 2011. **286**(6): p. 4364-70.
151. Oka, T., A.P. Schmitt, and M. Sudol, *Opposing roles of angiominin-like-1 and zona occludens-2 on pro-apoptotic function of YAP*. Oncogene, 2012. **31**(1): p. 128-34.
152. Ragni, C.V., et al., *Amotl1 mediates sequestration of the Hippo effector Yap1 downstream of Fat4 to restrict heart growth*. Nat Commun, 2017. **8**: p. 14582.
153. Mana-Capelli, S., et al., *Angiominins link F-actin architecture to Hippo pathway signaling*. Mol Biol Cell, 2014. **25**(10): p. 1676-85.
154. Whitmore, L. and B.A. Wallace, *DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W668-73.
155. Miles, A.J., S.G. Ramalli, and B.A. Wallace, *DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data*. Protein Science, 2022. **31**(1): p. 37-46.
156. Arakawa, T. and J. Wen, *Size-exclusion chromatography with on-line light scattering*. Curr Protoc Protein Sci, 2001. **Chapter 20**: p. Unit 20 6.
157. McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server*. Bioinformatics, 2000. **16**(4): p. 404-5.
158. Camilloni, C., et al., *Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts*. Biochemistry, 2012. **51**(11): p. 2224-31.
159. Bolik-Coulon, N., et al., *Chapter 3 - Experimental characterization of the dynamics of IDPs and IDRs by NMR*, in *Intrinsically Disordered Proteins*, N. Salvi, Editor. 2019, Academic Press. p. 65-92.
160. Cavanagh, J., et al., *Protein NMR spectroscopy: principles and practice*. 1996: Academic press.

161. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding*. Prog Nucl Magn Reson Spectrosc, 2013. **73**: p. 1-16.
162. Baker, K., et al., *Yorkie-Warts Complexes are an Ensemble of Interconverting Conformers Formed by Multivalent Interactions*. Journal of Molecular Biology, 2021. **433**(4): p. 166776 -166788.
163. Rheinemann, L., et al., *Interactions between AMOT PPxY motifs and NEDD4L WW domains function in HIV-1 release*. J Biol Chem, 2021. **297**(2): p. 100975.
164. Cai, D., et al., *Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression*. Nat Cell Biol, 2019. **21**(12): p. 1578-1589.
165. Banani, S.F., et al., *Biomolecular condensates: organizers of cellular biochemistry*. Nat Rev Mol Cell Biol, 2017. **18**(5): p. 285-298.
166. Mitrea, D.M. and R.W. Kriwacki, *Phase separation in biology; functional organization of a higher order*. Cell Communication and Signaling, 2016. **14**(1): p. 1.
167. Martin, E.W. and T. Mittag, *Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions*. Biochemistry, 2018. **57**(17): p. 2478-2487.
168. Harmon, T.S., et al., *Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins*. Elife, 2017. **6**.
169. Favier, A. and B. Brutscher, *Recovering lost magnetization: polarization enhancement in biomolecular NMR*. J Biomol NMR, 2011. **49**(1): p. 9-15.
170. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
171. Lee, W., M. Tonelli, and J.L. Markley, *NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy*. Bioinformatics, 2015. **31**(8): p. 1325-7.
172. Johnson, B.A., *Using NMRView to visualize and analyze the NMR spectra of macromolecules*. Methods Mol Biol, 2004. **278**: p. 313-52.
173. Kjaergaard, M., S. Brander, and F.M. Poulsen, *Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH*. J Biomol NMR, 2011. **49**(2): p. 139-49.
174. Kjaergaard, M. and F.M. Poulsen, *Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution*. J Biomol NMR, 2011. **50**(2): p. 157-65.
175. Schwarzingler, S., et al., *Sequence-dependent correction of random coil NMR chemical shifts*. J Am Chem Soc, 2001. **123**(13): p. 2970-8.
176. Zhu, G., et al., *Protein dynamics measurements by TROSY-based NMR experiments*. J Magn Reson, 2000. **143**(2): p. 423-6.
177. Miles, A.J., S.G. Ramalli, and B.A. Wallace, *DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data*. Protein Sci, 2022. **31**(1): p. 37-46.
178. Lees, J.G., et al., *A reference database for circular dichroism spectroscopy covering fold and secondary structure space*. Bioinformatics, 2006. **22**(16): p. 1955-62.

179. Yeung, Y.T., et al., *Dysregulation of the Hippo pathway signaling in aging and cancer*. *Pharmacol Res*, 2019. **143**: p. 151-165.
