Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Joseph L. Mertz Date

Next Generation Proteomics in the Nervous System

By

Joseph L. Mertz
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Neuroscience
James T. Laney Graduate School

_____
Junmin Peng, Ph.D.
Advisor

_____
Gary Bassell, Ph.D.
Committee Member

_____
Tamara Caspary, Ph.D.
Committee Member

_____
Peng Jin, Ph.D.
Committee Member

_____
James Zheng, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Next Generation Proteomics in the Nervous System

By

Joseph L. Mertz
B.S., Cornell University, 2005

Advisor: Junmin Peng, Ph.D.

An abstract of
a dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Neuroscience
2015

Abstract

Next Generation Proteomics in the Nervous System
By Joseph L. Mertz

Classical approaches to protein signaling pathways in the nervous system disregard the vast complexity of biological systems in general, and the brain especially. Large gaps remain in our understanding of the physiological and pathogenic roles of critical proteins such as the E3 ubiquitin ligase Mind bomb 1 (Mib1), the Alzheimer's disease (AD) hallmark Amyloid progenitor protein (APP), and the spliceosome subunit, U1-70K. To examine the proteins and signaling networks that interact with Mib1, APP, and U1-70K, and how they might relate to disease, we developed two novel interactomics methodologies exploring the affinity-stratified brain interactome of Mib1 and a highly controlled and affinity-stratified *in vivo* AD interactome of APP. Further, we thoroughly characterized the hippocampal proteomic perturbations at 3, 6, and 12 months in a newly developed transgenic mouse expressing N40K, a truncated form of U1-70K found in abundance in AD. Our findings greatly expanded the known interactomes of Mib1 and APP, established a role for Mib1 in dendritic spine development, and characterized its antagonistic interaction with the potent neurodevelopment regulator, CDKL5. We found numerous overlaps between N40K proteome perturbations and those seen in neurodegeneration (both in animal models and human), including strong decreases in the synaptic vesicle protein Synaptophysin. These methods represent improvements to the system-wide study of protein-protein interactions and the results greatly increase our understanding of the roles of these molecules, in addition to providing numerous avenues for future investigation.

Next Generation Proteomics in the Nervous System

By

Joseph L. Mertz
B.S., Cornell University, 2005

Advisor: Junmin Peng, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Neuroscience

2015

Acknowledgments

My utmost thanks go to my advisor, Dr. Peng, who taught me to look at the big picture by looking at all the minute details at the same time, and to be a happy, healthy person while doing so. To members of the Peng Lab and St. Jude Proteomics Core Ping-Chung Chen, Hong Wang, Yanling Yang, Zhiping Wu, Haiyan Tan, Chan Na, Drew Jones, Ash Mishra, Andy High, Vishu Pagala, Kanisha Kavdia, Kiran Kodali, Ji-Hoon Cho, Xusheng Wang, Tim Shaw, Yuxin Li, and especially Bing Bai, my lab brother, I want to thank you all for spirited scientific – and nonscientific – discussion, and for creating a great environment for me to learn how to be a better scientist and person. To my thesis committee Dr. Bassell, Dr. Caspary, Dr. Jin, and Dr. Zheng, thank you truly for your generous and remote yet ever-present guidance and your commitment to my scientific education. To the rest of the Emory Neuroscience community, of which there are too many amazing people to list here, I owe you my deepest thanks for welcoming me with an educational and personal experience so thoroughly life enriching that I will probably never be able to fully appreciate it. Thank you to Dr. Susan Culican, who first taught me what it means to be a scientist, and put me on the path to becoming one. To Christina, you are so beautiful, I just like to be around you and make you smile – and laugh hysterically. You are my constant motivation and my home. Thank you. To my family, to my sister Mary Jo, to my brother John, and especially to my parents John and Mary Anne Mertz, I can't thank you enough. You are a testament to the power of simple love. I feel like I can do anything just because you exist.

Table of Contents

List of Figures and Tables

CHAPTER ONE

Introduction

There are an estimated 3.72×10^{13} cells in the human body (1) of which 86×10^{9} are neurons (2) connected by an estimated 10^{14} synapses (3). Over 10,000 proteins are expressed in the brain (4), and there are approximately 3×10^{9} total protein copies in every cell (5). There are currently over 14,000 documented interactions between proteins (6), which of course does not include myriad interactions with DNA, RNA, and small molecules that constantly ebb and flow during cellular life. It is almost incomprehensible that this awe-inspiringly complex assembly of molecules, cells, and tissues can even function as a single unit. Yet it does function. It not only functions, it is self-aware and aware of its surroundings to the point of amassing a worldwide body of knowledge that has grown exponentially for 400 years, doubling in size roughly every 9 years (7), examining everything from the smallest aspect of itself to the farthest reaches of the universe.

The challenge faced by neuroscientists, cell biologists, molecular biologists, and scientists in many other fields is monumental. Nevertheless, the scientific community has proven to be up to the task, slowly building a reliable and reproducible interpretation of how cells, organs, and organisms function. As understanding of the complexity and nuance of the subject has grown, so has the realization that the scale of our approaches must grow to match. The human genome project is a prime example of this – it represents a tremendous international collaborative effort in attempts to systematically understand the entire breadth our genetic underpinnings (8). The impact of this effort has been invaluable and incalculable, providing a vast substrate for continued study in fields from molecular biology to evolutionary biology and adding, by some estimates, $1 trillion into the US economy (9). The influx of information was so great that the time before the sequencing of the genome has been referred to 'antediluvian', i.e. "before the flood" (10).

Genomics is only the tip of the iceberg (though contrastingly, a fundamental one) of the ongoing –omics revolution which broadens the scope of biological research to entire systems, at various levels. The genome begets the transcriptome, which begets the proteome which together with the metabolome and lipidome, beget the cytome, which in the brain begets the connectome, and so on. Each of these are rapidly evolving fields, that have huge potential to influence the way we diagnose and treat disease. My experiments over the course of my doctoral research have focused on the development of proteomics approaches to fill methodological gaps in the pursuit of our understanding of both

early neuronal development and neurodegeneration late in life. Our major points of focus were the protein network interactions of the E3 ubiquitin ligase Mib1, the Alzheimer's disease hallmark APP, and an aberrant form of the spliceosome U1-70K found in Alzheimer's disease.

Quantitative Proteomics

Proteomics, though complementary to genomics and transcriptomics in terms of overall examination of cellular activity, is uniquely powerful in that it focuses directly on the major effectors of the cell (11). After translation, and therefore beyond the direct control of DNA or RNA, the life of a protein can go in many directions, such as interfacing with other proteins, being trafficked to various locations inside and outside the cell, undergoing a huge range and combination of post-translational modifications (PTMs), and being subject to widely varying rates of turnover. Proteomic analysis is an ideal methodology to analyze these aspects of protein activity, especially on large scales. The most prevalent, and in many ways most powerful (12) methodology in large-scale proteomics entails tryptic digestion of protein samples into much smaller peptides, followed by liquid chromatography (LC) coupled online to an electrospray ionization (ESI). Analysis is performed by tandem mass spectrometry (MS/MS) in which the mass to charge ratio (m/z) of peptides are first measured, then fragmented by high energy collision with inert gas, and the m/z of the fragments measured again (13). This is referred to as bottom-up, or shotgun proteomics or by the acronym LC-MS/MS, as proteins are sequenced and identified by peptide fragments, making it the proteomic equivalent to

shotgun genome sequencing. In contrast, top down proteomics entails analysis of intact proteins, followed by fragmenting and reanalysis of fragments and is predominantly used in less complex samples in which it has particular advantages (14).

In the early days of proteomics, the major focus was on simply identifying protein analytes, but it gradually became clear that quantitative information was necessary to propel the field onward, and to match the varied and established applications of SDS-PAGE analysis. Early methods for protein mass spec quantification relied directly on the spectra of the peptides themselves, such as counting the number of peptide sequence matches (PSMs) attributed to each protein (15) – and relating this to abundance via specially designed algorithms (16). Quantification by this avenue was taken further by measuring the extracted ion current (XIC), or the area under the curve as a peptide elutes from the LC column (17). This has evolved and is still a popular approach as MS instrument sensitivity and resolution have made the method more accurate and reproducible. Furthermore, highly sophisticated algorithms have been developed to complement this approach (18) and it involves no comparatively expensive labeling of cells or samples used in other popular methods. The major drawback of these label-free quantification (LFQ) methods, however, is that analysis of multiple samples requires multiple MS runs, and run-to-run variability is inherently high. Widely used label-based quantitative methods address this and do allow for simultaneous MS analysis of multiple samples, thus greatly reducing variability, and further, reducing instrument time helps to offset the cost. The two

most commonly used label-based techniques utilize isotopes that are undetectable from the cell to the final LC steps but are discernible by the MS instrument. In stable isotope labeling in cell culture (SILAC), up to five cell culture samples are grown containing different isotope labeled amino acids, and relative levels of the isotopes representing each sample are measured by the MS (19,20). This method has been further expanded to the labeling of entire animals by stable isotope labeling in mammals (SILAM), in which the only protein source in the diet is isotope labeled to the point of 100% incorporation (21). In methods using tandem mass tag (TMT) labeling (22), which evolved from isotope coded affinity tags (ICAT) (11), labels are added post-trypsin digestion by highly efficient targeting of primary amines (n-terminals and lysines), and currently up to 10 samples can be compared. LFQ and label-based quantification both have advantages and drawbacks, as well as adherents and detractors, and both have already contributed significantly and hold great promise for the future.

Interactomics

Without examining interaction between proteins, even exhaustive identification and quantification of the proteome provides an incomplete picture of the state of a cell or tissue. In their incredibly diverse roles as effectors of cellular physiology, one of the most important roles proteins perform is to regulate and cooperate with other proteins; thus, analyzing these interactions is key to understanding the state and activity of a cell, tissue, or organ (23). Protein-protein interactions (PPIs) have long been studied using chemical cross-linking, affinity chromatography, co-immunoprecipitation (Co-IP), co-fractionation,

colocalization in immunochemical assays, and more recently by methods such as yeast 2-hybrid assays (Y2H) (24), protein-fragment complement assays (PCA) (25), and Fluorescence resonance energy transfer (FRET) (26). Each has particular strengths and weakness, for instance FRET is able to detect protein interactions in real-time and in a relatively unperturbed context, but requires ectopic expression of at least two fluorescent tagged-proteins and is essentially limited to analysis of those proteins of interest.

Roughly 15 years ago the field took a considerable leap and high-throughput systems-based approaches were developed using Y2H, and by coupling affinity-based techniques with mass spectrometry (AP-MS). The advent of these techniques helped carry the -omics revolution to the study of PPIs, thus initiating the study of the interactome, first named in 1999 (27). As the field has progressed, AP-MS has slowly begun to dominate as it has the potential to dissect an interactome directly from unperturbed tissue of interest (i.e. human disease brain tissue versus the nucleus of a yeast cell), it allows examination of entire complexes as opposed to binary interactions, and technological advances on several fronts have outpaced those related to Y2H assays (28).

Continually increasing sensitivity in LC-MS/MS technology (29,30) has allowed for the analysis of increasingly complex interactome samples (31) and the development of quantitative techniques, such as isobaric TMT labeling, aids in distinguishing specific interaction partners from nonspecific background (32). However, the extreme sensitivity of AP-MS, one of its key advantages, can also produce a vast number of false positive results, and a delicate balance lies

between eliminating nonspecific background and losing weakly bound, but truly functional, interaction partners (33). For instance, weak and transient interactions can be lost and reorganized during biochemical manipulation under native purification conditions (34). These issues are the primary drivers of continued technical refinement in the field. Tandem-affinity purification (TAP-MS) utilizes a doubly-tagged bait protein – which precludes it from most *in vivo* analysis – to limit nonspecific contaminants, but this also limits loosely bound interactors and requires larger quantities of starting materials (35). *In vivo* crosslinking with MS (XL-MS) helps to prevent loss of loosely bound interactors, but introduces a new potential source of nonspecific background (34,36). The state of the field therefore stands at exhaustive recovery of interacting proteins combined with sequential or multidimensional subdivision by various properties such as subcellular localization, binding affinity, confidence of interaction, and reproducibility. A recent methodological tweak aimed at this, affinity enrichment mass spectrometry (AE-MS), allows for both broad examination of an interactome and some differentiation by binding strength via more extensive use of experimental controls, but has yet to be implemented outside of yeast (37). Thus, there is substantial space, need, and potential for innovation in this field.

Ubiquitin Proteasome System, and the E3 Ligase Mind bomb 1

Mass spectrometry proteomics holds a rather unassailable advantage in the study of PTMs, particularly as they are beyond the scope of other systems-based approaches such as genomics and transcriptomics (12). Further, large scale proteomics analyses can detect multiple modification types with little to no

change in instrument settings, and can be greatly improved by fairly simple targeted purification methods (38). Phosphorylation sites can be identified by the characteristic 98 m/z shifts of (for H_3PO_4) and purified by TiO_2 columns, among others. Ubiquitination sites can be both purified and detected post-trypsinization by their tell-tale dual glycine (diGly) m/z shift, in which a diGly remnant from ubiquitin is attached to peptides from ubiquitinated proteins (39). With both of these, not only are the modifications identifiable, their location (within the modified peptide) is as well, and top-down methods can determine multiplicity and combination of modification sites within the same species (40). Additionally, modifications such as glycosylation (41), methylation, and acetylation are also easily amenable to MS analysis and represent rich fields of their own.

The ubiquitin proteasome system (UPS) holds particular interest as it is critically involved in almost every cellular process (42). It generally regulates degradation of protein targets, and is therefore a major factor in the abundance of most proteins, but it also initiates a variety of downstream protein behaviors such as changes in localization, endocytosis (43), and transcription regulation (44). Because of the aforementioned fact that ubiquitination is a post-translational modification and because the numbers involved are quite large - over 600 E3 ligases and system-wide protein tagging by ubiquitin (45) – proteomics is ideally suited to the study of this powerful pathway (39). Determining the substrates of a particular E3 ligase is a prominent biological question because it can potentially afford strong control over these substrates, which could have ramifications in disease treatment. Significant strides have been made, including the

development and systems-based implementations of an antibody targeting the signature diGly peptide (46). Considerable hurdles remain, however, including the fact that E3 ligases and their substrates are generally in contact for a short time, and that the system is very tightly regulated by the proteasome, deubiquitinases, and abundant expression of the multiple poly-ubiquitin genes themselves.

Mind bomb 1 (Mib1) is a potent E3 ubiquitin ligase which originally received its name from zebrafish mutagenesis screens in which Mib1 mutations caused excessive numbers of neurons during development (47). It was found to be a critical regulator of the Notch signaling pathway, as an essential activator of Notch Delta/Serrate/lag-2 (DSL) ligands via ubiquitination (47). In addition to its role in cell fate determination during early development, Mib1 is also abundantly expressed in the adult brain (48), regulates neurite outgrowth in postmitotic neurons (49), is highly enriched in the PSD fraction, and is important to long-term potentiation (LTP) and synaptic plasticity (48). Great advances have been made towards understanding the protein networks responsible for the varied aspects of neuronal morphogenesis, such as axon growth (50), dendrite formation (51), dendritic spine formation (52), and synaptogenesis (53), but large gaps still remain in the connections between these highly related processes. Mib1 is already known to interact with many proteins including SMN1, which contributes to spinal muscular atrophy (54), RYK, a receptor involved in the Wnt pathway (55), and CDK5 plus its activator p35 (49). The widespread importance of Mib1 in neuronal development, and the fact that the list of Mib1 interaction partners

continues to grow piece by piece, makes it a prime candidate for thorough systems-based interactome analysis.

Alzheimer's Disease and the Roles of APP

Alzheimer's disease is the 6th leading cause of death (56) and the cause of over 70% of dementia cases in the United States (57). While mortality from other major diseases – e.g. heart disease and cancer – is decreasing, AD mortality grew by 68% from 2000 to 2010 (56). In the coming decades, due to the overall aging of the population, it is projected that the number of AD sufferers in the United States will increase from roughly 4 million to 14 million (58) and that 43% of all elderly people who die in 2050 will have AD (59). In 2013 AD cost the US over $400 billion – with direct and indirect costs accounting for half of this total as even caregivers are deeply affected, suffering from lost wages, higher rates of depression, and declines in physical health (58). The modern view of AD describes a true disease continuum. The preclinical stages involves AD pathophysiological processes but the absence of evident symptoms (60), mild cognitive impairment (MCI) entails the emergence of predementia symptoms (61), and AD dementia involves declines from previous cognitive levels, significant functional impairment, impairment in learning and recall of recently learned information, language and visuospatial deficits, and executive dysfunction (62). Its prevalence (new cases plus duration) doubles every 5 years of life after 65 (63) and its incidence peaks near 7% for those over 95 (64).

The canonical pathological hallmarks of AD are intracellular neurofibrillary tangles (NFT) and extracellular amyloid plaques (65). Both are associated with

dystrophic neurites, synapse loss, and gliosis (66), and widespread cortical atrophy (67). NFTs are primarily filamentous polymers of the microtubule associated protein, tau and Amyloid plaques are primarily composed of Aβ, a 40 or 42 residue species proteolytically cleaved from the extracellular portion of APP. Numerous other gene actors have been associated with the disease via GWAS: polymorphisms in APOE, EPHA1, CR1, PICALM, SORL1, and MEF2C, among others are known to confer AD risk (68). Numerous Interactions between APP and other proteins have also been uncovered. For instance, the cleavage processes that transform APP into potentially pathogenic Aβ species are rather well characterized. Amyloidogenesis is aided by both γ–secretase, which includes presenilin-1/2 and cleaves at a position within the membrane bilayer, and β–secretase, or BACE1, which cleaves APP on the extracellular side. The α–secretase family competes with β-secretase to cleave APP extracellularly and is anti-amyloidogenic (69). The APP intracellular domain (AICD) is thought to translocate to the nucleus, and contains the NPXY motif which interacts with many proteins, such as the amyloid precursor binding A and B families (APBA and APBB) (70), JIP1(71), and GSK3β (72).

Despite the fact that the sheer magnitured of the combined effort of APP and AD analysis is nothing short of remarkable, our understanding of AD pathogenesis stands at a major deficit in comparison to the ongoing impact of this debilitating and deadly disease. Continued advances in interactomics technology and study design provide new opportunities to evaluate the physiological and pathogenic properties of APP.

This dissertation

The following dissertation presents fruits of projects focused on the expansion of methods examining protein-protein interactions and protein signaling networks in the brain on a systems-based scale. Dividing the experiments into four portions, two focused on the brain interactome of the broadly important E3 ligase Mib1 and E3 ligases more generally, and the other two focused on promising and critical molecular aspects of Alzheimer's disease. An emphasis on analysis by mass spectrometry proteomics underlies each portion, as this powerful methodology provides access to untapped depths of the molecular biology of the brain, and our group is exceptionally well suited to pursue this avenue of inquiry.

Our results greatly expanded the known putative interactome of Mib1, and stratified it by binding strength via novel techniques (Chapter Two). Our new contributions to the Mib1 interactome include several members of the catenin family, the deubiquitinase Usp9x, the endocytosis regulator Scyl2, and the kinase CDKL5. We revealed a previously undiscovered role of Mib1 in dendritic spine outgrowth, which involves ubiquitination of CDKL5. We developed and tested a promising systems-based ubiquitin ligase substrate tagging assay using a highly efficient bacterial biotin ligase to place a biotin tag on ubiquitin molecules in living cells, which are then attached to substrates of the E3 ligase of your choice and can be detected or purified downstream using the extremely strong streptavidin-biotin interaction (Chapter Three). This technique unfortunately faces highly challenging caveats in the context of live cells and will require extensive, if not

impossible, technical refinement if it is to be pursued moving forward. We devised and implemented a unique approach to examining the *in vivo* APP, and practically any other, interactome (Chapter Four). This approach stratified the interactome by affinity much the same as our experiments with Mib1, and also used comparisons against multiple control purifications, multiple test antibodies, and was performed using invaluable human AD tissue. Finally, we examined age-specific, overall, and age-related proteomic changes in a new transgenic mouse model overexpressing N40K, a shortened form of the spliceosome member U1-70K that is found in abundance in human AD brain (Chapter Five). The studies reported here forming just a portion of the overall analysis of this highly intriguing mouse model, we observed perturbations to pathways involving the actin cytoskeleton, the extracellular matrix, the complement cascade, ciliogenesis, and others. Several of these are highly similar to more traditional AD mouse models and suggest great promise for the dissection of the role of this molecule in AD pathogenesis.

CHAPTER TWO

Sequential Elution Interactome Analysis of the Mind Bomb 1 Ubiquitin Ligase Reveals a Novel Role in Dendritic Spine Outgrowth

Adapted from: *Molecular & Cellular Proteomics*

Joseph Mertz, Haiyan Tan, Vishwajeeth Pagala, Bing Bai, Ping-Chung Chen, Yuxin Li, Ji-Hoon Cho, Timothy Shaw, Xusheng Wang, Junmin Peng

Volume 14, Issue 7, 1 July 2015, Pages 1898-910

Mind bomb 1 (Mib1), an E3 ubiquitin ligase, is a critical regulator of metazoan development with a large but ever expanding number of functions through interactions with a variety of protein partners. Mib1 mutants were first found in zebrafish mutagenesis screens (47), in which the mutants had neurogenic defects, most notably supernumerary primary neurons, and additional deficits in the development of somites (73), ear (74), and vasculature (75). These phenotypes are predominantly the consequences of impaired Notch signaling, as Mib1 is an essential activator of Notch Delta/Serrate/lag-2 (DSL) ligands (47).

Mib1 also controls the development of several other organ and tissue systems, including gastrointestinal tract (76), limb bud (77), and the immune system (78). Mib1 is highly conserved across species. For instance, zebrafish Mib1 protein is 68%, 94% and 94% identical to its fly, mouse, and human orthologs, respectively (79). Moreover, Mib1 has a paralog (Mib2) that shares 38% identical protein sequence with Mib1 in mouse (80). Mib2 is only abundantly expressed in adult tissue, however and thus does not function in early development. Consistently, Mib1 knockout in mice results in embryonic mortality (43); Mib2 deletion has no obvious effect on mouse development (77).

In addition to its role in cell fate determination during early development, Mib1 is also abundantly expressed in the adult brain (48) and plays an important role in neuronal morphogenesis (49). Neurons usually have two basic polarized structures, a single extended axon for sending signals and multiple branched dendrites (or more precisely, the somatodendritic compartment) for receiving signals. Many principal neurons in mammals further grow dendritic spines that are tiny protrusions extended from dendritic branches, creating local postsynaptic compartments for the formation of excitatory synapses. In these synapses, the postsynaptic density (PSD) is an electron-dense membrane thickening aligned with the presynaptic active zone at synaptic junctions. During neuronal morphogenesis, axonal growth and path finding(50), dendrite formation (51), dendritic spine assembly (52), and synaptogenesis (53) are independent but highly related processes controlled by genetic elements and environmental cues. Although dramatic progress has been made in identifying the signaling cascades

responsible for these processes, large gaps still remain in the connection of individual signaling components as well as in the coordination of multiple pathways. Our previous proteomics analysis identified that Mib1 is highly enriched in the PSD fraction, and regulates neurite outgrowth in postmitotic neurons (49). Mib1 conditional knockout study suggests its role in long-term potentiation (LTP) and synaptic plasticity (48). In addition, further intriguing actions of Mib1 continue to be discovered. Mib1 was found to mediate the degradation of survival motor neuron 1 (SMN1), which contributes to spinal muscular atrophy (54). Mib1 was reported to be essential for Wnt3A activation of beta-catenin signaling through the receptor RYK (55), and a recent yeast two-hybrid screen indicated that Mib1 interacts with 81 candidate proteins beyond the canonical Notch pathway (81). The ongoing identification of new Mib1 interaction partners and functions underscores the need to characterize the Mib1 interactome *en masse* with high confidence.

The combination of affinity purification and liquid chromatography-tandem mass spectrometry (LC-MS/MS) has emerged as a powerful method for analyzing protein interaction networks. Technological advances in LC-MS/MS have continually increased the sensitivity of protein detection (29,30), allowing for the analysis of complex samples (82). The primary advantage of this technique, however, has also proven to be its greatest weakness: without stringent washes and data filtering, a vast number of false positives are included in the resulting datasets (33). Methods such as tandem-affinity purification (83) have been developed to remove nonspecific contaminants, but two-step purification requires

large quantities of starting materials and reduces sensitivity to loosely bound proteins. Removing contaminants by buffers containing high concentrations of salt and detergents can help limit false positives, but a delicate balance lies between rinsing contaminants and losing weakly bound but true interaction partners, and thus inflating false negative results. In addition, in vivo crosslinking and quantitative analysis are used to enhance the capture of transient interacting proteins (84,85).

To this end, we attempted to characterize the Mib1 interactome by combining glutathione S-transferase (GST) protein affinity purification and advanced quantitative mass spectrometry. In our sequential elution strategy, Mib1 interaction partners were bound to affinity resins coated with GST-Mib1 domains, then eluted in three sequential buffers of increasing stringency. Proteins in these three eluents were identified and quantified by an isobaric labeling Tandem Mass Tag (TMT) method (86). The elution profile of each protein reflected its binding affinity to the GST-Mib1 resins. The strategy not only provides high sensitivity to recover weakly bound partners, but also allows for the affinity-based classification of the interactome and the removal of contaminants. By this approach, we were able to recover 817 putative Mib1 binding partners in adult rat brain and accepted about half of the proteins with high confidence. This study also uncovered that Mib1 interacts with CDKL5, a protein kinase implicated in early infantile epileptic encephalopathy-2 (EIEE2), a severe form of epilepsy and mental retardation in females (87). We then found that Mib1 acts to downregulate CDKL5 and inhibits its promotion of dendritic spine outgrowth.

MATERIALS AND METHODS

Plasmids and Antibodies

For affinity purification analysis, GST-Mib1 domains containing residues 1-401 or residues 384-801 from Mib1 were cloned onto the 3' end of the GST tag within the pET21a bacterial expression vectors. For immunocytochemistry, western blot analysis, and dendritic spine morphology analysis, full length and C985S point mutant Mib1 were cloned onto the 3' ends of the HA tag in pcDNA3.1 vector and the GFP tag in peGFP vector. CDKL5 was cloned into the 3' end of the HA tag in pcDNA3.1. The Usp9x and alpha-, beta-, and delta-catenin constructs were gifts from Drs. Y. Zheng (Johns Hopkins) and A.P. Kowalczyk (Emory University). The antibodies used included HA rabbit polyclonal and mouse monoclonal antibodies (Abgent), mouse monoclonal GFP Ab and FITC-conjugated GFP Ab (Abcam), Streptavidin-HRP (Life Technologies),Usp9x, and catenin antibodies (Santa Cruz).

Affinity Purification and Sequential Elution

GST-Mib1 domain fusion proteins were expressed in *E. coli* and isolated by Glutathione-Sepharose resins (Amersham/Pharmacia). The resins coated with GST fusion proteins were directly utilized to prepare affinity columns (1V = 1 bed volume, with 1 ml resin containing at least 1 mg of protein). Highly concentrated rat brain lysate (20V, ~10 mg/ml protein, in buffer A: 20 mM Hepes, pH 7.2, 0.1 M NaCl, 0.1% Triton X-100, 1 mM DTT, 1 mM EDTA, 1 mM EGTA, 15% glycerol and protease inhibitors) was prepared, pre-cleared by ultracentrifugation

(~200,000 *g* for 1 h), and loaded on the columns. The columns were extensively washed with buffer A (40V, 20V per fraction), then sequentially eluted with buffer B (the same as buffer A except 0.3 M NaCl, 4V, ~0.5V per fraction), buffer C (the same as buffer A except 0.6 M NaCl and 0.5% Triton X-100, 4V, ~0.5V per fraction), buffer D (the same as buffer A except 2 M NaCl and 2% Triton X-100, 4V, ~0.5V per fraction), and finally cleaned with 2% SDS (4V). The fractions were analyzed by an SDS gel followed by silver staining.

TMT Labeling of Digested Peptides

Equal volumes of the eluents from each buffer condition were pooled to form two duplicated samples: 3 buffers × 2 replicates = 6 total samples. The samples were mixed with 5x loading buffer (20% Ficoll, 10% SDS, 50 mM Tris-HCl, pH 6.8, 0.1% bromophenol blue) and fresh DTT to 2 mM, heated at 90°C for 5 min, and then cooled to room temperature. Fresh iodoacetamide was added to 20 mM for 15 min alkylation at room temperature in the dark. The samples were run on an SDS gel until the dye front was 3 mm past the stacking gel border to ensure complete stacking but little separation (88). After visualization via GelCode Blue (Pierce Biotechnology) staining, protein bands were excised and cut into 1 mm^3 pieces for standard in-gel digestion, except that 5 mM HEPES buffer (pH 8.5) was used to replace 50 mM ammonium bicarbonate to avoid amine group reactions with TMT reagents downstream. The digested peptides were extracted from the gel pieces, dried and resuspended in 20 µL 50mM HEPES (pH 8.5) for labeling with the 6-plex TMT labeling kit (Thermo Scientific). Briefly, six peptide samples were labeled with isobaric TMT tags: TMT 126 and

129 for Low Stringency Elution (buffer B), TMT 127 and 130 for Medium Stringency Elution (buffer C), and TMT 128 and 131 for High Stringency Elution (buffer D). Labeling efficiency of each sample was verified by MS analysis. The labeling reaction was quenched by 5% hydroxylamine, pooled and dried down by SpeedVac. Finally, the sample was desalted using a Ziptip (Millipore), then dried and dissolved in 5% formic acid for LC-MS/MS analysis.

Long Gradient LC-MS/MS Analysis of TMT Labeled Peptides

The analysis was essentially based on an optimized long gradient LC-MS/MS system (89). TMT labeled peptides were loaded on a long C18 column (~1 m × 75 µm) packed with 1.9 µm resin (Dr. Amish GmbH, Germany), and eluted during a 9 h gradient (~0.15 µl/min; 20%-55%; buffer A: 0.2% formic acid, 5% DMSO; buffer B: 0.2% formic acid, 5% DMSO, and 65% ACN). The column was heated to 65^{o}C by a butterfly portfolio heater (Phoenix S&T) to reduce backpressure. The eluted peptides were analyzed on an Orbitrap Elite MS (Thermo Fisher Scientific) with one MS scan (30,000 resolution, 1×10^{6} automatic gain control, and 100 ms maximal ion time) and top 10 high resolution MS/MS scans (HCD, 5×10^{4} automatic gain control, 200 ms maximal ion time, 2 *m/z* isolation window, 37 normalized collision energy, and 30 s dynamic exclusion). The long gradient LC-MS/MS run was repeated once for the sample.

Acquired MS/MS raw files were converted into mzXML format and searched by Sequest algorithm (version 28 revision 13) against a composite target/decoy database (39,90) to estimate false discovery rate (FDR). The target protein database was downloaded from the Uniprot rat database (28,863 protein

entries) and the decoy protein database was generated by reversing all target protein sequences. Spectra were searched with ± 20 ppm for precursor ion and product ion mass tolerance, fully tryptic restriction, static mass shift for alkylated Cysteine (+57.02146) and TMT-tagged N-terminus and Lysine (+229.162932), two maximal missed cleavages, and three maximal modification sites. Only *a*, *b*, and *y* ions were considered during the search. Assigned peptide spectra matches were first filtered by MS mass accuracy (±4 standard deviations, ~2 ppm, which was determined by all empirical good matches of doubly charged peptides with Xcorr at least 2.5). These good matches were also used for global mass recalibration prior to the filtering. The survived matches were grouped by precursor ion charge state and further filtered by Xcorr and ΔCn values. The cutoff values for XCorr and ΔCn were adjusted until a protein FDR lower than 1% was achieved. If one peptide was matched to multiple proteins, the peptide was represented by the protein with the highest peptide-spectrum matches (PSM) according to the rule of parsimony. Similar results were obtained using the JUMP program, a tag-based hybrid search engine recently developed (91). Raw data are available via the PRIDE database (www.proteomexchange.org, project accession: PXD001255).

Protein Quantification by TMT Labeled Peptides

Quantification of TMT labeled peptides was carried out by an in-house program in the following steps. (i) TMT reporter ion intensities of each identified PSM were extracted and recorded. (ii) The raw intensities were corrected according to isotopic distribution of each labeling reagent. For instance, the

TMT126 reagent produced 91.8% 126 *m/z* reporter ion, 7.9% 127 *m/z* reporter ion, and 0.3% of 128 *m/z* reporter ion. (iii) The average of all six reporter ion signals was used as a reference to compute a relative intensity between each sample and the average. (iv) The relative intensities of PSMs were averaged for identified proteins. (v) The average reporter ion intensities for Low, Medium or High Stringency eluents were compared to derive the $\log_2$ ratios. (vi) To analyze experimental variations, the intra-sample comparisons between technical duplicates were viewed as null experiments, showing an average standard deviation of 0.32. (vii) Finally, we selected a $\log_2$ ratio cutoff of 0.8 (~2.5 fold of the null standard deviation) for comparing different samples. The eluted proteins were compared in a sequential fashion (L→M and M→H), resulting in the values of $\log_2$(M/L, Medium vs. Low Stringency) and $\log_2$(H/M, High vs. Medium Stringency).

Interaction Network Analysis

Enrichment of Kyoto Encyclopedia of Genes and Genomes pathways (KEGG, (92)), and Gene Ontology terms was determined by analyzing the dataset using the NIH DAVID Bioinformatics Resources 6.7 Functional Annotation Clustering tool (93). STRING-DB (94) was used to evaluate interconnectivity between members of the pathways and processes determined to be enriched by DAVID.

Protein Preparation and Western Blot Analysis

For preparation of protein extract from HEK293 cell culture, cells were rinsed, dislodged, and transferred to chilled centrifuge tubes with ice-cold PBS. Cells were centrifuged at 21000 × *g* for 30 s at 4°C and lysed in 1X LDS sample buffer (Life Technologies) with cOmplete protease inhibitor cocktail (Roche Applied Science) and 10 mM DTT. The lysates were subsequently sonicated 6 × 3 s at 25% amplitude at 4°C. An aliquot of the total cell lysate was used for SDS-PAGE and Western blotting.

In vitro Ubiquitination Assay

Recombinant GST-Mib1 was expressed and purified from *E. coli* as previously reported (49); and HA-CDKL5 was expressed and purified from HEK293 by immunoprecipitation with magnetic anti-HA beads (Pierce Biotechnology). Ubiquitination was performed according to manufacturer's specifications using a Ubiquitinylation kit (Enzo Life Sciences). Briefly, HA-CDKL5-bound beads was incubated with biotinylated ubiquitin in ubiquitination buffer containing DTT and ATP, with the addition of different combination of the UBA1 E1, the UBCH5B E2, the GST-Mib1 E3 in a total reaction volume of 20 µL at 37°C for 1 h. The reaction was then quenched by EDTA (5 mM). The proteins were extracted and resolved by SDS-PAGE followed by Western blot or LC-MS/MS analysis.

Primary Hippocampal Neuron Culture and Transfection

Embryonic rat brains were harvested from E21 pups, rinsed in ice-cold DPBS, and placed in ice-cold complete Hibernate E media (BrainBits LLC).

Hippocampi were dissected, rinsed, and resuspended in 30°C Papain enzymatic solution (Worthington Biochemical) for 5 min. The tissue was spun briefly, rinsed and triturated using a fire-polished Pasteur pipet into Hibernate E. Finally, the cells were collected by centrifugation, resuspended in NBActiv4 (BrainBits LLC), plated at 1×10^5 cells per well on glass coverslips in 24-well plates and maintained at 37°C and 5% CO_2. At day in vitro (DIV) 7, the neurons were transfected via $CaPO_4$ precipitation for morphological analysis (95), with 2 µg plasmid DNA per well, including 0.5 µg of EGFP plasmid, 0.75 µg of each experimental plasmid as indicated, and empty pcDNA3.1+ plasmid to equalize total DNA to 2 µg.

Immunocytochemistry

At DIV 14, 7 days post transfection, primary neuron cultures on glass coverslips were prepared for immunocytochemistry, confocal microscopic imaging, and subsequent morphological analysis of dendritic spines. All cells were harvested, fixed in ice-cold 4% paraformaldehyde in phosphate buffered saline (PBS) for 25 min, permeabilized with 0.05% Triton X-100 in PBS for 5 min, incubated with blocking solution (3% normal goat serum in PBS) for 40 min at room temperature, and then incubated with FITC-conjugated anti-GFP antibody (Abcam) in 3% blocking solution for 1 h. The glass coverslips were washed, mounted, and sealed using ProLong Gold mounting media (Life Technologies) for confocal imaging. The transfection and staining analyses in other cells (e.g. HEK293) were performed by similar methods.

Dendritic Spine Morphological Analysis of Primary Hippocampal Neurons

Neurons were imaged on a Nikon (Tokyo, Japan) TE2000 C2 laser scanning confocal microscope. Images were acquired using a 60x/1.45 NA oil-immersion objective at 2048 x 2048 pixel (0.105 µm/pixel) resolution, and z-stacks were collected at 0.25 µm intervals. For analysis of dendritic spine morphologies, secondary dendrites were traced manually in ImageJ (National Institutes of Health, version 1.44p 64-bit) and spines were counted and assigned into morphological categories: stubby/mushroom-shaped, thin-headed, and filopodia-like. Stubby/mushroom-shaped spines were defined as having a head width much greater than spine neck width and spine width approximately equal to spine length. Thin-headed spines exhibited a head width greater than spine neck width and spine length much greater than spine width. Filopodia-like spines were defined as those with spine length much greater than spine width and showing no prominent head. All analyses were performed in a blind manner. Totals for each category along traced lengths of dendrites were combined to calculate the overall number of spines per 10 µm for spine density analysis. Statistical significance was tested using Student's *t* test for unpaired samples.

RESULTS

Mib1 Affinity Purification from Rat Brain and Sequential Elution

To achieve a comprehensive Mib1 interactome in the brain, we combined Glutathione S-transferase (GST) fusion protein affinity purification with quantitative mass spectrometry. The Mib1 protein can be divided into three main

domains (**Fig. 1A**): the N-terminal domain that contains five kelch repeats and one ZZ zinc finger and mediates substrate binding, the middle domain composed of 12 structural ankyrin repeats, and the C-terminal domain that possesses three RING finger motifs for ubiquitin E3 ligase catalytic activity (49). Our previous small scale study showed that only the N-terminal domain binds a large number of proteins (49). We further confirmed this result in a test affinity purification analysis, in which the N-terminal domain yielded a large array of proteins, whereas the middle domain isolated only one visible band shown on a stained SDS gel (**Fig. 1B**). Thus, we determined to focus on the interactome analysis of the N-terminal domain of Mib1.

Instead of using one-step elution used in the majority of AP-MS experiments, we sequentially eluted Mib1 binding proteins with increasing buffer stringency to differentiate strong and weak binding partners from co-purified contaminants (**Fig. 1C**). Nonspecific binding to either bait proteins or columns produces false positives in affinity purification experiments, and proteins that are more strongly bound - and therefore resistant to elution by buffers of low salt and detergent concentrations - are more likely to be genuine interaction partners. Despite this, simply using more stringent washes to remove weakly bound proteins without detection may result in a misleading number of false negatives. To solve this issue, we designed three sequential elution steps and analyzed all eluted proteins quantitatively by mass spectrometry, which allowed the evaluation of protein binding affinity. As visualized in the gel image (**Fig. 1D**), distinct subsets of proteins, both in identity and abundance, were eluted in the

Low and Medium stringency buffers, and a remaining few proteins were recovered primarily in the final elution by the High stringency buffer. Moreover, protein patterns in all three eluents were vastly different from that in the onput (i.e. rat brain lysate), indicating the enrichment of a subset of proteins by the Mib1 N-terminal domain.

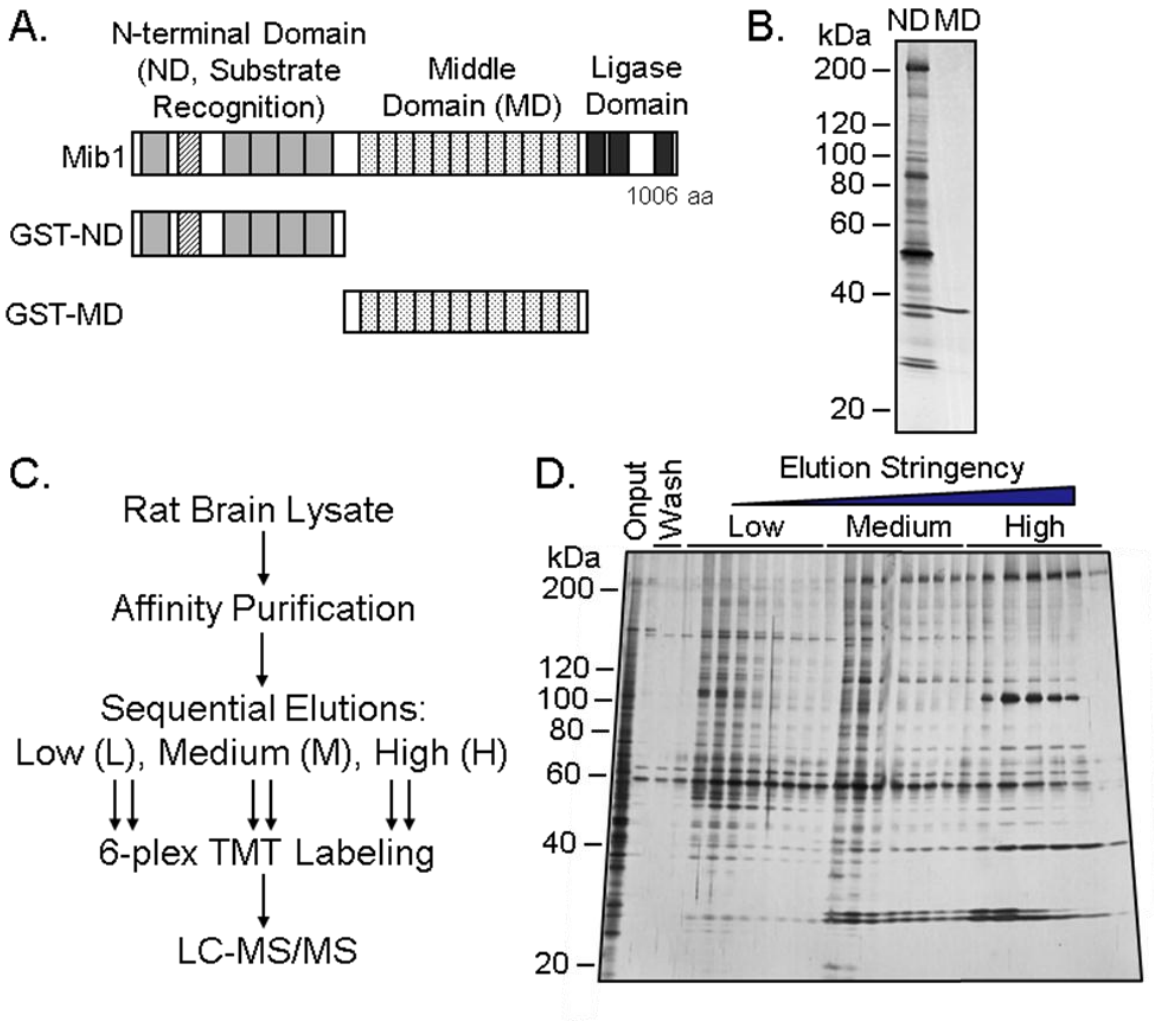


Figure 2.1. Sequential elution strategy of Mib1 affinity purification.

(A) Domain structure of Mib1 and recombinant proteins, depicting full-length Mib1 at 1006 aa, the ND (1-401aa), and MD (384-801aa). Mib1 contains 5 Kelch repeats (gray boxes), 1 ZZ-type zinc finger domain (diagonal lined boxes), 12 ankyrin repeats (dotted boxes), and 3 RING domains (black boxes).

(B) Gel analysis of total elution from ND and MD. Total elution from ND recovered many more proteins than MD.

(C) Overview of the purification, sequential elution, TMT labeling, and LC-MS/MS strategy. Rat brain lysate was incubated with GST fusion protein beads, washed and eluted. Eluents were digested and labeled by TMT isobaric tags, mixed, and analyzed by LC-MS/MS.

(D) Sequential elution from ND. Gel analysis of sequential elutions from ND demonstrates disparate band profiles for each buffer, including protein differences and intensities, as well as overall abundance.