180. Chen, H.I., et al., *Characterization of the WW domain of human yes-associated protein and its polyproline-containing ligands*. *J Biol Chem*, 1997. **272**(27): p. 17070-7.
181. Ray, G., P.T. Schmitt, and A.P. Schmitt, *Angiomotin-Like 1 Links Paramyxovirus M Proteins to NEDD4 Family Ubiquitin Ligases*. *Viruses*, 2019. **11**(2).
182. Vogel, A., A. Crawford, and A. Nyarko, *Multivalent Angiomotin-like 1 and Yes-associated protein form a dynamic complex via an 'anchor and sample' mechanism*. In review in *Protein Science*, 2021.
183. Duning, K., et al., *KIBRA modulates directional migration of podocytes*. *J Am Soc Nephrol*, 2008. **19**(10): p. 1891-903.
184. Genevet, A., et al., *Kibra is a regulator of the Salvador/Warts/Hippo signaling network*. *Dev Cell*, 2010. **18**(2): p. 300-8.
185. Swaroop, B.S., et al., *KIBRA Connects Hippo Signaling and Cancer*. *Exp Cell Res*, 2021: p. 112613.
186. Olsen, J.G., K. Teilum, and B.B. Kragelund, *Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness*. *Cell Mol Life Sci*, 2017. **74**(17): p. 3175-3183.
187. Johnson, B.A. and R.A. Blevins, *NMR View: A computer program for the visualization and analysis of NMR data*. *J Biomol NMR*, 1994. **4**(5): p. 603-14.
188. Staub, O., et al., *Regulation of the epithelial Na⁺ channel by Nedd4 and ubiquitination*. *Kidney Int*, 2000. **57**(3): p. 809-15.
189. Yan, S., et al., *NEDD4-1 Is a Key Regulator of Epidermal Homeostasis and Wound Repair*. *J Invest Dermatol*, 2022. **142**(6): p. 1703-1713 e11.
190. He, H., et al., *An outlined review for the role of Nedd4-1 and Nedd4-2 in lung disorders*. *Biomed Pharmacother*, 2020. **125**: p. 109983.
191. Wang, Z.W., et al., *NEDD4 E3 ligase: Functions and mechanism in human cancer*. *Semin Cancer Biol*, 2020. **67**(Pt 2): p. 92-101.
192. Xu, D., et al., *The role of Nedd4-1 WW domains in binding and regulating human organic anion transporter 1*. *Am J Physiol Renal Physiol*, 2016. **311**(2): p. F320-9.
193. Kanelis, V., et al., *Structural determinants for high-affinity binding in a Nedd4 WW3* domain-Comm PY motif complex*. *Structure*, 2006. **14**(3): p. 543-53.
194. Henry, P.C., et al., *Affinity and specificity of interactions between Nedd4 isoforms and the epithelial Na⁺ channel*. *J Biol Chem*, 2003. **278**(22): p. 20019-28.
195. Panwalkar, V., et al., *Multiple WW domains of Nedd4-1 undergo conformational exchange that is quenched upon peptide binding*. *FEBS Lett*, 2017. **591**(11): p. 1573-1583.
196. Zhu, K., et al., *Allosteric auto-inhibition and activation of the Nedd4 family E3 ligase Itch*. *EMBO Rep*, 2017. **18**(9): p. 1618-1630.

197. Wang, Z., et al., *A multi-lock inhibitory mechanism for fine-tuning enzyme activities of the HECT family E3 ligases*. Nat Commun, 2019. **10**(1): p. 3162.
198. Mari, S., et al., *Structural and functional framework for the autoinhibition of Nedd4-family ubiquitin ligases*. Structure, 2014. **22**(11): p. 1639-49.
199. Wiesner, S., et al., *Autoinhibition of the HECT-Type Ubiquitin Ligase Smurf2 through Its C2 Domain*. Cell, 2007. **130**(4): p. 651-662.
200. Woody, R.W., *Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins*. Eur Biophys J, 1994. **23**(4): p. 253-62.
201. Koepf, E.K., et al., *Characterization of the structure and function of W --> F WW domain variants: identification of a natively unfolded protein that folds upon ligand binding*. Biochemistry, 1999. **38**(43): p. 14338-51.
202. Chen, I., B.M. Dorr, and D.R. Liu, *A general strategy for the evolution of bond-forming enzymes using yeast display*. Proc Natl Acad Sci U S A, 2011. **108**(28): p. 11399-404.
203. Dorr, B.M., et al., *Reprogramming the specificity of sortase enzymes*. Proc Natl Acad Sci U S A, 2014. **111**(37): p. 13343-8.
204. Ying, J., et al., *Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data*. J Biomol NMR, 2017. **68**(2): p. 101-118.