Sequential Elution Profiling of Mib1 Affinity-Purified Proteins by Isobaric Labeling

We then profiled the Low, Medium, and High stringency eluents by the 6-plex isobaric TMT strategy with replicates. The six samples were matched by equal volumes rather than protein amounts, allowing comparison of relative protein abundances eluted in each fraction. Briefly, the six samples were run on a short SDS gel to remove detergents and salt, followed by in gel digestion that was modified to remove all amine-reactive buffers. The resulting peptides were fully labeled by TMT reagents, pooled and analyzed twice on long gradient LC-MS/MS (~1 m x 75 μm, 9 h gradient). After database search, a total of 817 unique proteins were identified with a false discovery rate of approximately 1%, and their protein abundances were further obtained by the corresponding TMT tag-derived reporter ions. Nearly all proteins show variable levels among the three elution conditions (**Supplemental Table S1**).

Global histogram analysis of the three conditions also indicated large differences (**Fig. 2A**). For example, the intra-group replicate comparison (e.g. M:M) of $\log_2$ ratios displayed a normal distribution centered at zero. In contrast, the inter-group comparisons (e.g. M:L or M:H) had small overlap with the M:M comparison, reflecting large differences in protein composition of the three samples. To assess the experimental variations, we treated the intra-group comparisons as null experiments and found that the averaged standard deviation (SD) of $\log_2$-ratios was 0.32. We then selected a $\log_2$-ratio cutoff of 0.8 (i.e. 2.5 fold of the null SD) that was outside a 99% confidence interval from the mean of the normal distribution. The eluted proteins were compared in a sequential

fashion (L→M and M→H). As each inter-elution comparison resulted three possibilities: up, no change, and down, we delineated 9 *elution group profiles* in the two comparisons (M:L and H:M, **Fig. 2.2B**).

Proteins in group 1 (*n* = 6) and group 4 (*n* = 2) were primarily eluted by the High stringency buffer (**Fig. 2C**), indicating that these proteins were firmly bound to the column and had the highest likelihood of interaction. Because of the stringency of the sequential elution profiling, only 8 proteins (1% of identified proteins) were classified into these two groups, including Dll4, a Notch ligand previously known to interact with Mib1 (43); Scyl2, an endocytotic protein that induces internalization of the Wnt receptor Fzd5 (96); Gbas, a possible regulator of vesicular transport (97); and Usp9x, a deubiquitinating enzyme that stabilizes Smn1 (98) and functions in neuron migration and axon growth (99).

Proteins in group 2 (*n* = 48) were recovered equally by the Medium and High stringency buffers, whereas proteins in group 3 (*n* = 335) were eluted predominantly by the Medium stringency buffer, showing slightly weaker affinity than those in group 2. Group 2 includes Nips1, a paralog of Gbas (also termed Nips2), Pde4d, an enzyme that mediates memory through cAMP degradation (100) and Syngap1, a major component of the PSD involved in dendritic spine formation(101). Group 3 includes Cdk5, a regulator of Mib1 level and neuronal morphogenesis; and Smn1 (98), an mRNA processing protein contributing to spinal muscle atrophy (**Fig. 2.2C**).

Proteins in group 5 (*n* = 110) were eluted at almost equal levels by all three buffers, while proteins in group 6 (*n* = 263) eluted at similar levels in the

Low and Medium stringency buffer but decreased significantly in the High stringency elution. Groups 8 (*n* = 19) and 9 (*n* = 34) components were eluted largely by the first, Low stringency buffer. Collectively, these proteins are unlikely to be direct interaction partners, and the majority was simply background contaminants that would be false positives if all purified proteins were eluted together by a single buffer. Nonetheless, some proteins in these four groups might exhibit functional, but very weak affinity to Mib1, or bind Mib1 indirectly through strongly bound partners. For example, Snx5 in group 5 colocalizes with the zebrafish ortholog mind bomb in early endosomal compartments. Another group 5 protein, Beta-catenin (i.e. Ctnnb1), is a component of the Wnt pathway. Group 6 includes Ctnnd2, another member of the catenin family, and Synapsin II which associates with synaptic vesicles. Group 8 contains Notch signaling component Numb1; and group 9 contains Ap2m1, which is a component of clathrin-mediated endocytosis.

Theoretically, proteins in group 7 (*n* = 0) would have been bound very weakly to be eluted by the Low stringency buffer, resistant to the Medium stringency buffer, and then also eluted at high levels by the High stringency buffer. This is unlikely, if not impossible in theory. Consistently, we found that no proteins fit this theoretical profile, which supports our assumption that proteins would elute reliably based on buffer salt and detergent concentrations in our sequential elution strategy.

In summary, the affinity of Mib1-interacting candidates can be estimated from the above eight possible elution profiles: group 4 > 1 > 2 > 3 > 5 > 6 > 8 > 9

(**Fig. 2.2C**). Groups 4, 1, and 2 were accepted as high affinity binding partners, group 3 were assigned as low affinity binding partners, and the remaining four groups were likely false positives (e.g. co-purified background proteins) or extremely weak binding proteins. Therefore, we essentially accepted proteins identified in group 1-4 as candidates in the Mib1 interactome for subsequent studies.

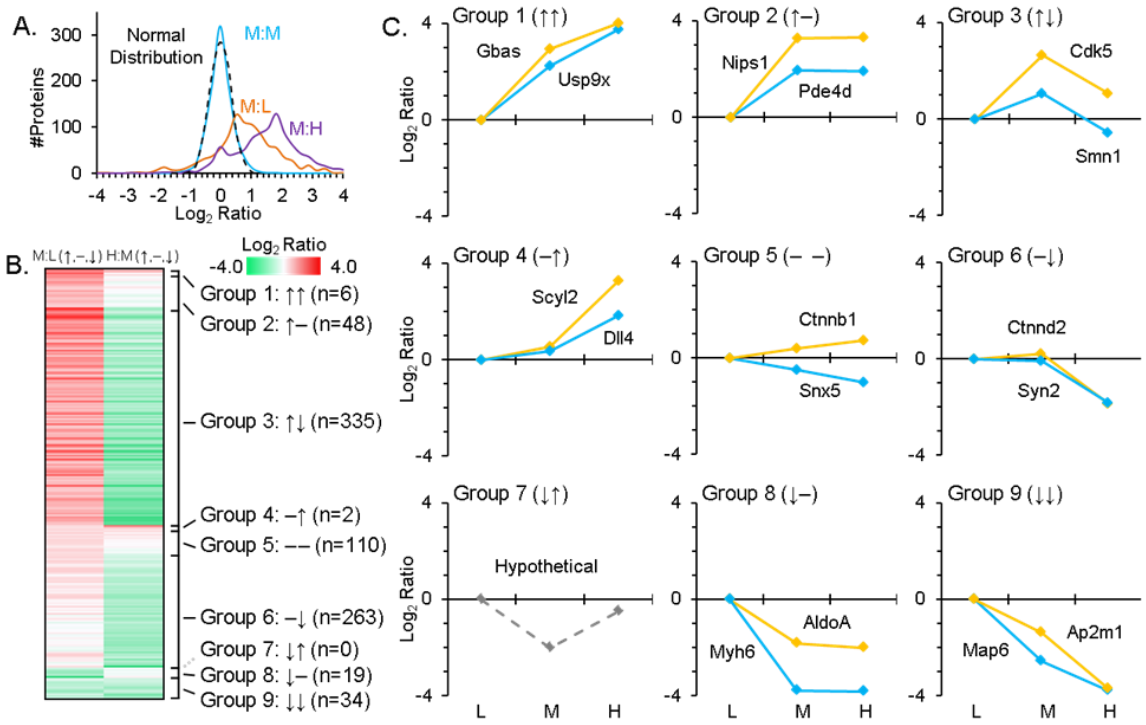


Figure 2.2. Grouping proteins by their sequential elution profiles.

(A) Intra-elution null comparison and inter-elution experimental comparisons. Null comparison of protein abundances from Medium Stringency buffer technical replicates shows a tight distribution around a $\log_2$ value of zero approaching a normal distribution. Meanwhile experimental comparisons between Medium and Low Stringency buffers and between Medium and High Stringency buffers display widely varied distributions, shifted towards positive values suggesting increased overall abundances in the Medium Stringency elution.

(B) Heat map showing $\log_2$ values for each protein from Low to Medium elutions and Medium to High elutions. Using a log2 value of 0.8 as a threshold, proteins exhibiting increases, decreases, and no change from the Low Stringency elution to the Medium Stringency elution as well as Medium to High were determined.

(C) Detailed elution profiles of representative proteins from each group.

To better understand Mib1 function in the brain, pathway and molecular function analysis by DAVID functional analysis software (102) revealed that Mib1-interacting proteins participate in a wide variety of biological processes: KEGG pathways (92) enriched included Spliceosome ($p < 0.001$), RNA degradation ($p < 0.001$), Gap junction ($p = 0.02$), LTP ($p = 0.03$), Neurotrophin

signaling (p = 0.05), and Adherens junction (p = 0.05, **Supplemental Table S2**) . GO terms enriched in Groups 1 through 4 included several related to RNA processing, chromatin assembly, cellular respiration, cellular morphogenesis and development, cell migration, cell adhesion, synaptic transmission and plasticity, and protein catabolism processes. We emphasized 5 key processes enriched in the Mib1 interactome, divided by groups 1 to 4 that showed relative binding affinity to Mib1 (**Fig. 2.3**), including Notch and Wnt pathways, endocytosis and vesicle transport, ubiquitin-proteasome system, morphogenesis, and synaptic activities. Further elaboration of these pathways using STRING-DB and including members of groups 5 to 9 shows high degrees of interconnectivity (**Fig. 2.4, Supplemental Fig. S1-S4**), and possible means of indirect binding for many group 5 to 9 members.

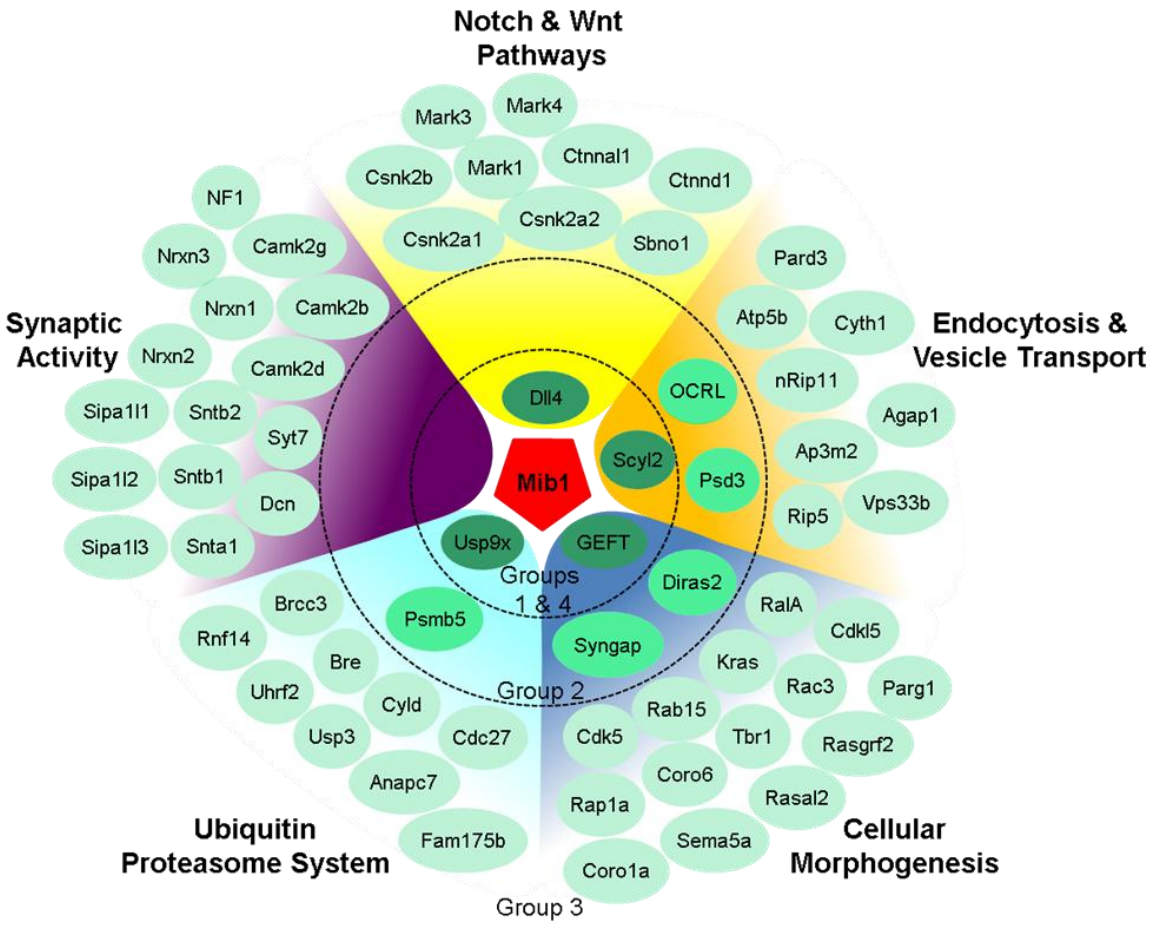


Figure 2.3. Mib1 interaction partners participate in several important signaling pathways.

Highlight of several important pathways and biological functions enriched in Groups 1 through 4 in our dataset. Notch signaling showed a modest enrichment, with an isoform of canonical Mib1 interaction partner delta, DLL4, exhibiting strong binding affinity and appearing in group 4. The Wnt pathway, however, was strongly enriched, with several Catenin and Casein kinase family members appearing in Group 3. Proteins involved in Endocytosis and Vesicle Transport and Morphogenesis were very strongly enriched across all four elution profile groups, with clathrin-mediated endocytosis regulator Scyl2 and RhoA interaction partner GEFT showing highest affinity binding. The Ubiquitin Proteasome System was also highly enriched across elution profile groups and the Usp9x deubiquitinase displayed one of the strongest elution profiles in addition to high abundance in our analysis. While many proteins located at Synapses and regulating LTP were recovered in our analysis, none exhibited extremely high binding affinity.

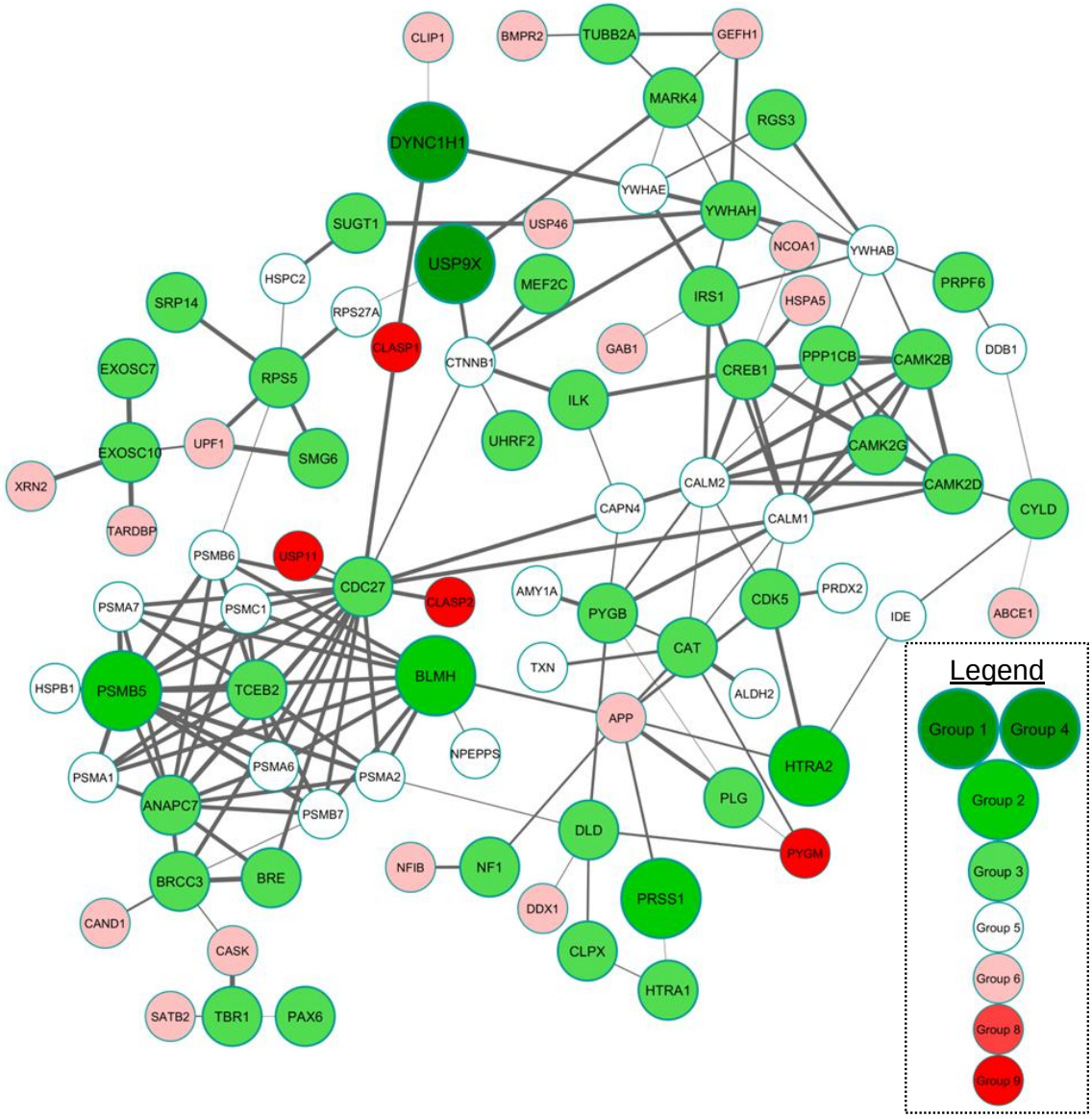


Figure 2.4. Interconnectivity in Mib1 Ubiquitin Proteasome System interactome.

Potential Mib1 interaction partners in the UPS pathway show high levels of interconnectivity. Several constituents of the Proteasome itself are included, with Proteasome Subunit Beta 5 showing highest affinity. The deubiquitinating enzyme USP9x is shown to interact with several proteins of interest as well. Edge thickness represents 'combined score' from STRING-DB analysis.

Mib1 Interacts with Usp9x and Catenin Family Members

One of the most strongly bound and highly abundant proteins in our proteomics analysis, Usp9x, is a deubiquitinating enzyme that regulates embryonic development (103). The dataset also included several members of the Catenin family, which participate in the Wnt signaling pathway and regulate cell adhesion (104). We analyzed these interactions with Mib1 via immunocytochemistry. When recombinant Mib1 was expressed in HEK293 cells, Mib1 colocalized almost entirely with endogenous Usp9x in puncta throughout the cytoplasm (**Fig. 2.5A-5C**). When Mib1 was co-expressed with alpha-catenin (a member of group 3), the two proteins also colocalized strongly throughout the cytoplasm (**Fig. 2.5D-F**). Interestingly, Mib1 staining pattern was drastically altered (comparing **Fig. 2.5A-5D**), showing the dispersal of Mib1 from puncta. Similar results were obtained with beta-catenin (in group 5) and delta-catenin (i.e. p120, also in group 3, **Fig. 2.5G-L**), but beta-catenin appeared to have less influence on the Mib1 staining pattern (comparing **Fig. 2.5A**-**G**). The apparent distinction in the strength of Mib1 pattern alterations seen in conjunction with alpha- and delta-catenin expression versus beta-catenin expression may be due to the strength of their interaction, as suggested by our sequential elution strategy: alpha- and delta-catenin in group 3 have higher Mib1 affinity than beta-catenin in group 5.

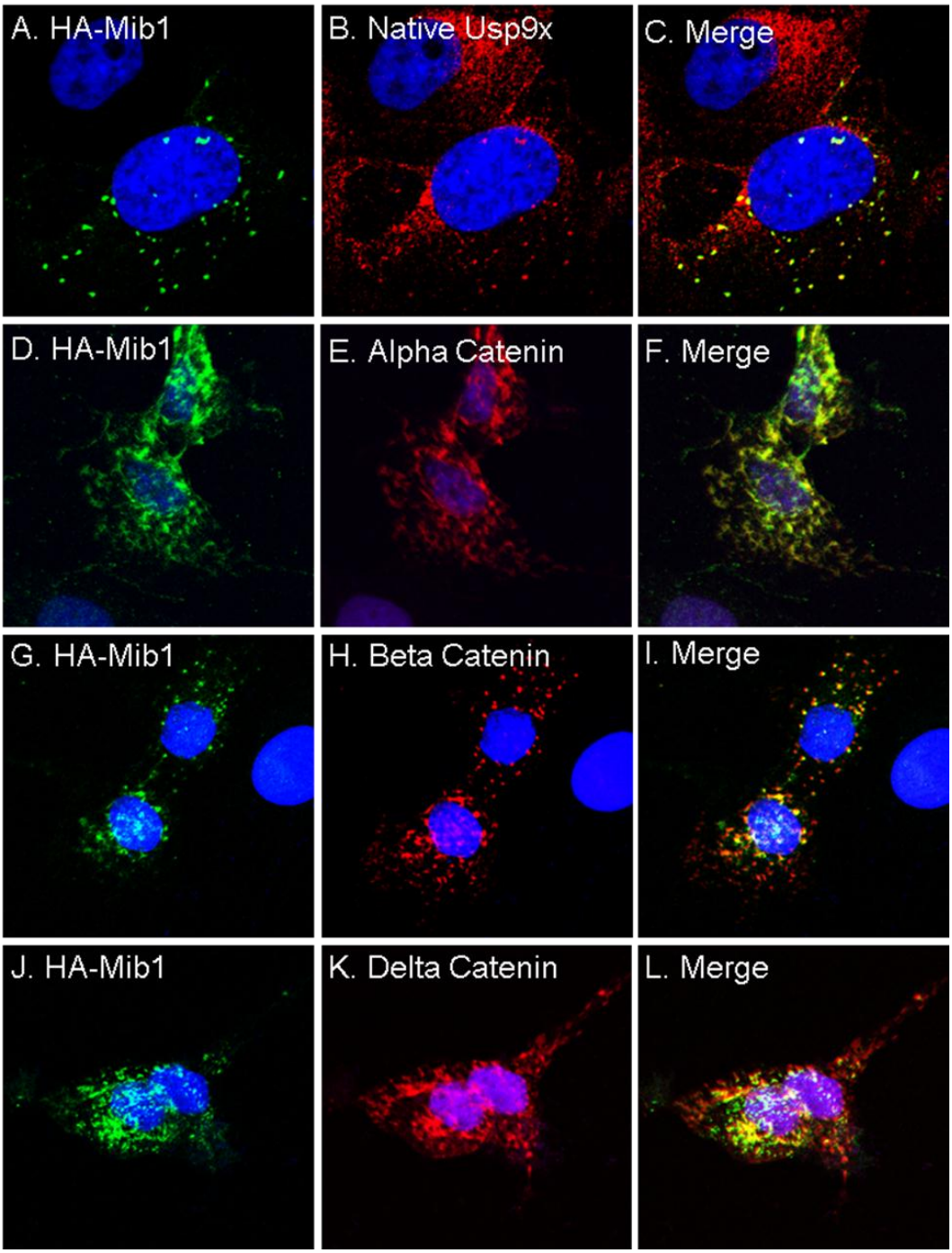


Figure 2.5. Mib1 colocalizes with Usp9x (FAM) and 3 members of the catenin family.

HEK 293 cells were transfected with HA tagged Mib1, and/or other catenin proteins, followed by immunofluorescence staining.

(A-C) Immunostaining of recombinant Mib1 and native Usp9x.

(D-L) Coexpression of HA-Mib1 and different catenin family members and subsequent immunostaining.

Mib1 Ubiquitinates CDKL5 and Alters Its Localization, Abundance, and Functional Effects on Neuron Morphogenesis

Our Mib1 interactome analysis also revealed CDKL5 in group 3 (**Fig. 2.6A**), which is known to regulate neuronal morphogenesis (105) and its genetic mutations cause early infantile epileptic encephalopathy-2, a severe form of mental retardation with defects in neurodevelopment (106). When expressed in HEK293 cells, CDKL5 exhibited primarily nuclear localization (**Fig. 2.6B, C**). In contrast, simultaneous expression of Mib1 caused a drastic shift of CDKL5 out of the nucleus towards large cytoplasmic puncta, where the two proteins colocalized strongly (**Fig. 2.6D-F**). Moreover, Mib1 coexpression also led to strong downregulation of CDKL5 in a dose-dependent manner (**Fig. 2.6G**). This downregulation was abolished, and indeed may have been reversed by a point mutation (C985S) that renders the Mib1 ligase domain inactive (49) (**Fig. 2.6H**), suggesting that the Mib1 ligase activity is required for the induced change in CDKL5 abundance, likely through ubiquitination and proteasomal degradation. Equal loading was determined by Ponceau S staining of membranes (supplemental **Fig. S5**). We also tested if Mib1 can directly ubiquitinate CDKL5 by in vitro ubiquitination assay. Recombinant CDKL5 and Mib1, mixed with E1, E2 enzymes, and Ub, formed a distinct Ub-positive smear with molecular size larger than CDKL5 (110 kD). Without the addition of either E1/E2 enzymes or the Mib1 ligase, no Ub-positive smear was detected. MS analysis of the smear indicated the ubiquitination of CDKL5, and self-modification of Mib1 and E1 enzymes (**Fig. 2.6I**). In addition, a slight increase in the CDKL5 level was

observed with the overexpression of the Mib1 C985S mutant (**Fig. 2.6H**), suggesting a potential dominant negative effect of C985S. Following this clue, affinity pull-down and immunocytochemical analysis showed that Mib1 binds to and colocalizes with its own C-terminal domain (**Fig. 2.6J-K**). Therefore, Mib1 may self-associate via this domain and over-expressed C985S may form a complex with endogenous Mib1 to disrupt its activity. The data provides additional evidence for the downregulation of CDKL5 by Mib1.

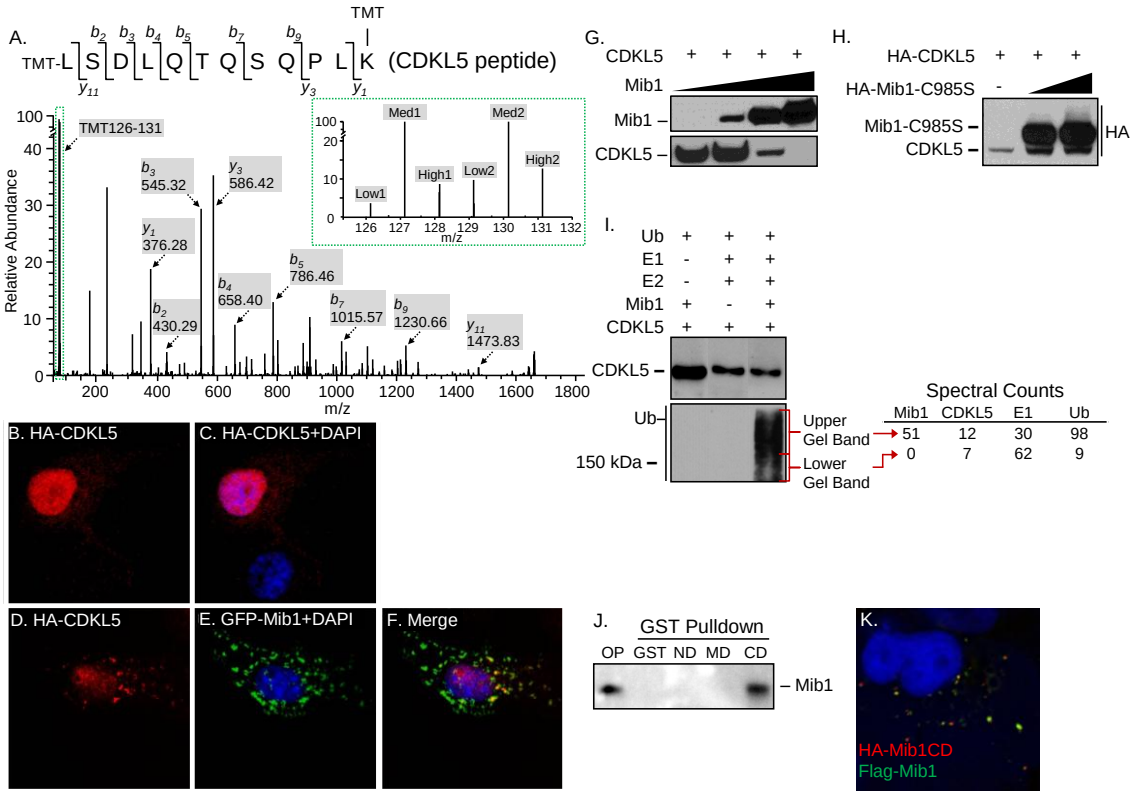


Figure 2.6. Mib1 colocalizes with and downregulates CDKL5.

(A) Representative MS2 spectrum depicting peptide 826-837 from CDKL5 protein and TMT reporter ion intensities (inset).

(B-C) Typical HA-CDKL5 localization in HEK293 cells is largely nuclear when expressed alone.

(D-F) Coexpression of HA-CDKL5 along with GFP-Mib1 led to redistribution of CDKL5 from the nucleus to puncta in cytoplasm.

(G-H) Western blot analysis of HA-CDLK5 and GFP-Mib1 co-transfected HEK293 cells, dependent on Mib1 ligase activity. C985S is a mutated form of Mib1 with abolished ligase activity.

(I) Western blot and LC-MS/MS analysis of ubiquitination of CDLK5 by Mib1.

(J-K) Western blot and immunofluorescence staining analysis show self-association of Mib1 through its C-terminal domain.

We further tested functional interaction of Mib1 and CDKL5 during neuronal morphogenesis (**Fig. 2.7**). The two proteins were reported to have opposing effects on neurite outgrowth (49,107). The literature is somewhat

conflicting in regards to the effects of CDKL5 on overall dendritic spine density – Ricciardi et al. reported that shRNA knockdown of CDKL5 results in increased dendritic spine density (108), whereas Zhu et al. reported the opposite effect of RNAi knockdown of CDKL5 (109). But both groups agree that CDKL5 promotes dendritic spine maturation in terms of spine shape. The function of Mib1 has not been examined in this developmental process. When individually expressed in cultured hippocampal neurons, Mib1 decreased spine density by 1.19 spines/10μm versus negative control (p = 0.046), whereas CDKL5 caused an increase of 1.43 spines/10μm (p = 0.049). When co-expressed, the spine-promoting effect of CDKL5 was eliminated. These neurons exhibited 0.68 fewer spines per 10μm than control neurons (p = 0.252), and far fewer (2.11/10μm) than CDKL5 alone (p = 0.0004) (**Fig. 2.7F**). Moreover, the effects of Mib1 alone and on CDKL5 are abolished by the C985S point mutation, suggesting that the role of Mib1 is dependent on its ligase activity (**Fig. 2.7G**). Moreover, CDKL5 itself caused stark changes in spine morphology, shifting strongly away from immature filopodia (p = 0.004) towards more mature thin-headed and stubby/mushroom shaped spines (p = 0.002). Mib1 caused no significant effect on spine shape when compared to control, though it did limit the effect of CDKL5 when overexpressed: Mib1/CDKL5 coexpressing neurons showed a minor shift away from filopodia and thin-headed shaped spines towards stubby/mushroom shape – the effect barely reached statistical significance for the proportion of stubby/mushroom spines when compared to control (p = 0.046). Statistical significance was not reached for the coexpressing neurons compared to neurons

expressing only Mib1 (p = 0.614, 0.129, and 0.104 respectively) or CDKL5 alone (p = 0.092, 0.297, and 0.448 respectively) (**Fig. 2.7H**). Together, our results also suggest that elevated CDKL5 increases spine width and maturity, and our data are consistent with the paper of Zhu et al. in regards to overall spine density. Importantly, we found that Mib1 interacts with and downregulates CDKL5, and thus restricts the effects of CDKL5 on dendritic development.

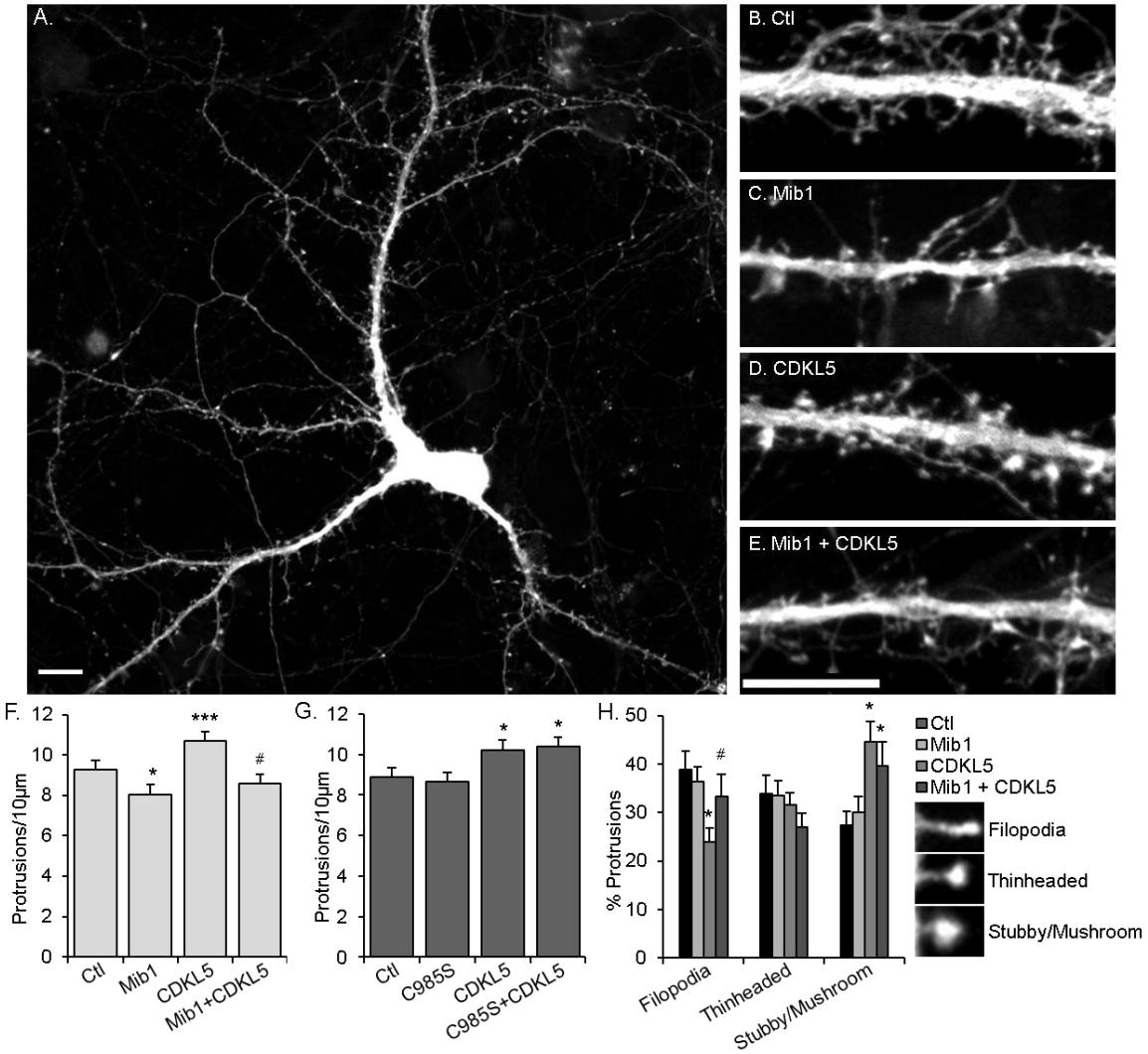


Figure 2.7. Mib1 inhibits dendritic spine outgrowth and limits pro-outgrowth effects of CDKL5 in neuronal culture.

(A) Representative rat hippocampal neuron transfected at DIV7 with EGFP and stained for EGFP at DIV14.
(B-E) Higher magnification representative images of dendritic spines in control (Ctl, GFP only), Mib1, CDKL5, and Mib1 + CDKL5 transfected neurons, showing decreased outgrowth phenotype with Mib1 and Mib1 + CDKL5, and increased with CDKL5 and C985S + CDKL5.
(F-G) Quantitative analysis of dendritic spine density from these neurons. Analysis of the spine density showing mean ± s.e.m. per 10 μm, n = 12 for Ctl, n = 16 for Mib1, n = 14 for CDKL5, n = 11 for Mib1 + CDKL5 in experiment one; n = 11 for Ctl, n = 34 for C985S, n = 27 for CDKL5, n = 12 for C985S + CDKL5 in experiment two (* $p \leq 0.05$ from Ctl, *** $p \leq 0.001$ from Ctl, # $p \leq 0.05$ from CDKL5, t-test).
(H) Classification and analysis of dendritic spine shapes in these neurons. Shape proportion analysis of the same neurons delineated between filopodia-, thinheaded-, and stubby/mushroom-shaped protrusions order of maturation. * $p \leq 0.05$ from Ctl, # $p \leq 0.05$ from CDKL5, t-test). Scale bars, 10 μm.

DISCUSSION

Affinity purification coupled with mass spectrometry is a common and robust method for mining protein-protein interaction networks. The basic technique has evolved (35) and been elaborated upon (110) since its inception well over a decade ago (111), but there remain important drawbacks. The challenge of nonspecific binding partners is significant. With technological advances in LC-MS/MS that improve sensitivity, the detected level of nonspecific background increases as well. The method we employed is based on a well-established concept that proteins can be eluted sequentially by buffers of increasing stringency. To our knowledge, the sequential elution strategy has not, however, been used to evaluate binding partners in a quantitative manner. Using this strategy, our dataset of 817 proteins was pared down to a small subset of 56 extremely highly confidence and 335 high confidence interaction partners,

whereas lower confidence potential interaction partners and possible indirect binding partners were also recovered in the remaining list. Purifying and detecting large arrays of potential binding partners and distinguishing them by binding strength rather than reproducibility during biological replicates is particularly suitable for detecting low-copy number interaction partners – an ongoing technical frontier in the proteomics field (112).

We have expanded the Mib1 signaling network in breadth and functional importance. Its critical regulation of Notch-Delta signaling alone incorporates it into almost every aspect of nervous system development (43,77,113-118). More recently, an interactome study revealed that it also plays a role in the equally vital Wnt signaling pathway (55), and several other potent developmental modulators beyond these key pathways have also been revealed to interact with Mib1, such as SMN1 (49,54) and the CDK5/p35 complex. In our previous study, Choe et al. presented evidence for Mib1 interaction with several other kinases in addition to proteins involved in membrane trafficking, the UPS, the cytoskeleton, and cell adhesion. These categories of Mib1 protein interaction were reiterated and elaborated upon by a yeast two-hybrid screen, revealing 81 putative binding partners for Mib1 and Mib2 (81). The present study reports an even larger list of potential Mib1 binding partners from adult rat brain, further expanding the Mib1 interaction network, reinforcing several pathways mentioned above, and introducing more components. As expected, the UPS, notch pathway, morphogenesis/cytoskeleton regulation, and endocytosis/vesicle transport pathways were strongly enriched in this dataset. One somewhat unexpected

result is the number of Wnt pathway members that may interact with Mib1. As mentioned previously, Mib1 had been shown to interact with one Wnt pathway member, RYK (55). Our study enlarged upon this, producing a range of powerful Wnt pathway interaction partners that includes several members of the Catenin family (alpha, beta, delta1 and delta2), several MAP/microtubule affinity-regulating kinases (Mark1, 3, and 4), and Casein kinase family members.

A potential linking factor between the role of Mib1 in both Notch signaling (its canonical role), and Wnt signaling is the endocytic pathway. It is well established that Mib1 plays a critical role in Notch signaling via DSL ligand ubiquitination and endocytosis, though this molecular mechanism is still not completely clear. What is known about the process is it is clathrin mediated, requires actin and epsin adaptors, and it may generate a mechanical pulling force on the juxtaposed Notch extracellular domain to activate signaling (119). Meanwhile, the Wnt pathway also requires clathrin mediated receptor endocytosis for successful signaling (120). Among the most highly abundant and strongest elution profiles in our dataset, Scyl2 (aka CVAK104) is a relatively unstudied clathrin coated vesicle-associated kinase that binds with both the Wnt ligand receptor Frizzled 5 and the Wnt scaffold protein Dishevelled (96). This may be a common effector for Mib1 activity in these pathways, and warrants further investigation.

The synaptic Mib1 binding partners revealed in our study included the Ca^{2+}/calmodulin-dependent kinase family (Camk2b, d, and g), the Neurexin family (Nrxn1, 2, and 3), and the Signal-induced proliferation-associated 1 like

family (Sipa1l1, 2, and 3). These span functions from dendritic spine morphology (121), plasticity at glutamatergic synapses (122), and structural connection at the synaptic cleft (123). These functions are not unlike the known Mib1 functional roles: structural changes through cytoskeletal remodeling and cell-to-cell connections, suggesting that Mib1 is a central agent to bridge these processes. Indeed, there is significant overlap between synaptic modulators and proteins that promote cellular and neuronal morphogenesis. Proteins regulating morphogenic activity were even more strongly enriched in our data than synaptic proteins, supporting the role of Mib1 in regulating neurite outgrowth (49), cell polarity (118), and synaptic plasticity (48). Moreover, our data show Mib1 suppresses dendritic spine outgrowth and acts as an opposing force to CDKL5 and its spine outgrowth and maturation effects.

Regulation of CDKL5 by Mib1 suggests Mib1 controls multiple stages of neuronal and synaptic morphogenesis. CDKL5 has been demonstrated to promote neuronal survival (124), neurite outgrowth (107), dendritic spine formation (108,109), and is an excellent example of a Mib1 interaction partner that regulates both cytoskeleton morphology and synaptic activity – bridging these two interrelated processes. Mutations in CDKL5 cause extreme variants of Rett Syndrome (125), and EIEE2 (126) which are severe and progressive forms of mental retardation. Our experiments demonstrate that when Mib1 is in abundance in relative to CDKL5, dendritic spine outgrowth is impaired. This may be an important consequence of the mutations seen in CDKL5-associated diseases. Removal of CDKL5 via deletions or mutations resulting in

nonfunctional protein may tip the balance away from signaling pathways that lead to cytoskeletal outgrowth and neuronal maturation, resigning developing neurons to an outcome with shorter processes and fewer synapses.

In summary, using our novel and widely applicable improvement to affinity purification proteomics, we uncovered the most comprehensive view of the Mib1 interactome to date and partitioned it by binding strength and confidence of interaction. Our experimental design utilized several recent advances in MS-based proteomics – TMT isobaric tags and high resolution/high sensitivity mass spectrometry to attain this deep and distinguishing analysis. Our data reinforce the pivotal role of Mib1 in neuronal development and identify a large number of novel interaction partners. These interaction partners spread across such varied and critical signaling networks suggesting Mib1 may act as a hub to link these networks and perform a widely powerful governing role for cell, tissue, and organ development. Further work on this critical protein is certainly warranted as interactions with many well- and lesser-known molecules are highly likely to have functional ramifications.

CHAPTER THREE

Biotin and Ubiquitin Labeling of Ligase Substrates (BULLS) Proteomics Analysis of Mib1 Substrates

The comprehensive and high-confidence detection of E3 ligase substrates is a prominent ongoing challenge in the molecular biology and biochemical fields. With over 600 human E3 ligases currently known (45), combined with the pervasiveness and significance of the ubiquitin proteasome system (UPS) in virtually every biological process (127,128), the drive for a high-throughput technique to discover the ubiquitination targets of a given E3 ligase is powerful. Significant challenges are posed by the fact that these interactions are generally transient, substrates are often quickly degraded, the system as a whole is under highly efficient control via deubiquitinases and the proteasome, the ubiquitome is simply very large and complex, the topology of polyubiquitin chains can be very complicated, and even the notably high rate of endogenous ubiquitin expression can pose problems. Nonetheless, the UPS also lends itself to examination – in

comparison to other tag-based systems such as phosphorylation – as ubiquitin is a gene-encoded protein tag and therefore allows for direct genetic manipulation and creation of modified gene constructs. To date, the most successful systems-based methodology compares ubiquitination sites throughout the proteome using an antibody for the signature ‘di-Glycine’ peptide of ubiquitinated proteins combined with knockout of the E3 ligase of interest (46). This, however, is an indirect measure as it does not provide evidence of direct enzyme-substrate interaction.