205. Johnson, B.A., *Using NMRView to visualize and analyze the NMR spectra of macromolecules*. Methods Mol Biol, 2004. **278**: p. 313-352.
206. Shen, Y. and A. Bax, *Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks*. J Biomol NMR, 2013. **56**(3): p. 227-41.
207. Maltsev, A.S., J. Ying, and A. Bax, *Deuterium isotope shifts for backbone (1)H, (1)(5)N and (1)(3)C nuclei in intrinsically disordered protein alpha-synuclein*. J Biomol NMR, 2012. **54**(2): p. 181-91.
208. Farrow, N.A., et al., *Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation*. Biochemistry, 1994. **33**(19): p. 5984-6003.
209. Couderc, C., et al., *AMOTL1 Promotes Breast Cancer Progression and Is Antagonized by Merlin*. Neoplasia, 2016. **18**(1): p. 10-24.
210. Spagnol, G., et al., *Structural Studies of the Nedd4 WW Domains and Their Selectivity for the Connexin43 (Cx43) Carboxyl Terminus*. J Biol Chem, 2016. **291**(14): p. 7637-50.
211. Qi, S., et al., *Structural and biochemical basis for ubiquitin ligase recruitment by arrestin-related domain-containing protein-3 (ARRDC3)*. J Biol Chem, 2014. **289**(8): p. 4743-52.
212. Acuner Ozbabacan, S.E., et al., *Transient protein-protein interactions*. Protein Eng Des Sel, 2011. **24**(9): p. 635-48.
213. Jiang, S., et al., *Protein phase separation and its role in tumorigenesis*. Elife, 2020. **9**.
214. Xu, Y., et al., *Liquid-Liquid Phase-Separated Systems from Reversible Gel-Sol Transition of Protein Microgels*. Adv Mater, 2021. **33**(33): p. e2008670.

215. Lu, J., et al., *Emerging Roles of Liquid-Liquid Phase Separation in Cancer: From Protein Aggregation to Immune-Associated Signaling*. *Front Cell Dev Biol*, 2021. **9**: p. 631486.
216. Alberti, S. and D. Dormann, *Liquid-Liquid Phase Separation in Disease*. *Annu Rev Genet*, 2019. **53**: p. 171-194.
217. Felli, I.C. and R. Pierattelli, *(13)C Direct Detected NMR for Challenging Systems*. *Chem Rev*, 2022. **122**(10): p. 9468-9496.
218. Zhu, P., et al., *A Highly Versatile Expression System for the Production of Multiply Phosphorylated Proteins*. *ACS Chem Biol*, 2019. **14**(7): p. 1564-1572.
219. Frey, S. and D. Gorlich, *A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins*. *J Chromatogr A*, 2014. **1337**: p. 95-105.
220. Webb, C., et al., *Structural features and ligand binding properties of tandem WW domains from YAP and TAZ, nuclear effectors of the Hippo pathway*. *Biochemistry*, 2011. **50**(16): p. 3300-9.
221. Franco, R., et al., *Probing Conformational Exchange Dynamics in a Short-Lived Protein Folding Intermediate by Real-Time Relaxation-Dispersion NMR*. *J Am Chem Soc*, 2017. **139**(3): p. 1065-1068.
222. Morin, S., et al., *relax: the analysis of biomolecular kinetics and thermodynamics using NMR relaxation dispersion data*. *Bioinformatics*, 2014. **30**(15): p. 2219-20.
223. Welker, L., J.C. Paillart, and S. Bernacchi, *Importance of Viral Late Domains in Budding and Release of Enveloped RNA Viruses*. *Viruses*, 2021. **13**(8).
224. Garcia, M.L., et al., *Functional characterization of the putative hepatitis B virus core protein late domain using retrovirus chimeras*. *PLoS One*, 2013. **8**(8): p. e72845.
225. Rost, M., et al., *Gamma-adaptin, a novel ubiquitin-interacting adaptor, and Nedd4 ubiquitin ligase control hepatitis B virus maturation*. *J Biol Chem*, 2006. **281**(39): p. 29297-308.
226. Han, Z., et al., *Modular mimicry and engagement of the Hippo pathway by Marburg virus VP40: Implications for filovirus biology and budding*. *PLoS Pathog*, 2020. **16**(1): p. e1008231.
227. Sakurai, A., et al., *Regulation of human T-cell leukemia virus type 1 (HTLV-1) budding by ubiquitin ligase Nedd4*. *Microbes Infect*, 2004. **6**(2): p. 150-6.