Recently, an enzyme-substrate tagging methodology was developed for detection of direct protein interactions (129) utilizing the highly efficient and highly specific biotinylation of a small acceptor peptide tag (AP) by the E. coli derived biotin ligase BirA (130). The reaction requires a proximity of 50Å for ligation to occur and BirA has no mammalian substrates while the AP is not recognized by any mammalian ligases. This leads to labeling of the AP tag with high fidelity and abundance, based on very close proximity between molecules. Subsequently, the covalently bound biotin can be exploited for downstream analysis by numerous potential avenues of detection and purification through the extremely strong streptavidin-biotin bond. This method has been implemented with some success to examine binary interactions between two proteins in live cells by coding the BirA enzyme in frame with one protein of interest and coding the AP in frame with the second protein (130).

We sought to extend this methodology into a more systems-based and ubiquitin-focused technique dubbed BULLS (biotin and ubiquitin labeling of ligase

substrates). In this approach, we fused the BirA enzyme to Mib1, our E3 ubiquitin ligase of interest, and the AP to ubiquitin. This results in the biotinylation of AP-tagged ubiquitin molecules utilized by Mib1, with biotin carried along onto ubiquitination substrates of Mib1. Downstream detection and identification of these substrates can then be achieved by streptavidin antibodies for immunocytochemical imaging or western blots, and for affinity-based pulldowns, which can be analyzed by mass spectrometry proteomics.

MATERIALS AND METHODS

Plasmids and Antibodies

Prior to creation of fusion proteins, we optimized BirA and BirA Acceptor Peptide (AP) codons for *Homo sapiens* expression. This was particularly important for two reasons: first, these are both naturally occurring E. coli proteins with no mammalian homologs and therefore not ideal for eukaryotic expression, second, endogenous ubiquitin is a very highly abundant protein, and in order to contribute significantly to the cellular pool of ubiquitin, expression must occur in a highly efficient manner.

The 15 aa AP and 6x His sequences were cloned onto the N-terminal of a ubiquitin gene in a *pFUGw* backbone under a ubiquitinC promoter. We also created a WHE13-15GEF variation of the AP sequence via site-directed mutagenesis of *TGGCAYGAR* sequence to *GGNGARTTY*.

Using the InFusion cloning kit (Clontech), BirA was cloned into a position at the carboxyl terminal of an HA tag and the amino terminal of Mib1 in a

pcDNA3 plasmid backbone, with a short 4 amino acid chain as linker between BirA and Mib1. The ligase domain of Mib1 is very close to the carboxyl terminal of the protein – residue 985 being critical for ligase activity (49) – and attachment of the 35kD BirA at the N-terminal will therefore be less likely to disrupt Mib1 activity than attachment at the C-terminal. Three negative control fusion proteins were created – BirA-Mib1(C985S), BirA-Mib1(803X), and BirA-Mib1(401X). Each of these were created using a site directed mutagenesis kit (Clontech), with C985S receiving the modification to switch the Cysteine codon to Serine, and the nonsense mutations receiving double stop codons at the noted location.

The antibodies used included HA mouse monoclonal antibody (Abcam), and Streptavidin-HRP (Life Technologies).

Cell Culture & Transfection

HEK 293 cells were grown in DMEM + serum supplemented growth media and transfected using FuGene HD (Promega). Total plasmid DNA amount of 1 ug for each well of a 12 well plate, a ratio of 3 uL FuGene to 1 ug DNA ratio, and a post-transfection incubation period of 72 hrs was derived from the Promega FuGene protocol. BirA-Mib1 and BirA-LD plasmid levels were varied to optimize conditions, with EGFP plasmid used to balance total DNA levels. Various post-transfection incubation periods were tested in attempts to optimize conditions, with 72 h, 48 h, 24 h, and 12 h each utilized. Cells in these experiments were transfected one day after replating at 33% confluency, harvested at the selected time-points, then protein concentrations tested via BCA to ensure equal loading in subsequent SDS-PAGE analysis.

Protein Preparation & Western Blot

Cell lysis was performed using ice cold 2% SDS buffer in 50 mM HEPES, pH 8, with cOmplete protease inhibitor cocktail (Roche Applied Science) and 10 mM iodoacetic acid (IAA). IAA was utilized to inhibit deubiquitinases activity as integrity of protein ubiquitination was key to this experiment. After aspiration of growth media, lysis buffer was added directly to cells, mixed, pipetted, transferred to microfuge tube, sonicated at 4°C, and either analyzed directly or frozen at -80°C.

RESULTS

Biotinylation and Ubiquitination by BirA-Mib1, BirA-C985S, and Free BirA

In an attempt to generate a system for detection of E3 ubiquitin conjugation substrates with high confidence, specificity, and flexibility, we created a fusion protein of Mib1 E3 ligase with bacterial BirA biotin ligase and a fusion protein of Ubiquitin with the specific BirA biotinylation target sequence. A previously established ligase dead form of Mib1 was used for generation of a negative control BirA fusion protein, and free BirA was used as a positive control.

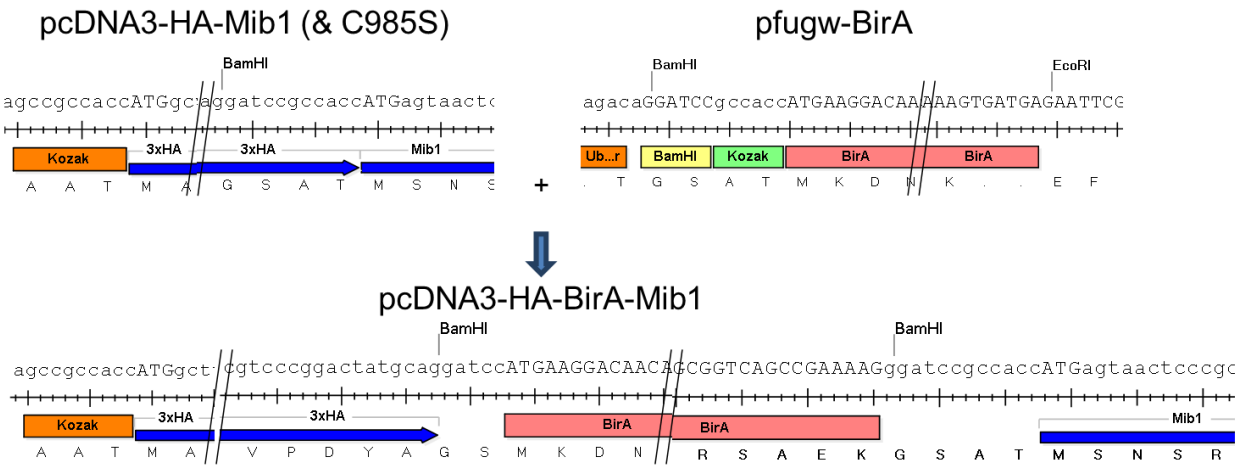


Figure 3.1. Cloning of the BirA sequence into the HA-Mib1 construct results in HA tagged BirA-Mib1.

A BamHI site near the end of the HA sequence allowed us to clone BirA in frame between the HA tag and Mib1 protein using the InFusion cloning kit from Clontech.

To verify that BirA and Mib1 maintained their enzymatic activity in the fusion protein, we expressed BirA-Mib1 and AP-Ub in HEK 293 cells followed by western blot analysis. We used biotin-detecting streptavidin-horse radish peroxidase in a one-step probe method that allowed high affinity detection of biotin and biotin-labeled species. Lane 7 (Fig 3.1) depicts the very low levels of HEK 293 endogenous biotin containing protein species as detected by this method. As predicted, the biotin labeled species in the free BirA positive control (Fig 3.1 lanes 8 & 9) exhibits a highly intense smear of biotinylated and ubiquitinated protein species – quite similar in pattern to a probe for ubiquitin itself. Also as predicted, BirA-Mib1 with AP-Ub exhibits a slightly less intense smear, with many ubiquitinated and biotinylated species (Fig 3.1 lanes 1-3). This is indeed predicted as the Mib1 E3 ligase is known to be promiscuous and to bind many potential substrates (49,131). Unfortunately, this pattern was virtually unchanged in the E3 ligase-dead BirA-C985S negative control (Fig 3.1 lanes 4-

6), suggesting that there may be off-target background signal and the system may need stronger control.

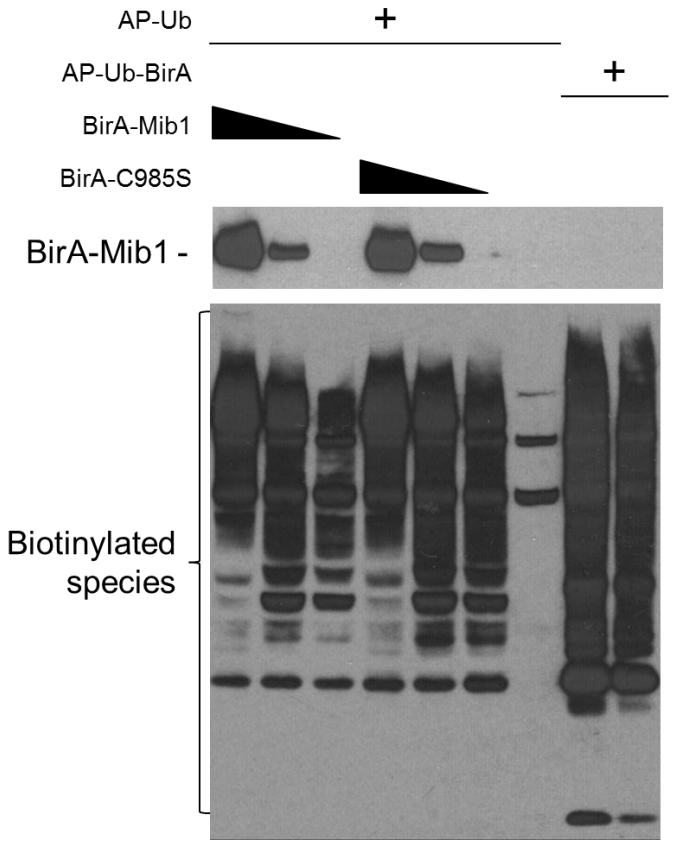


Figure 3.2. BirA-C985S ligase dead negative control produced similar levels of positive signal to active form.

While proving the usefulness of the concept overall, our BirA-C985S negative control produced almost identical signal to the test construct. Free BirA produced much stronger and widespread biotinylation, with abundant mono-ubiquitin bands.

Biotinylation and Ubiquitination by Mib1 Truncation Mutants

To establish more rigorous negative controls, we created two BirA-Mib1 truncation mutants (Fig 3.2). BirA-803X contained BirA fused to Mib1 with the entire C-terminal catalytic domain removed, and BirA-401X was an even more

severe version with both the C-terminal catalytic domain as well as the structural middle domain removed.

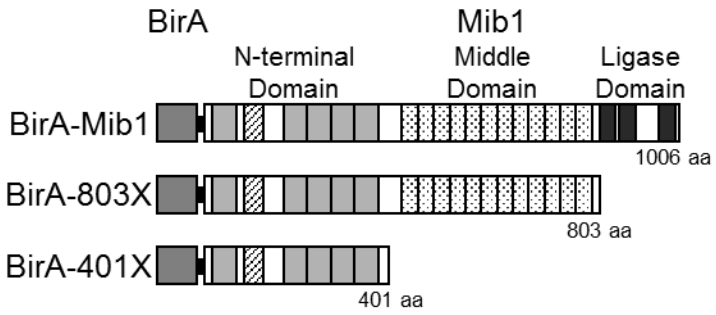


Figure 3.3. BULLS gene constructs included a full length test protein and two truncations used for negative controls.

The full length test construct contained BirA cloned in frame onto the n-terminus of full length human Mib1 protein. The two negative control truncate constructs included BirA cloned onto the n-terminal of a form of Mib1 with no c-terminal ligase domain and one with no middle domain or ligase domain.

We examined these new controls in comparison to the experimental BirA-Mib1 construct in HEK 293 expression via Western blot (Fig 3.3). These negative controls exhibited slightly less intense ubiquitination smears than the test construct, but unfortunately the decrease was not suitable to move forward with purification and MS examination. Contrary to our predictions, and perhaps revealing as to the excessive background seen in these experiments, the truncation mutants also produced what we believe to be a biotinylated but unconjugated ubiquitin species seen as a ~10kD band (Fig 3.3). Of particular interest, and potentially encouraging, BirA-Mib1 does not produce this species. This suggests AP-Ub biotinylated by this construct is subsequently conjugated to substrates, while the ligase dead mutants are unable to complete the second part of this reaction.

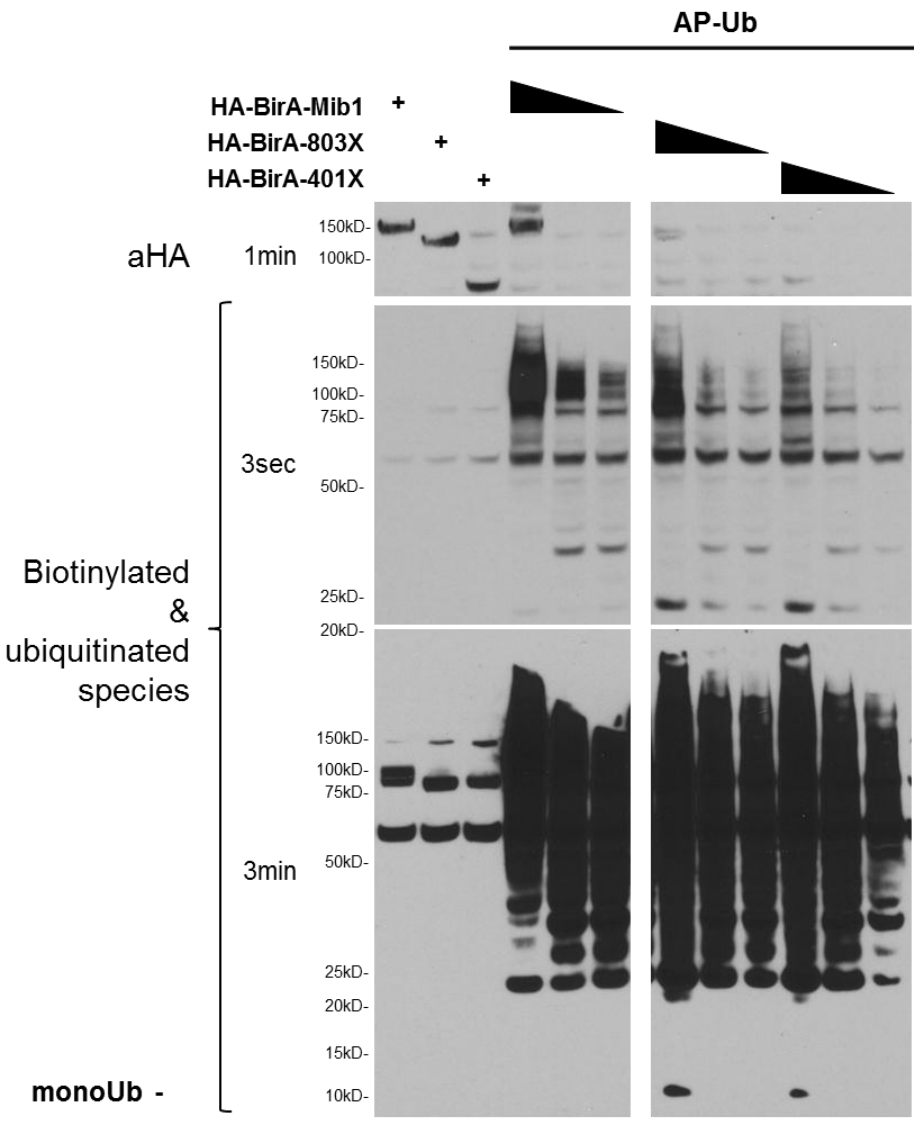


Figure 3.4. Western blot examination of truncation mutants shows decreased but still substantial positive signal.

We still observed substantial background signal from our more extreme negative controls. Interestingly, the truncates also produce a pool of biotinylated monoUb (11kD band), which could contribute to the nonspecific biotin-ubiquitination in these samples.

Modification of the BirA Acceptor Peptide Sequence

To further increase the signal-to-noise ratio in our system, we followed the methods of the Ting group who, struggling with similar difficulties, took a somewhat broad mutation screening approach on the acceptor peptide sequence

(130). They found that by modifying residues 12-15 of the protein sequence from WHE to GEF, they could obtain a ~5-fold boost in signal-to-noise. We therefore performed the same modification via site-directed mutagenesis in our AP-Ub construct and tested via Western blot from HEK 293 lysate. We saw a significant decrease in signal from the negative control, as expected, but also saw decreased signal from the experimental condition. This largely negated any gains in signal-to-noise we would have otherwise obtained.

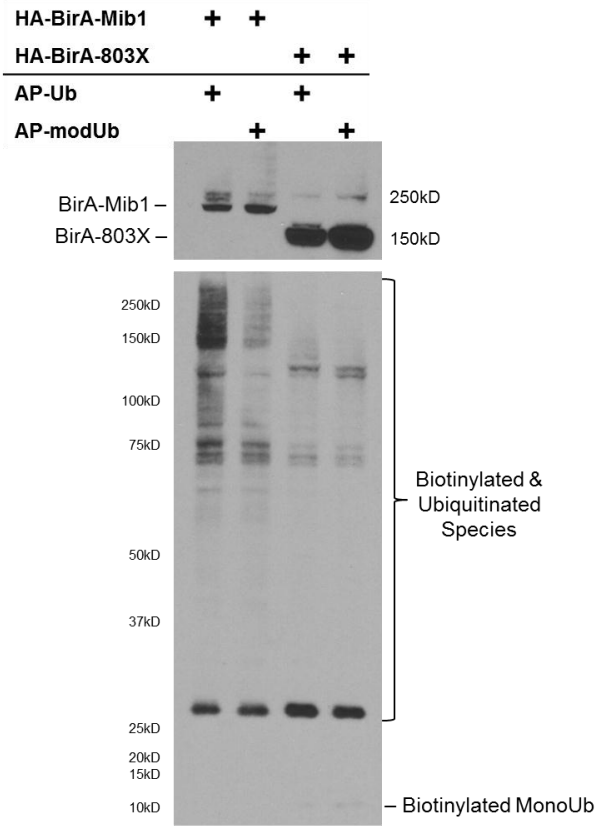


Figure 3.5. Modified of the AP sequence decreases signal in both test and negative control samples.

Modified AP sequence may have slightly reduced background, but also the positive signal in the test sample, which lead to similar overall signal-to-noise ratio

To address the possibility that biotinylated ubiquitin was being recycled in the cellular system and to limit the opportunity for off target reactions, we attempted to optimize the reaction by examining a range of post-transfection incubation periods shorter than the Promega recommended 72 h. Optimization was attained by transfecting each group of cells with the desired plasmid constructs at the same time, harvesting protein after the specified time period, and equalizing protein loading for western blot analysis by preliminary BCA assay. For example, because all cells were transfected below 100% confluence, and the 12 h group would thus replicate fewer times and produce less protein than the 72 h group, we simply had to test a larger percentage of the protein from 12 h than 72 h to match overall protein levels.

What resulted were correspondingly less intense biotin/ubiquitin smear patterns for each decrease in incubation period (Fig 3.5). Unfortunately, the intensity decreased for both the test group and the negative controls, rather than any disproportionate large decreases in the negative controls as we had hoped. Signal-to-noise thus remained relatively unchanged and unsuitable for further pull-down experiments.

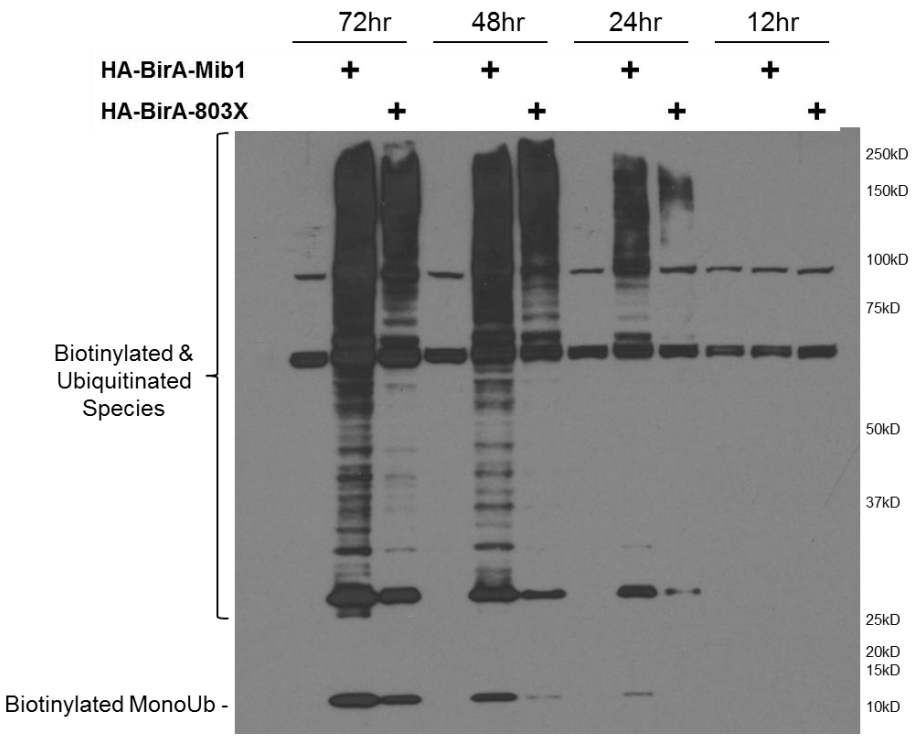


Figure 3.6. Examination of multiple incubation periods shows increases in signal from both test and negative control.

As expected, decreased incubation times led to decreased signal, with somewhat promising signal-to-noise ratio at 24 hrs.

DISCUSSION

The search for comprehensive detection of direct E3 ligase substrates steered us toward the expansion of a previously established protein-protein interaction methodology. The potential of this system was exciting, as it utilized a small molecular tag, a small and highly efficient enzyme, required a very tight molecular proximity, was amenable to use in live cells, and provided for extremely broad downstream analyses. It did indeed produce strong and specific biotinylation, with test and positive control conditions exhibiting ubiquitin-like

smears and with negative controls (either no AP or no BirA) producing negligible numbers of biotinylated species. Further, the banding produced by BirA-Mib1 were distinct in pattern and abundance from those produced by free BirA positive controls, suggesting that the BirA-Mib1 biotinylation was a subset of the overall ubiquitinated proteome.

Despite these encouraging observations, we also encountered a continuing issue of excessive signal from our would-be negative controls. As with any systems-based approach, we expected off-target activity and false positive signal, but designed the various null mutant BirA-Mib1 fusion proteins to establish a measure of these factors in the experiment. As we underwent extensive optimization and manipulation of these controls, we were unable to limit the background signal to the point at which we would be confident in designating true substrates from our BirA-Mib1 construct. The prevalence of signal we observed from fusion proteins that presumably possessed no E3 ligase activity, in addition to rather abundant biotinylated monoubiquitin, led us to believe that unpredicted, detrimental, and largely unavoidable processes were taking place in this system that would prevent further pursuit of these methods.

We attributed these issues primarily to two potential processes that were unforeseen at the time of experimental design (Fig 3.5). First, the BirA-Mib1 constructs – both ligase active and ligase-dead controls – might have been binding and biotinylating AP-ubiquitin in the absence of any ubiquitination substrate, then subsequently releasing this biotinylated ubiquitin and allowing any other E3 ligase present in these cells to conjugate it to its substrates (Fig

3.5B). This likely happened in parallel to the conjugation of ubiquitin by Mib1, but clearly would lead to detection and identification of off-target, false-positive substrates. Secondly, the biotinylation and ubiquitination of Mib1 substrates may have taken place with the fidelity we desired – and required – but the biotinylated ubiquitin was subsequently removed by deubiquitinases and/or by proteasome activity (Fig 3.5C).

Deubiquitinases, analogous to phosphatases in the phosphorylation system, cleave and release ubiquitin monomers, serving both to antagonize the activity of ubiquitin ligases, and to provide the very ubiquitin molecules on which ubiquitin ligases act. There are over 100 human deubiquitinase genes. Some cleave ubiquitin monomers from the pro-protein polyubiquitin chains in which ubiquitin is initially expressed, and others have even been shown to closely associate with E3 ligases to negatively regulate them (132). Indeed, Mib1 has been shown to bind strongly with Usp9x (131), a deubiquitinase which removes ubiquitin from SMN1 (133), a known Mib1 ubiquitination substrate (54). DUBs and DUB domains are also present in proteasomes, the protein destruction machines of the cell, and remove ubiquitin from proteins prior to degradation (134). Following removal via either of these pathways, ubiquitin molecules are simply recycled by the cell and attached to new substrates. In our system, it is highly likely that biotinylated AP-Ub is being recycled and reattached alongside endogenous Ub, and therefore we observe a preponderance of off-target false positive signal.

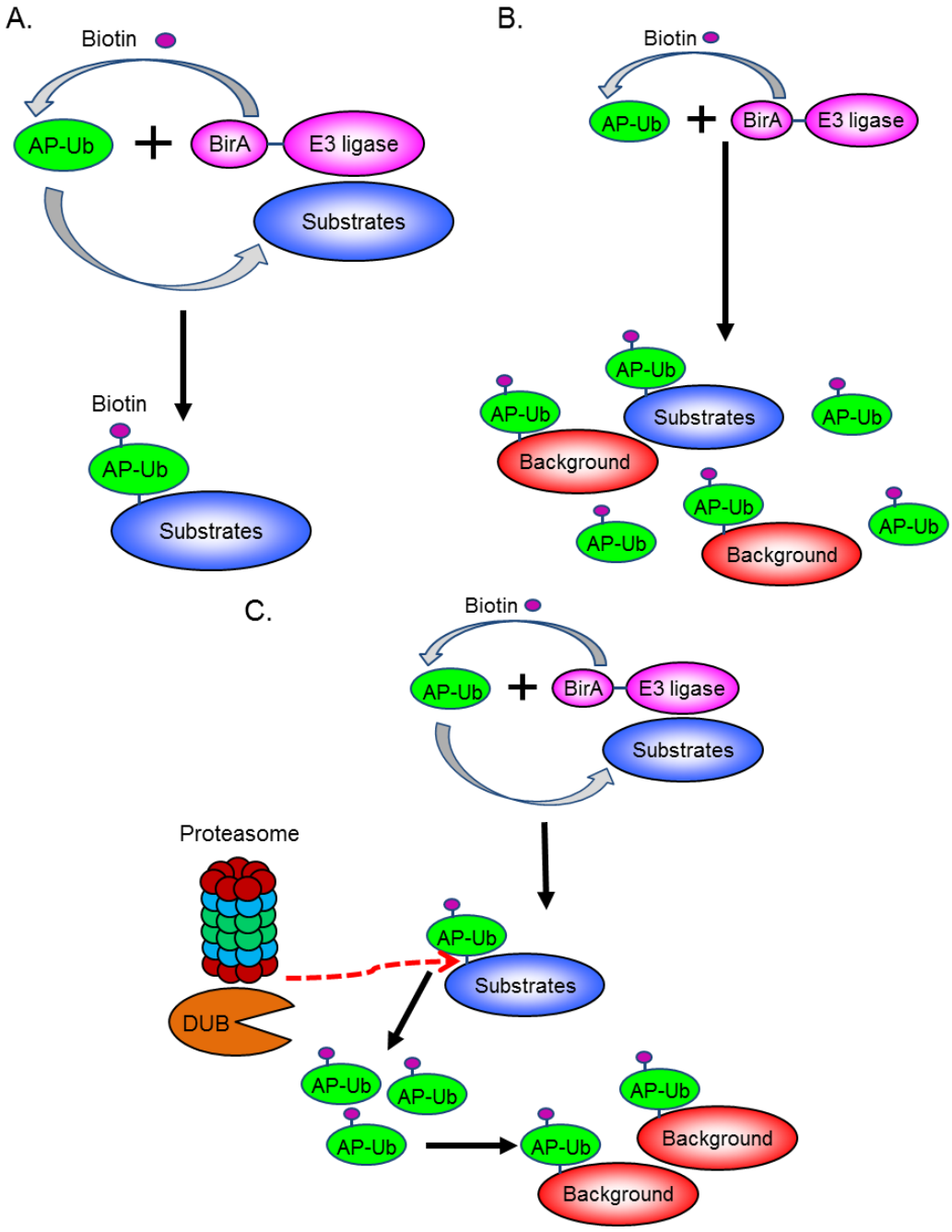


Figure 3.7. Proposed models of technical pitfalls.

(A) Desired model for successful biotinlyation and ubiquitination

(B) The BirA-Mib1 and BirA-Mib1LD enzymes may be biotinylating AP-Ub in the absence of a substrate – or ligation to a substrate, then releasing these species into the general cellular pool.

C) Biotinylated ubiquitin might be released into the cellular pool after proteolysis by the proteasome and/or by deubiquitinases detaching the tagged ubiquitin. In both cases, the tagged ubiquitin is re-entered into the cycle for ubiquitination by any E3 ligase

The efficacy of the BirA-AP method has encountered complications in other groups as well. As mentioned, the Ting group, who originally developed the system for cell-surface interactions before 2005 (135), advanced the approach several years later for intracellular interactions (130). In that publication, they also presented experiments to improve signal-to-noise (130), and a paper using this system for live inter-neuron imaging from 2010 (136) was later retracted. Since the retraction, this group has seemingly abandoned the approach. Together, these developments suggest the BirA-AP system itself has significant pitfalls, and these applications did not involve the UPS, which we believe was the source of our most insurmountable difficulties. A modification of the technique, dubbed BioID by its creators, utilizes a mutant BirA that possesses biotin activation activity but lacks ligase activity (137). This modified enzyme, also used as a fusion product with a protein of interest, subsequently releases activated biotin (biotinoyl-5'-AMP), which readily attaches to primary amines in the vicinity – which include lysines, arginines, and amino terminals of nearby proteins. This method was applied to an E3 ligase with results published in the last year showing improvement over more traditional Flag tag affinity-purification MS proteomics (138). Our efforts, while promising, were ultimately unable to solve the continuing challenge of comprehensive capture and identification of E3 ligase substrates. However, the knowledge, technology, and techniques seem to be within reach, and this advance may not be far off.

CHAPTER FOUR

Differential Enrichment and Elution Proteomics (DEEP) Analysis of the APP Interactome

Adapted from: *Molecular & Cellar Proteomics*

Bing Bai, Joseph Mertz, Ping-Chung Chen, Yuxin Li, Ji-Hoon Cho, Timothy Shaw, Hong Wang, Gang Yu, Thomas G. Beach, Xusheng Wang, Junmin Peng

Forthcoming

In 1907, Alois Alzheimer described “A peculiar disease of the Cerebral Cortex” presenting severe cognitive impairments, emotional disturbances, and miliary foci throughout the cerebral cortex (139). Today Alzheimer’s disease (AD) is the 6th leading cause of death (56), cause of over 70% of dementia cases in the United States (57), and costs the US economy over $400 billion and rising per year (12, 13). At preclinical stages AD pathophysiological processes are ongoing but symptoms evident (60), predementia symptoms emerge

subsequently in the mild cognitive impairment (MCI) stage (61), and finally, AD dementia involves declines from previous cognitive levels, significant functional impairment, deficits in learning and recall of recently learned information, language and visuospatial deficits, and executive dysfunction (62). AD pathology, beginning in preclinical stages, is typified by intracellular neurofibrillary tangles (NFT) and extracellular amyloid plaques (65). Senile plaques, and neuritic plaques specifically, are associated with dystrophic neurites, synapse loss, and gliosis (66). Aβ, a cleavage product of the amyloid precursor protein (APP), was found to be the major species in amyloid plaques (140,141), and causal mutations in APP, PSEN1, and PSEN2 – two proteases known to process APP – were identified in early-onset forms. These findings resulted in the Amyloid cascade hypothesis (142) and the amyloid pathway becoming the most widely studied biological pathway thereafter (143).

Aβ is a 40 or 42 amino acid peptide proteolytically cleaved from the extracellular and transmembrane domains of APP, which is a one pass transmembrane protein with a large N-terminal extracellular domain and smaller C-terminal cytoplasmic tail. The APP protein shares 98.8% identity with its mouse ortholog and is highly similar to its mammalian paralogs APLP-1 and APLP-2 (144). All three members of the family share E1 and E2 domains in the extracellular region and are highly similar in the cytoplasmic NPXY motif, but APP is unique from its family members in the region that becomes Aβ.

APP – and in many cases Aβ –interacts with a vast array of interaction partners. The most prominent are the secretase complexes that process full

length APP into Aβ. The γ–secretase complex, including presenilin-1/2, cleaves APP in an intramembrane position, which helps contribute to Aβ formation, and is therefore amyloidogenic. The complementary amyloidogenic enzyme β–secretase, or BACE1, is a membrane-spanning enzyme that cleaves APP on the extracellular side of the membrane, completing Aβ formation. The anti- or non-amyloidogenic α–secretase family cleaves APP to release its extracellular portion, and helps prevent Aβ formation. ADAM10 protease, is a member of this family and has been shown to compete with β-secretase and suppress Aβ production (69). Several caspases, predominantly CASP3, 8, and 9 also interact with APP (145,146).

APP undergoes a separate cleavage event to produce a species known as the APP intracellular domain (AICD), which has been shown to participate in clathrin mediated endocytosis (147), axonal trafficking, and transcriptional regulation. The AICD NPXY motif interacts with a large array of proteins, such as the phosphotyrosine-binding domain (PTB) family (148) which includes amyloid precursor binding A and B families (APBA and APBB) (70), the c-Jun NH2-terminal kinase-interacting proteins (JIPs), and Src-homology-2 domain-containing (Shc) protein families (71). JIP1 facilitates MAPK8 phosphorylation of AICD (149), binds to phosphorylated APP and mediates its axonal transport by kinesin (150). Kat5 colocalizes and complexes with APP intracellular domain through APBB1 to induce transcription of itself, APP, and BACE (151,152). APPBP2 binds AICD post caspase-cleavage and leads to increased APP levels on cell surface, and cell death in primary neurons and HEK cells (153). GSK3β

binds AICD and promotes its kinase activity, leading to reduced Wnt signaling and increased neurite outgrowth (72). Meanwhile the canonically Parkinson's disease associated Alpha-synuclein colocalizes and co-aggregates with Aβ *in vivo*, and may form pore-like oligomers that decrease cell membrane integrity (154,155).

Despite the combined efforts of many individual APP protein interaction analyses, there remains a large and important gap in our current understanding of actual AD pathogenesis, largely due to the lack of thorough, well-controlled large-scale interactome studies. Previous large-scale studies focused on processed APP, in the form of soluble Aβ oligomers (156) or did little to address differential binding affinity to APP (157). A thorough and relevant analysis of the APP interactome must entail the capture and identification of the broad range of strong, weak, and transient APP candidate interactors from human AD tissue and systematic determination of the confidence that these candidates represent true biological interactions.

We thus undertook a systems-based analysis of APP interaction partners by co-immunoprecipitation mass spectrometry proteomics in human AD brain tissue. Traditionally, affinity and antibody-based purification methods utilize strong washes to limit background, but this is not necessarily an accurate reflection of the biology – truly functional interactions can occur between weakly binding proteins. Recovering weakly bound interaction partners while still differentiating them from more strongly bound partners greatly expands the potential for exhaustive examination of protein function (131). The widespread

lack of agreement between different antibodies for the same target protein is another considerable issue in co-IP/MS experiments, leading to variability and lack of reproducibility in data (158,159).

We sought to address both of these issues via thorough long gradient LC and high resolution quantitative MS analysis of multiple elution conditions from co-IPs using multiple test antibodies. Our methodology utilized differential enrichment and differential elution proteomics, and we coined it DEEP. This exhaustive approach allows us to broaden our dataset across binding affinities as well as antibodies in order to reduce false negatives, while also allowing for comparison across these parameters to examine agreement and limit false positives. Further, this was an examination of AD brain tissue, and thus provides key insight into pathological interactions taking place in this crippling disease.

Overall, we identified 7804 proteins from 8 unique co-IP fractions and 2 replicates of the AD brain lysate input. As expected, APP was among the most highly enriched overall. Of the 8 co-IP fractions, 5 were from different elution conditions of 3 unique APP antibodies, and 3 were from different elution conditions of control IgG antibodies. Our multi-directional comparisons of these test and control antibodies provided us with differential enrichment and differential elution datasets within which we examined agreement and overlap. Of the 7804 proteins, we found 696 proteins to be differentially enriched to various degrees and roughly 40 reached extremely high confidence of interaction. Within these 40, pathways such as Aβ peptide metabolism, neural development,

synaptic regulation, trophic signaling, intracellular trafficking, mitochondrial function, and RNA splicing were strongly enriched.

MATERIALS AND METHODS

AD Brain Tissue Lysate Preparation

Frozen human postmortem brain tissue (superior frontal cortex) with clinically and pathologically confirmed Alzheimer's disease were provided by Banner Sun Health Research Institute. The postmortem intervals of these samples were between 2~4 h. Brain tissues were homogenized in low salt buffer [LS, 20 mM Hepes buffer, pH7.4 with 1% CHAPS and 1x protease inhibitor cocktail (Roche)] at 10 ml buffer per gram tissue. After the protein concentration was measured by BCA, all samples were run by SDS-PAGE to evaluate their protein integrity. 10 cases with similar sample quality were finally chosen and equally mixed. The pooled sample in LS buffer was centrifuged at 20,000 g for 1 min to collect the supernatant (Sup1), then the pellet was subjected to the same volume of high salt buffer (HS, 20mM Hepes, pH 7.4, 300mM NaCl, 1x protease inhibitor cocktail). After brief microsonication, the sample was centrifuged again to collect the supernatant (Sup2). Sup1 and Sup2 were mixed as input (IP) (~5 μg/μl) for co-immunoprecipitation.

Co-Immunoprecipitation (Co-IP)

Protein G Dynabeads (Life Technologies) were used for Co-IP. 100 μl slurry (~5ul beads, binding capacity: ~30 μg) was incubated with excess antibody or control IgG for 15 min. After brief wash with the Co-IP buffer (mixture of LS

and HS buffers), the saturated beads were added to 100 µl input (~0.5 mg protein) and incubated for 2 hrs at room temperature. After incubation, the vials were placed on a magnetic rack for beads to be attracted on the side wall in each vial, and the flow-through was removed. For the unwashed eluate, 100 µl PBS was added for gentle rinse without disturbing the beads to reduce the residual input. For the washed eluate, the vials were taken off the magnetic rack and the beads were washed by 100 µl TBST buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.05% Tween-20) three times as per common Co-IP procedures. Both unwashed and washed beads were eluted using 1x SDS-PAGE sample loading buffer with 10 mM DTT for further processing. Of the three anti-APP antibodies used for Co-IP, 6E10 and 4G8 are mouse monoclonal from BioLegend, Y188 is rabbit monoclonal from Abcam. Purified IgG from the serum of mouse or rabbit were used as controls (Sigma).

10-plex TMT-based Quantitative LC-MS/MS Analysis

10 DEEP samples that included two input replicates, five unwashed Co-IP eluates and three washed eluates in 1x SDS-PAGE sample loading buffer (10 mM DTT) were heated at 80°C for 10 min, run for a short polyacrylamide gel, then each gel piece was cut and digested overnight by ~10 ng/µl trypsin (enzyme: protein = ~1:20) in 50 mM TEAB buffer, pH8.5, 10% acetonitrile. Digested peptides were extracted, dried, and resuspended in 100 µl 50 mM HEPES buffer, pH 8.5 with 25% acetonitrile. TMT 10-plex reagents (ThermoFisher Scientific) were added to each vial according to the commercial instructions and incubated for 1 h at room temperature and equally mixed. After

drying, the labeled peptide mixture was desalted on Sep-Pak C18 column (Waters). The cleaned peptides were then fractionated on a XBridge C18 column (Waters, 3.5 μm resin, 4.6 mm x 250 mm), by a gradient of solvent B (90% acetonitrile in 10 mM ammonium formate, pH8.0): 5% for the first 10 min, then steadily increased to 15%, 20%, 35%, 50% and 95% at 11, 20, 95, 115 and 120 min, respectively. Solvent A (10 mM ammonium formate, pH 8.0) was used to generate the solvent B gradient. In total, 60 fractions were collected and dried.

The peptides in 5% formic acid were separated by a C18 column (75 μm × 30 cm, 1.9 μm resins) at 0.3 μl/min flow rate. The main gradient was 20%~45% of buffer B (65% acetonitrile, 0.2% formic acid, and 5% DMSO) for nearly 2 h generated by buffer A (0.2% formic acid and 5% DMSO). The gradient was slightly adjusted between samples according to the appearance of the LC chromatography curve. Eluted peptides were scanned on the Q Exactive HF MS analyzer (Thermo Fisher Scientific) by MS1 scan (120,000 resolution, 1×10^6 automatic gain control, and 50 ms maximal ion time), followed by MS2 scans top 20 (HCD collision with 35 normalized collision energy, 60,000 resolution, 1×10^5 automatic gain control, 128 ms maximal ion time, 1.0 *m/z* isolation window with 0.3 *m/z* isolation offset, and 15 sec dynamic exclusion).

52 MS raw files were acquired and converted into mzXML format for peptide matching by the JUMP algorithms (91). A combination of target database from Uniprot and a decoy database generated from reversed peptide sequences of the target database was used. The mass (+229.162932) of TMT tags was

added to peptide N-termini and Lysines as static modifications. Only a, b, and y ions were chosen for peptide spectrum matches (160).

Protein Quantification by TMT Labeled Peptides

Quantification of TMT labeled peptides was carried out by an in-house program in the following steps. (i) TMT reporter ion intensities of each identified PSM were extracted and recorded. (ii) The raw intensities were corrected according to isotopic distribution of each labeling reagent. For instance, the TMT126 reagent produced 91.8% 126 *m/z* reporter ion, 7.9% 127 *m/z* reporter ion, and 0.3% of 128 *m/z* reporter ion. (iii) The average of all ten reporter ion signals was used as a reference to compute a relative intensity between each sample and the average. (iv) The relative intensities of PSMs were averaged for identified proteins. (v) The combined reporter ion intensities for the Input replicates, washed and unwashed mouse IgG, washed and unwashed APP Antibodies 1 and 2, and unwashed rabbit IgG and APP Antibody 3 were compared to derive the $\log_2$ ratios. (vi) Visualizing these comparisons via histogram analysis, the data were fit to Gaussian distributions, and thereby mean and standard deviation estimated. (vii) To analyze experimental variation, the intra-sample comparisons between the input replicates were examined as null experiments, showing an average standard deviation of 0.10. (viii) Finally, Z-scores were calculated using the fitted Gaussian parameters, and we selected a Z-score cutoff of 1.8 for comparing different samples. For enrichment, the eluted proteins of the APP antibodies (both washed mouse APP antibodies, both unwashed mouse APP antibodies, and the unwashed rabbit APP antibody

fractions) were compared to the average of the two Input fractions. For differential enrichment, the APP antibody fractions were compared to their respective control IgG (washed mouse APP Antibodies to washed mouse IgG, unwashed mouse APP Abs to unwashed mouse IgG, and the unwashed rabbit APP ab to the unwashed rabbit IgG).

Interaction Network Analysis

Enrichment of Kyoto Encyclopedia of Genes and Genomes pathways (KEGG) (92) was determined by analyzing the dataset using DAVID Bioinformatics Resources 6.7 Functional Annotation Clustering tool (93,102), and cell compartment and gene Ontology using PantherDB Online Classification System (161,162). STRING-DB (94) was used to evaluate interconnectivity between members of the pathways and processes determined to be enriched by DAVID and PantherDB.

RESULTS

LC-MS/MS Analysis of APP Co-Immunoprecipitation

7804 unique proteins were identified and quantified via automated database search and manual verification with a false discovery rate of 1%. The relative abundance of each protein in each fraction was determined by quantification of the reporter ion derived from their corresponding TMT isobaric tag. APP was highly prevalent in our purifications: a disproportionately high number of MS scans identified APP peptides overall, and the identified peptides originated from positions throughout the APP protein, as opposed to only the

area surrounding the antibody antigens or the Aβ region. This suggests that we recovered full length APP, and therefore were able to examine the endogenous interactome of APP, in addition to pathological interactions seen with Aβ after AD onset.