228. Zhao, T., et al., *HTLV-1 activates YAP via NF-kappaB/p65 to promote oncogenesis*. *Proc Natl Acad Sci U S A*, 2022. **119**(9).
229. Wang, Z., et al., *The Hippo Pathway and Viral Infections*. *Front Microbiol*, 2019. **10**: p. 3033.
230. Birnbaum, F. and M. Nassal, *Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein*. *J Virol*, 1990. **64**(7): p. 3319-30.
231. Zhou, S. and D.N. Standring, *Hepatitis B virus capsid particles are assembled from core-protein dimer precursors*. *Proc Natl Acad Sci U S A*, 1992. **89**(21): p. 10046-50.
232. Steven, A.C., et al., *Structure, Assembly, and Antigenicity of Hepatitis B Virus Capsid Proteins*, in *Virus Structure and Assembly*. 2005. p. 125-164.

233. Langerova, H., et al., *Hepatitis B Core Protein Is Post-Translationally Modified through K29-Linked Ubiquitination*. Cells, 2020. **9**(12).
234. Wynne, S.A., R.A. Crowther, and A.G. Leslie, *The crystal structure of the human hepatitis B virus capsid*. Mol Cell, 1999. **3**(6): p. 771-80.
235. Carver, J.P. and R.E. Richards, *A general two-site solution for the chemical exchange produced dependence of T2 upon the carr-Purcell pulse separation*. Journal of Magnetic Resonance (1969), 1972. **6**(1): p. 89-105.
236. Wang, B., et al., *Liquid-liquid phase separation in human health and diseases*. Signal Transduct Target Ther, 2021. **6**(1): p. 290.
237. Alberti, S., A. Gladfelter, and T. Mittag, *Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates*. Cell, 2019. **176**(3): p. 419-434.
238. Boyko, S. and W.K. Surewicz, *Tau liquid-liquid phase separation in neurodegenerative diseases*. Trends Cell Biol, 2022. **32**(7): p. 611-623.
239. Kanaan, N.M., et al., *Liquid-liquid phase separation induces pathogenic tau conformations in vitro*. Nat Commun, 2020. **11**(1): p. 2809.
240. Elbaum-Garfinkle, S., *Matter over mind: Liquid phase separation and neurodegeneration*. J Biol Chem, 2019. **294**(18): p. 7160-7168.
241. Ray, S., et al., *alpha-Synuclein aggregation nucleates through liquid-liquid phase separation*. Nat Chem, 2020. **12**(8): p. 705-716.
242. Kanekura, K. and M. Kuroda, *How can we interpret the relationship between liquid-liquid phase separation and amyotrophic lateral sclerosis?* Lab Invest, 2022.
243. Yang, J. and X. Yang, *Phase Transition of Huntingtin: Factors and Pathological Relevance*. Front Genet, 2020. **11**: p. 754.
244. Lin, Y., S.L. Currie, and M.K. Rosen, *Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs*. J Biol Chem, 2017. **292**(46): p. 19110-19120.
245. Lu, Y., et al., *Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression*. Nat Cell Biol, 2020. **22**(4): p. 453-464.
246. Li, R.H., et al., *A phosphatidic acid-binding lncRNA SNHG9 facilitates LATS1 liquid-liquid phase separation to promote oncogenic YAP signaling*. Cell Res, 2021. **31**(10): p. 1088-1105.
247. Tripathi, S., et al., *TAZ exhibits phase separation properties and interacts with Smad7 and beta-catenin to repress skeletal myogenesis*. J Cell Sci, 2022. **135**(1).
248. Yu, M., et al., *Interferon-gamma induces tumor resistance to anti-PD-1 immunotherapy by promoting YAP phase separation*. Mol Cell, 2021. **81**(6): p. 1216-1230 e9.
249. Chu, X., et al., *Prediction of liquid-liquid phase separating proteins using machine learning*. BMC Bioinformatics, 2022. **23**(1): p. 72.
250. Sanulli, S. and G.J. Narlikar, *Generation and Biochemical Characterization of Phase-Separated Droplets Formed by Nucleic Acid Binding Proteins: Using HP1 as a Model System*. Curr Protoc, 2021. **1**(5): p. e109.

251. Dao, T.P., et al., *Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions*. *Mol Cell*, 2018. **69**(6): p. 965-978 e6.
252. O'Flynn, B.G. and T. Mittag, *The role of liquid-liquid phase separation in regulating enzyme activity*. *Curr Opin Cell Biol*, 2021. **69**: p. 70-79.
253. Zheng Wang, G.Z., Hong Zhang, *Protocol for analyzing protein liquid-liquid phase separation*. *Biophysics Reports*, 2018: p. 1-9.