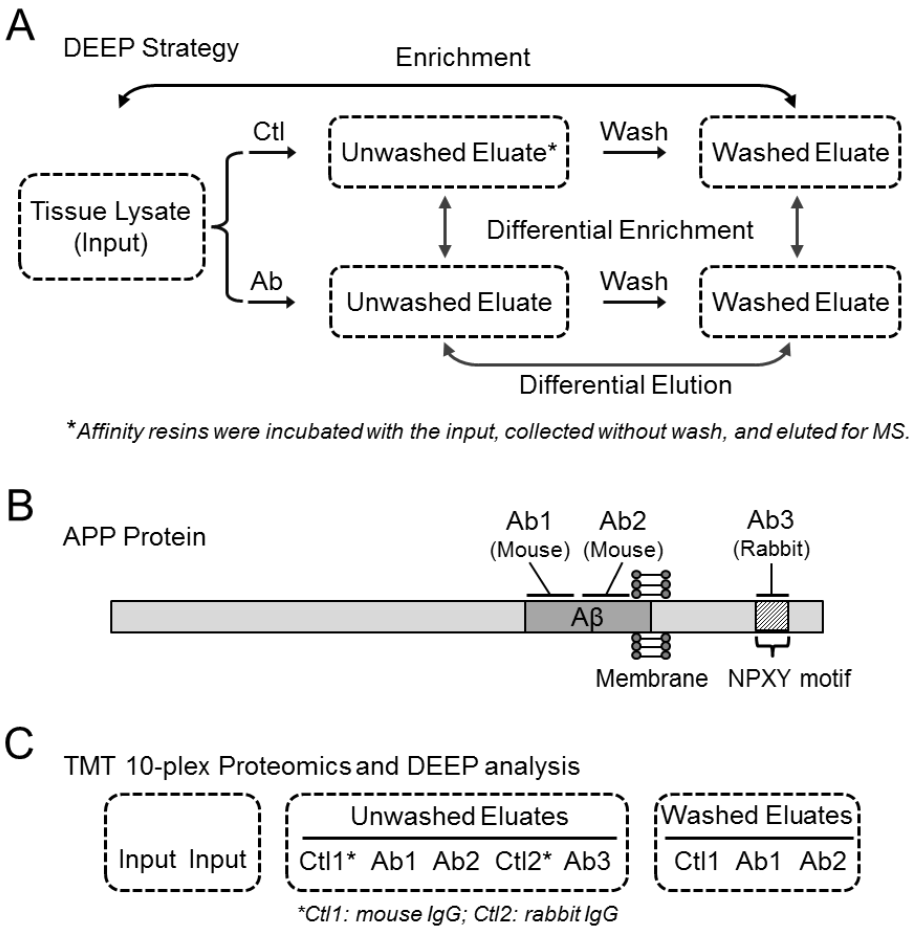


Figure 4.1. DEEP analysis utilizes multiple comparisons of APP IP-MS proteomics from human Alzheimer's disease tissue.

A. Immunoprecipitations were performed on human AD tissue using 5 control and 5 test conditions, then analyzed by quantitative proteomics. Each fraction was compared to input and relative IP control.

B. Two of the antibodies used to purify APP and interacting proteins targeted the Aβ region and one targeted the NPXY motif within the AICD.

C. Description of grouping used for each of the samples used in this analysis.

Null comparisons were made for each protein between the input (IP) technical duplicates, between the unwashed fractions of Ab1 and Ab2 APP antibodies, and between the washed fractions of Ab1 and Ab2 APP antibodies. The IP duplicates exhibited high reproducibility, with values clustering tightly around a log2 value of zero (SD = 0.1), while the APP antibody null comparisons were centered around a log2 of zero but less tightly. The unwashed fractions varied slightly more than the IP comparison and the washed fractions exhibited the most systematic variability (SD = 0.36 and 0.75, respectively). This increase in variability is to be expected (as greater technical variability is introduced into the system with greater experimental manipulation), though we would still expect it to center around zero as the manipulations should affect the vast majority of proteins in a largely stochastic manner.

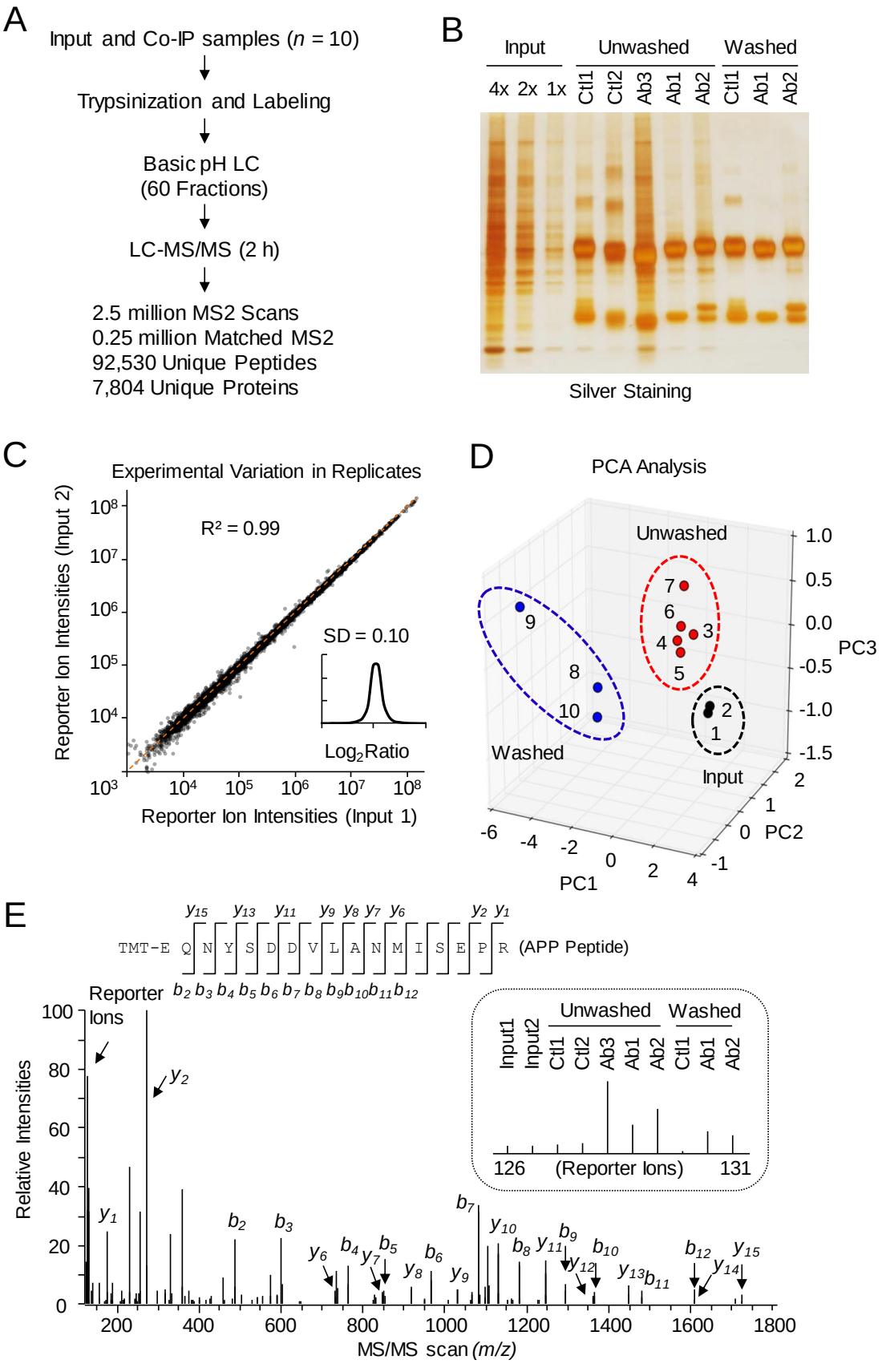

A
Input and Co-IP samples (n = 10)
Trypsinization and Labeling
Basic pH LC
(60 Fractions)
LC-MS/MS (2 h)
2.5 million MS2 Scans
0.25 million Matched MS2
92,530 Unique Peptides
7,804 Unique Proteins
B
Input
Unwashed
Washed
4x 2x 1x Ctl1 Ctl2 Ab3 Ab1 Ab2 Ctl1 Ab1 Ab2
Silver Staining
C
Experimental Variation in Replicates
R² = 0.99
SD = 0.10
Log2Ratio
Reporter Ion Intensities (Input 2)
Reporter Ion Intensities (Input 1)
D
PCA Analysis
Unwashed
Washed
Input
PC1
PC2
PC3
E
TMT-E Q N Y S D D V L A N M I S E P R (APP Peptide)
Reporter Ions
Relative Intensities
MS/MS scan (m/z)
Input1 Input2 Ctl1 Ctl2 Ab3 Ab1 Ab2 Ctl1 Ab1 Ab2
126 (Reporter Ions) 131

Figure 4.2. DEEP Co-Immunoprecipitation and LC-MS/MS Analysis

A. Experimental workflow demonstrating co-immunoprecipitation and thorough quantitative MS

B. Silver stain of samples used for proteomics analysis showing relative overall protein levels in each.

C. Correlation analysis of the input replicates exhibiting extremely strong correlation R^2 value of 0.99 and miniscule SD of 0.10.

D. Principle component analysis (PCA) chart displaying how overall patterns in the data group similarly prepared samples more tightly than those from other preparation methods.

E. A representative APP peptide sequenced in this analysis demonstrating a strong sequence ladder. Inset displays reporter ion values used for quantitation for this scan.

Enrichment

We measured traditional enrichment for each protein by comparing test antibody values to the average input value for that protein. As overall protein abundances between samples were matched, selectively precipitated proteins exhibited increases over their input levels. After fitting to a Gaussian distribution, we estimated enrichment mean and standard deviation (SD). The control IgG enriched proteins in a similar overall pattern as the test antibodies, with similar SD values for all three enrichments (Fig 4 A-B). In addition, there were substantial amounts of overlap between all 3 fractions. Nevertheless, both AB1 and AB2 strongly enriched APP, but control IgG did not. The APP $\log_2$ enrichment value for AB1 was 2.93, or 7.6-fold, and for AB2 was 2.12, or 4.3-fold, both of which are in the 96^{th} percentile of all proteins in their respective fractions, while IgG actually exhibited a negative $\log_2$ value for APP enrichment of -0.13. FSIP2, COL4A6, C1QA, and TRIM21 were among the proteins most highly enriched.

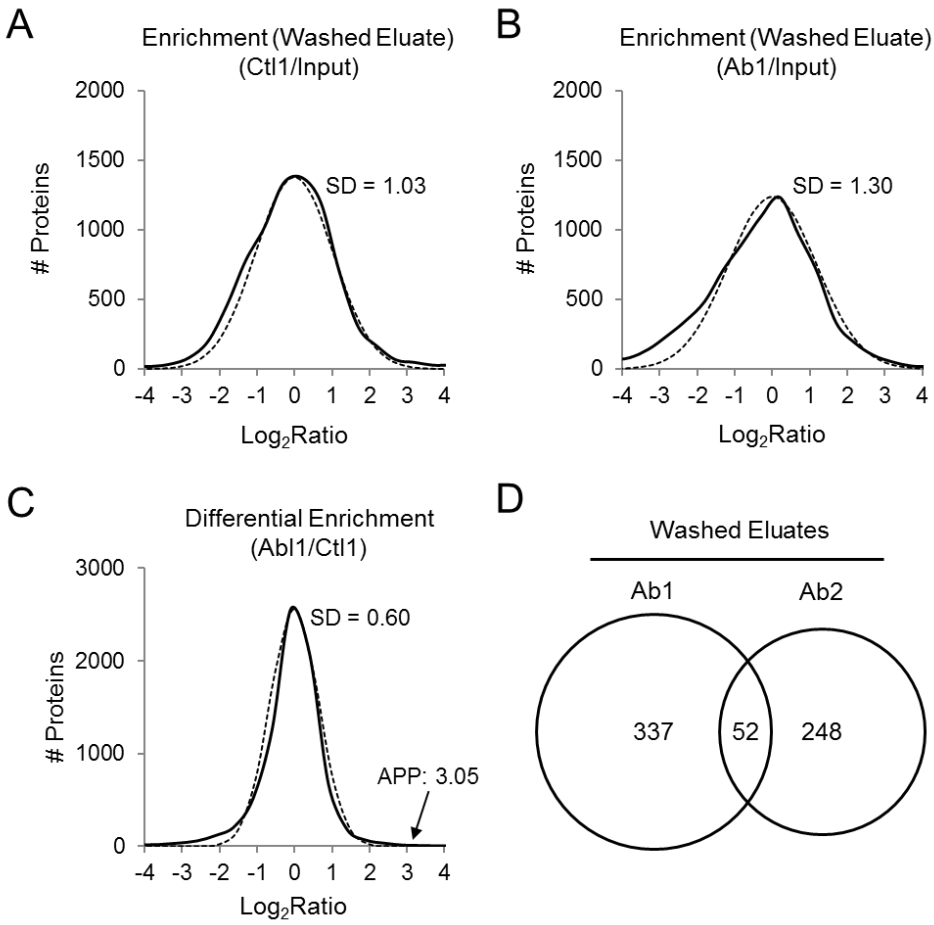


Figure 4.3. Histogram Analysis and Gaussian Fitting of Data from Ctl1 and Ab1 Comparisons.

A - C. Histograms depicting Enrichment for Ctl1 and Ab1, and Differential Enrichment for Ab1, plus the Gaussian curves which were fit to each comparison and used to calculate statistical significance.
D. Venn diagram of the washed eluates from Antibody 1 and Antibody 2 shows rather strong antibody specificity.

Differential Enrichment

We then measured differential enrichment by comparing between the test antibodies and the control IgG. Control IgG acts as a much stronger and relevant control for the experiment, as it accounts for nonspecific background binding to the beads. This was represented by the data when, after fitting them to a

Gaussian distribution (Fig 4), we once again observed strong overall reproducibility between the two samples, but also a slight systematic negative shift and strong tightening around the mean when compared to the enrichment analysis. Conversely, APP log_2 differential enrichment values increased, to 3.05 and 2.24 for AB1 and AB2, respectively. Both values are in the 99th percentile, with APP being the 16th and 34th most differentially enriched protein in the two fractions.

Using the fitted Gaussian parameters, we established a threshold z-score of 2 to determine significance and found 389 proteins differentially enriched by AB1, and 300 by AB2. Highlights of the proteins near the top of this list include GABRR2, DNAH2, RPLP1, and COX7A1.

Differential Elution

In order to include weakly associated – but potentially biologically relevant – binding partners that can be lost in wash steps, we extended our analysis via a differential elution approach. We accomplished this by performing the same co-immunoprecipitation, but with a minimally stringent wash. This included both test mouse ABs, the control mouse IgG, as well as both test and control rabbit antibodies, for which we performed no traditionally washed co-immunoprecipitation. For each of the three test antibodies, we used the relevant control antibody to analyze differential enrichment in the unwashed fractions, and again fit the data to Gaussian curves (Fig 5). Briefly, classical enrichment analysis – i.e. test Abs compared to input – showed a systematic positive shift from that of washed enrichment, as well as a tightening of the data around the

mean. The differential enrichment and elution analysis repeated these patterns – i.e. unwashed test Abs compared to unwashed control IgG – for these fractions as well.

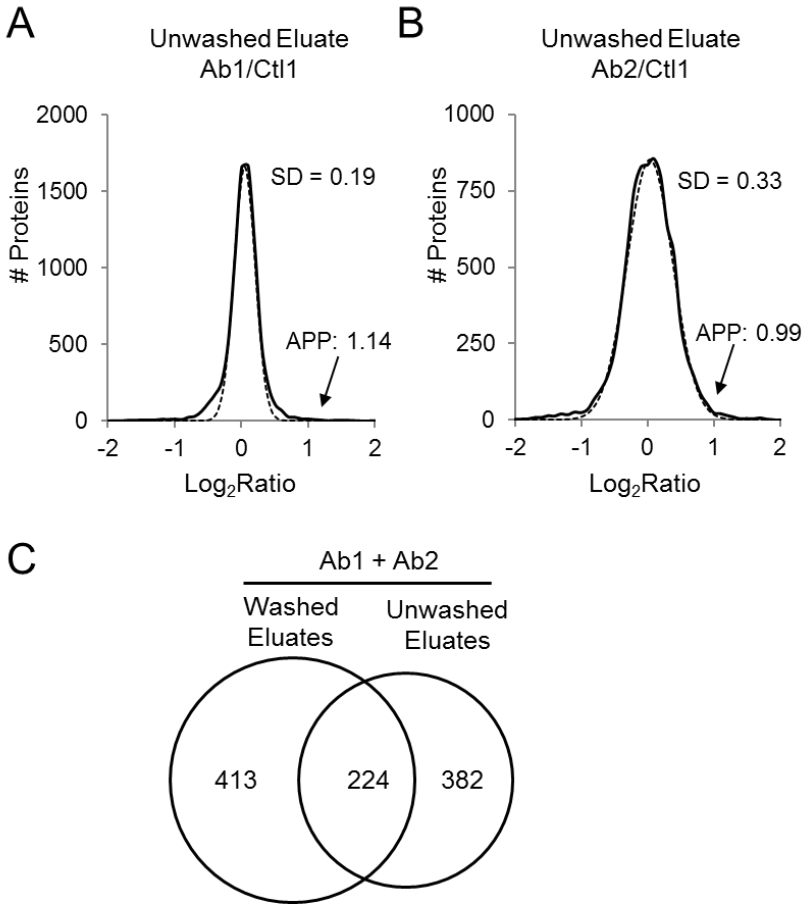


Figure 4.4. Histogram Analysis and Gaussian Fitting of Data from Differential Enrichment Comparisons for both Ab1 and Ab2.

A - B. Histograms depicting Differential Elution comparisons Ab1 and Ab2, plus the Gaussian curves which were fit to each comparison and used to calculate statistical significance.

C. Venn diagram displaying fairly strong overlap between overall washed and unwashed eluates.

The unwashed mouse antibodies strongly differentially enriched APP as well, particularly in comparison to the overall shape of the data, with $\log_2$ values

of 1.14 and 0.99 for Ab1 and Ab2, respectively, while the rabbit antibody differentially enriched APP with a $\log_2$ value of 0.41.

Using a threshold z-score of 2, we found 339 proteins were significantly differentially enriched by AB1 and 305 by AB2 in the unwashed mouse fractions. Additionally, roughly 50% of detected proteins were enriched in AB3, the rabbit fraction. Overall mouse antibody numbers resemble those of washed fractions due to our threshold method, and there is substantial overlap between different fractions of the same antibody. The numbers are so different for AB3 because the data distribution exhibited a bimodal pattern, and fitting this to a Gaussian model resulted in a mean well below zero, -0.73, and a high proportion of statistically significantly enriched proteins. In theory, all proteins in the washed fraction should be present in the unwashed, however, the presence of nonspecific background prevented us from relaxing thresholds to include all washed fraction interactors – though many had positive but subthreshold enrichment values. Proteins highly enriched in this analysis included HMX3, CBLN1, and BPTF.

	Enrichment			Differential Enrichment				
	Mean	SD	APP log_2	Mean	SD	Threshold log_2	Threshold Fold	APP log_2
mIgG	*0.45*	*1.10*	*-0.12*	-	-	-	-	-
mAb1	*0.25*	*1.30*	*2.93*	*-0.13*	*0.50*	*0.78*	*1.71*	*3.05*
mAb2	*0.05*	*1.10*	*2.12*	*-0.15*	*0.60*	*0.93*	*1.91*	*2.24*
Differential Elution								
$mIgG_{unw}$	*0.30*	*0.55*	*0.37*	-	-	-	-	-
$mAb1_{unw}$	*0.25*	*0.60*	*1.51*	*0.00*	*0.19*	*0.34*	*1.27*	*1.14*
$mAb2_{unw}$	*0.63*	*0.40*	*1.35*	*-0.01*	*0.33*	*0.58*	*1.50*	*0.99*
$rIgG_{unw}$	*0.28*	*0.60*	*0.35*	-	-	-	-	-
$rAb3_{unw}$	*0.10*	*0.80*	*0.76*	*-0.73*	*0.35*	*0.93*	*1.91*	*0.41*

Table 4.1. Summary of Key DEEP Analysis Values

Multiple Antibody DEEP Analysis

Each individual dataset is informative to the function and pathogenic capabilities of APP, and the value of our approach is further enhanced by combining them. We examined overlap between the differential enrichment and elution analyses, between the different antibodies, and stratified the interactome based on these different types of agreement. The exceedingly high numbers from the AB3 fraction precluded us from assigning it the same value as the AB1 and AB2 fractions. Rather, we used it to supplement those proteins also enriched by AB1 and AB2. Using the differential enrichment data from AB1, AB2, and AB3, plus both elution conditions for AB1 and AB2 provided 5 fractions and 1019 proteins in total with which to examine overlap (Fig 6).

The topmost tier, 1A, is comprised of seven proteins that were significantly differentially enriched in all five comparisons, and represents the highest confidence putative interaction partners. One step down, in tier 1B, are two proteins that were differentially enriched in all four mouse AB comparisons, but

not in AB3. Tier 2A is comprised of 23 proteins that were differentially enriched in three mouse AB fractions plus AB3, and tier 2B of 8 proteins differentially enriched in three mouse AB fractions but not AB3. Tier 3 is composed of 118 proteins differentially enriched in 2 mouse AB fractions plus AB3, while tier 4 rounds out our analysis with 536 differentially enriched in 2 of any of the 5 fractions.

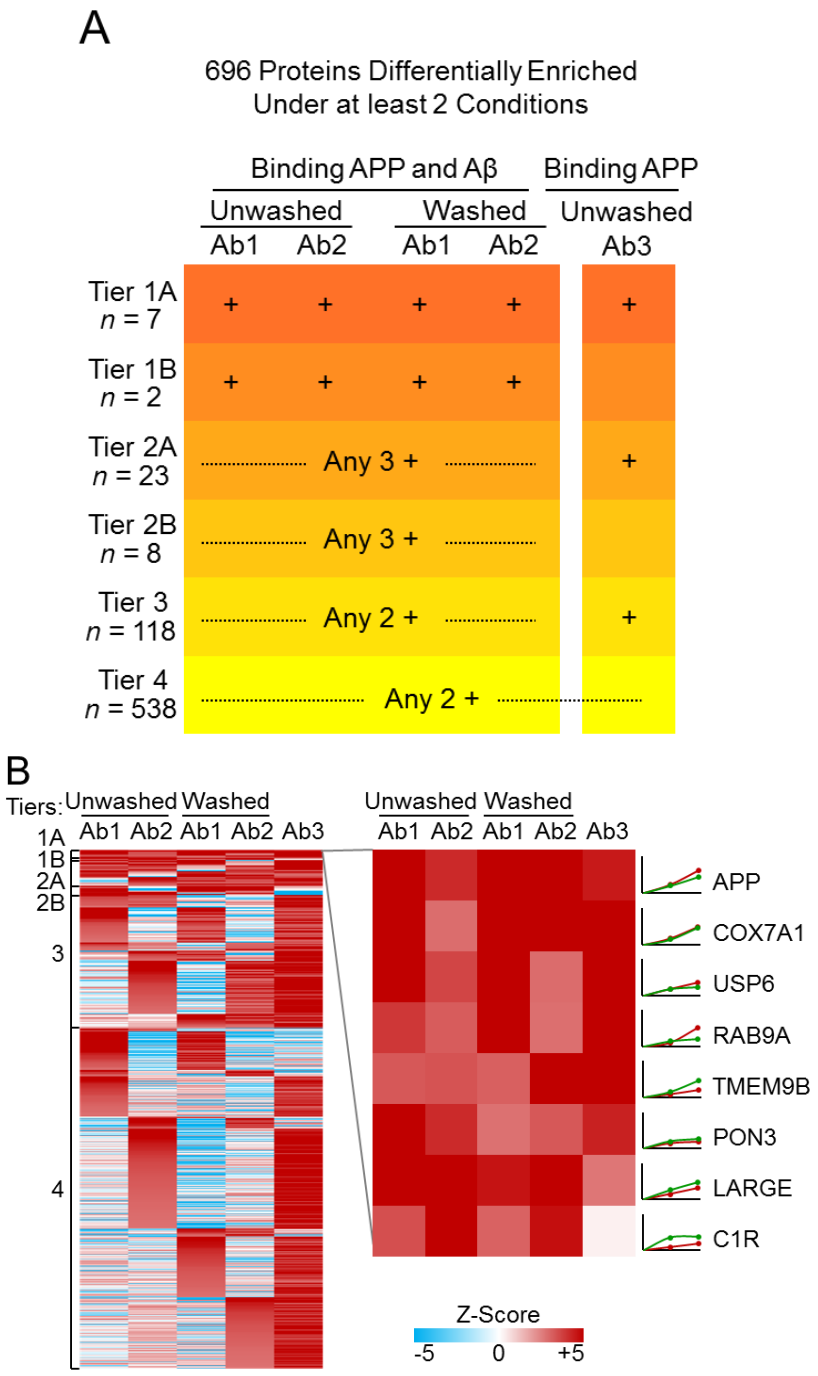


Figure 4.5. Multiple DEEP analysis

A. Graphic representation of the rules used in the multiple DEEP analysis primarily utilizing the more stringent mouse Abs 1 and 2 comparisons with supplement from the rabbit Ab 3 comparison to determine tier, and essentially confidence in interactoin. B. Heatmap depicting a summary of the 696 proteins assigned to tiers 1-4 and highlighting 8 of the proteins from Tiers 1A and 1B.

Pathway and Network Analysis

To understand the biological processes in which APP takes part, we performed cell compartment, pathway, gene ontology, and protein domain enrichment using DAVID v6.7 online web-accessible functional annotation tools, and PANTHER online gene ontology classification utilities. Mitochondria, the nucleus, the cytoskeleton, and various membranes in the cell top the list of most enriched cell compartments in our dataset. The most highly enriched pathways and gene ontologies include the complement cascade, ribosomes, spliceosomes, transcription regulation, signal transduction and kinase cascades, and metal ion homeostasis. Within the 40 proteins in tiers 1 and 2, pathways such as Aβ peptide metabolism, neural development, synaptic regulation, trophic signaling, intracellular trafficking, mitochondrial function, and RNA splicing were strongly enriched. Selected differentially enriched proteins from tiers 1-2 in selected compartments and pathways are highlighted in **Table 4.2**. Closer examination of the proteins within these individual networks reveals abundant interconnection between putative APP interaction partners, which suggests incorporation of APP into multi-protein complexes and reinforces the likelihood for a role of APP in these processes.

Top-tier proteins isolated by APP antibodies from Alzheimer's brain

				Differential Enrichment (*z* Score)[2]				
				Unwashed Eluates		Washed Eluates		*Unwashed*
Tiers[1]	Accession#	GN	Protein names	Ab1/Ctl1	Ab2/Ctl1	Ab1/Ctl1	Ab2/Ctl1	*Ab3/Ctl2*
Aβ peptide metabolism								
1A	P05067	APP	Amyloid precursor protein	6.01	3.02	6.34	3.98	*3.26*
1B	P00736	C1R	Complement C1r subcomponent	2.50	5.45	2.22	3.41	*0.21*
2B	Q9NZP8	C1RL	Complement C1r subcomponent-like protein	3.47	5.46	2.06	1.92	*-2.76*
2B	P23435	CBLN1	Cerebellin-1, a secreted protein similar to C1q	3.81	8.42	-1.13	3.52	*-2.99*
Neural development and synaptic regulation								
2A	Q9ULB5	CDH7	Cadherin-7	-0.72	2.89	2.30	2.31	*2.82*
2A	P29033	GJB2	Gap junction beta-2 protein	4.88	2.75	4.52	-1.32	*3.09*
2A	Q9H252	KCNH6	Potassium voltage-gated channel subfamily H member 6	2.75	2.62	0.99	2.43	*5.47*
2B	P28476	GABRR2	Gamma-aminobutyric acid receptor subunit rho-2	20.20	-2.51	22.66	13.11	*0.28*
2B	Q68DQ2	CRYBG3	Beta/gamma crystallin domain-containing protein 3	4.89	-4.64	10.39	4.33	*0.08*
Trophic signaling activities								
1A	Q9Y6X0	SETBP1	SET-binding protein	4.04	7.50	4.09	5.64	*4.33*
2A	P16591	FER	Tyrosine-protein kinase Fer	2.02	4.10	-0.29	2.81	*3.52*
2A	Q12830	BPTF	Nucleosome-remodeling factor subunit BPTF	2.19	5.04	3.91	1.54	*3.97*
2A	P82970	HMGN5	HMG nucleosome-binding domain-containing protein 5	10.21	7.66	3.46	0.65	*4.32*
2A	Q8IVB5	LIX1L	LIX1-like protein	3.22	0.77	4.62	2.16	*2.12*
2A	Q8WY91	THAP4	THAP domain-containing protein 4	3.37	0.24	25.17	23.66	*2.13*
2B	Q9C029	TRIM7	Tripartite motif-containing protein 7	20.11	0.95	5.34	2.21	*-1.66*
2B	Q69YN4	KIAA1429	Protein virilizer homolog	10.52	4.81	5.63	1.50	*-6.84*
Intracellular trafficking								
1A	P51151	RAB9A	Ras-related protein Rab-9A	2.84	2.31	5.49	2.04	*3.99*
1A	P35125	USP6	Ubiquitin carboxyl-terminal hydrolase 6	4.79	2.64	3.78	2.09	*4.36*
2A	P67936	TPM4	Tropomyosin alpha-4 chain	4.73	2.63	-4.40	3.14	*2.03*
2A	P55735	SEC13	Protein SEC13 homolog	1.49	2.00	2.37	2.14	*3.00*
2A	P51636	CAV2	Caveolin-2	1.87	4.75	2.36	3.23	*2.73*
2A	H0YJ05	STON2	Stonin-2	-0.62	3.46	2.34	2.34	*6.10*
2A	Q9Y2Y0	ARL2BP	ADP-ribosylation factor-like protein 2-binding protein	3.18	-1.32	4.99	3.03	*2.42*
2A	Q68CL5	TPGS2	Tubulin polyglutamylase complex subunit 2	2.78	2.89	1.12	5.10	*4.41*
2B	Q8IWB7	WDFY1	WD repeat and FYVE domain-containing protein 1	2.20	-0.04	4.20	3.88	*1.06*
Mitochondrial function								
1A	P24310	COX7A1	Cytochrome c oxidase subunit 7A1, mitochondrial	4.39	2.07	4.98	4.02	*6.00*
2A	Q9Y241	HIGD1A	HIG1 domain family member 1A, mitochondrial	6.54	1.77	4.55	2.76	*4.58*
2A	P0DJ07	PET100	Protein PET100 homolog, mitochondrial	-1.35	3.27	4.88	2.60	*7.27*
RNA splicing								
2A	P62314	SNRPD1	Small nuclear ribonucleoprotein Sm D1	2.87	1.05	2.93	2.55	*2.84*
2B	P62308	SNRPG	Small nuclear ribonucleoprotein G	3.82	0.46	4.15	3.50	*1.10*
Others								
1A	Q9NQ34	TMEM9B	Transmembrane protein 9B	2.36	2.44	2.25	4.19	*4.79*
1A	Q15166	PON3	Serum paraoxonase/lactonase 3	3.68	3.01	2.01	2.36	*3.15*
1B	O95461	LARGE	Glycosyltransferase-like protein LARGE1	3.91	3.88	3.33	4.18	*1.94*
2A	H3BR08	PMM2	Phosphomannomutase	3.82	2.04	3.57	1.34	*3.71*
2A	Q5T6J7	IDNK	Probable gluconokinase	2.03	3.39	2.08	1.42	*3.75*
2A	Q8WUY8	NAT14	N-acetyltransferase 14	0.20	5.02	2.80	3.32	*5.30*
2A	A0A087X2F8	ACACA	Acetyl-CoA carboxylase 1	1.68	2.36	2.45	2.35	*3.02*
2A	Q8NBT3	TMEM145	Transmembrane protein 145	2.95	2.23	0.96	2.17	*3.71*
2A	Q5BJH2	TMEM128	Transmembrane protein 128	0.49	2.28	4.01	6.05	*3.41*

[1]The Tier is assigned by counting the experimental conditions under which a protein is enriched (see details in Experimental Procedures).

[2]The *z* scores are derived on standard deviations in individual comparisons.

Table 4.2. Highlighted Pathways and Contained Proteins from Tiers 1 and 2.

DISCUSSION

Despite APP being the most widely studied single protein in biological research, its physiological roles and the pathogenic perturbations that lead to AD largely remain mysteries. There remains no efficacious treatment for AD, which is

the 6th leading cause of death (56), most prevalent cause of dementia (57) and it is only becoming more prevalent (59). Several high-throughput studies have examined APP interaction partners, but AD and APP itself present several challenges. No single animal model of AD – and there are many – fully recapitulates the pathology (163). The fact that even the early onset varieties do not appear until 45-60 years of age (164) makes it difficult to model in a mouse, and the language and nuanced cognitive deficits are of course nearly impossible to model (165).

Biochemically, APP is a transmembrane protein, which are notoriously difficult to purify and analyze, and the aggregates formed in the disease state are vast and even more intractable assemblies of proteins. Further, endogenous APP undergoes significant and intricate post-translational processing, which results in several distinct protein fragments. Amyloidogenic $A\beta_{42}$, the most well-known and disease-causing cleavage product, results from γ-and β- secretase cleavage, while numerous non-amyloidogenic cleavage events also occur in the same region. Meanwhile, a cleavage product from another region, the AICD intracellular domain, is cleaved from mature membrane-bound APP and translocates to the nucleus where it is affects transcription (151). These properties combine to produce a disease pathogenesis that appears to be highly specific, with effects that are widespread, and by processes that are difficult to study.

We sought to improve understanding of the APP interactome in a human AD context with three major focuses. First – differential enrichment – to strictly

control the co-IP/MS results by comparing test antibodies to conspecific IgG antibody controls. This control helped limit false positives resulting from abundant proteins, as well as from generally ‘sticky’ proteins that bind the column nonspecifically. Our second focus – differential elution – was to capture weak interaction partners and stratify a combined analysis of these with traditional co-IP/MS results. By doing this, we broadened the potential interaction partner pool to previously undetected low affinity binding partners, while also deepening it by allowing us to differentiate within the putative interactome via binding strength. Our third focus was to compare and examine overlap between co-IP/MS results from multiple antibodies. Doing this improved the analysis in two ways: it essentially created multiple simultaneous analyses from similar – but not identical – angles and allowed us to directly compare these analyses to reinforce overlapping results over disparate results. The widespread lack of precision and preponderance of off-target cross-reactivity from antibodies (166) – especially at a systems level (167) – necessitated this type of control and cross-validation.

Our differential enrichment method attained promising results – significantly improving the overall dataset and boosting the APP enrichment value for APP. As compared to the enrichment comparisons for AB1 and AB2, the means and SDs of the differential enrichment comparisons were substantially lower: means were 0.25 and 0.05 versus -0.13 and -0.15, and the SDs were 1.30 and 1.10 versus 0.50 and 0.60 – with a value of zero representing no change. The control IgG values for the vast majority of individual proteins in the 7804 we quantified – which can generally be assumed to be non APP interacting proteins

– were more elevated than the input values for these same proteins. These proteins were simply enriched due to nonspecific binding to the column and/or antibodies, resulting in a system-wide negative shift in the differential enrichment comparisons because the values were therefore higher in proportion to the test antibody values – in which they are presumably nonspecifically enriched. Meanwhile, the values for APP were increased: 2.93 and 2.12 vs 3.05 and 2.24, respectively. Values for true APP interacting proteins were presumably similarly increased due to enrichment through binding to APP, which eliminated many false positives. While it almost certainly created some false negatives, we expect the number to be negligible.

Our differential elution strategy produced promising and interesting results. Overall, the differential enrichment comparisons of the unwashed fractions were much more tightly clustered around a $\log_2$ value of zero, with very small SDs – means for mouse antibody fractions were 0.00 and -0.01 and SDs were 0.19 and 0.33. Therefore, unwashed IgG fractions were highly similar to unwashed test antibodies systemwide, a result we expected due to the simple fact that washing of co-IP fractions is an additional technical manipulation that will increase variation on this scale. Two unwashed co-immunoprecipitations are more similar to the input, and therefore to each other, than to washed fractions.

While there was sizable overlap between the unwashed and washed elutions, we were surprised that it was not larger. We expected most, if not all, proteins that were significantly differentially enriched in the washed fractions to at least have positive differential enrichment values in the unwashed fractions –

with additional low-affinity binding partners also showing positive values. This was not the case for many proteins, however, and there were certainly a large number of proteins that were not significantly differentially enriched in both fractions. We interpreted the fact that positive results could be negative in the unwashed fraction to mean that they may be false positives, and certainly proteins enriched in the unwashed fractions, but not the washed, must subject to stringent control. Therefore, in order to maintain statistical rigor, we placed equal weight on each fraction when combining the analyses, and used degrees of overlap between the fractions as the primary determinant of confidence.

There was also great deal of overlap, within comparisons between antibodies – but again it was not complete. Again this was expected, and was the idea behind examining co-IP using multiple antibodies. As the two mouse antibodies, AB1 and 2, performed similarly, and with matching stringency, we were able to weight them equally. In contrast, AB3, the rabbit antibody produced a preponderance of background and drove us to assign it a slightly diminished value in the overall analysis.

Taken together, the lack of overlap between each of the different analyses we performed helped display the effectiveness of our strategy. Each strategy on its own would have produced vast numbers of both false positive and false negative results. Only by performing these simultaneously, and in parallel, were we able to compare, contrast, and examine agreement between them, and this created a more inclusive analysis, but also allowed us to differentiate within our broadened results.

The pathways and gene ontologies most highly enriched include Aβ peptide metabolism, the complement cascade, neural development, synaptic regulation, trophic signaling, signal transduction and kinase cascades, mitochondria, ribosomes, spliceosomes, transcription regulation, intracellular trafficking, and metal ion homeostasis. Enrichment of these compartments, pathways, and ontologies aligns well with current understanding of the APP protein. APP has previously been shown to alter cytoskeletal (168), mitochondrial (169), and membrane dynamics (170), and it is a transmembrane protein with a domain that is known to be cleaved and transported to the nucleus where it acts to regulate transcription (171). Oxidative stress, inflammation, and the complement cascade, in particular, are strongly implicated in AD pathogenesis and progression (172), and RNA splicing has become a focus of AD research of late (95,173,174). Membrane and endosomal sorting may be important to AD pathogenesis and amyloidogenic processing of APP (175).

CHAPTER FIVE

Perturbations in the N40K Transgenic Mouse Proteome

Alzheimer's disease (AD) is the most common form of dementia and the sixth-leading cause of death in the U. S., affecting over 5 million people in the U. S., costing sufferers over $200 billion in care, and it is only becoming more prevalent (58). Insoluble protein aggregates are a prominent feature of AD pathology and the identification of Aβ and tau proteins in these aggregates provided the impetus for the bulk of AD research thereafter (140,176). A number of mouse models have been developed around this primary hypothesis of the amyloid cascade (142), but no one model exhibits AD pathology in entirety. Further, there remains no effective treatment, and our knowledge of pathogenesis is incomplete, which means proteomics technology and unbiased approaches have great potential to shed new light.

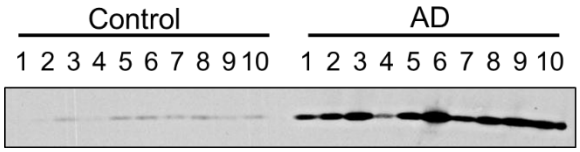
To this end, our group previously undertook several proteomic and subproteomic studies of AD tissue addressing PSDs, the detergent-insoluble

proteome, senile plaques using laser capture microdissection, polyubiquitin linkage, and the phosphoproteome (177-182). The detergent insoluble proteome holds particular interest in AD research because misfolded and aggregated proteins are often found in this fraction. A recent study by our group on the detergent-insoluble proteome of AD and several other neurodegenerative diseases in comparison sought to more fully characterize abnormally aggregated proteins in AD. We identified multiple subunits of the U1 small nuclear ribonucleoprotein (U1 snRNP), a constituent of the spliceosome complex, to be enriched in the AD insoluble proteome (173). Further examination in this study demonstrated widespread cytoplasmic aggregation of both U1-70K and U1A snRNP in neurons and global alterations in RNA processing in human AD brains via deep RNA sequencing. Moreover, U1-70K knockdown increases APP and Aβ levels in cellular models and U1-70K aggregates are present in cases of mild cognitive impairment, which can be a precursor of AD, suggesting that U1 perturbations may occur early in AD pathogenesis.

A

Accession#	Protein names	Spectral counts[2] Ctl1	Ctl2	AD1	AD2
Amyloid-beta peptide metabolism					
NP_000475.1	amyloid beta peptide[3]	9	31	169	196
NP_000032.1	apolipoprotein E	1	1	49	92
NP_115907.2	collagen, type XXV, alpha 1 isoform 2	1	0	23	24
NP_004369.1	cellular retinoic acid binding protein 1	0	0	9	7
Cytoskeleton maintenance					
NP_058519.2	microtubule-associated protein tau	10	11	824	989
NP_116757.2	dystrobrevin alpha	0	11	23	24
Inflammation					
NP_009224.2	complement component 4A preproprotein	7	7	77	128
NP_001002029.3	complement component 4B preproprotein	7	7	81	163
NP_000055.2	complement component 3	1	2	57	93
RNA splicing					
NP_003080.2	U1 small nuclear ribonucleoprotein 70 kDa	2	2	31	39
NP_004587.1	U1 small nuclear ribonucleoprotein A	0	1	12	22
NP_055644.2	ATP-dependent RNA helicase DDX46, Prp5	0	0	9	17

B



C

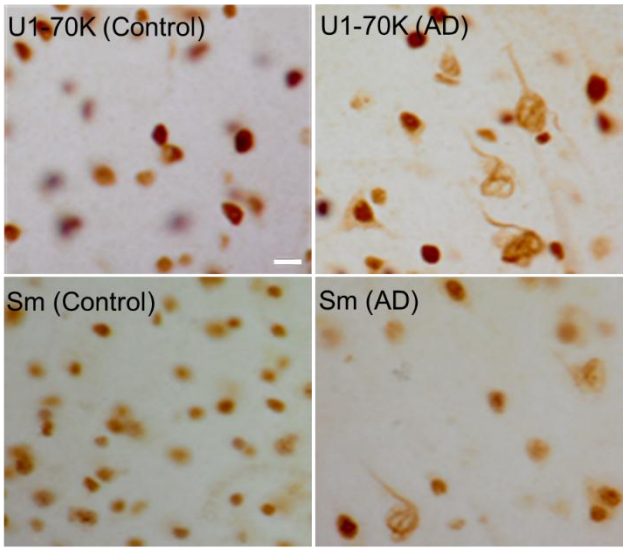


Figure 5.1. Spliceosome subunits are enriched in Alzheimer's disease.

Unpublished work from our group.

(A) Summary of AD proteomics analysis depicting strong enrichment of U1-70K and U1-A in human AD brains.

(B) Western blot analysis confirming a strong increase in U1-70K levels in multiple AD brain samples.

(C) Immunohistochemical analysis demonstrating U1-70K localization outside its canonical nuclear cellular compartment in tangle-like structures in AD brains. Sm, a spliceosomal protein shared between other snRNP complexes demonstrates similar perturbations.

During this analysis, we detected two distinct forms of U1-70K, a full length 70 kDa species and a truncated 40 kDa species we dubbed N40K (95). This truncation cannot be explained by alternative splicing, and appears to be the result of protease cleavage, a phenomenon common in neurodegenerative disease (183). Further study focused on this truncation discovered it contains the very first 300 amino acids of the full length form, its cleavage site, that it correlates with downregulation of the full-length species, and that it is toxic to primary neurons (95).

U1-70K is one of ten protein subunits of the U1 snRNP complex, which is itself one of the 5 major components of the spliceosome – the others being U2, U4, U5, and U6 (184). U1 recognizes the 5’ splicing sites in pre-mRNA and is crucial to initiation of the splicing process and recruitment of the other components (185). Other than U1-70K, U1 consists of two other proteins unique to it, U1-A, and U1-C, seven Sm proteins that are common between the other snRNP components, plus a small nuclear RNA (snRNA) (Fig 5.2A). U1-70K, a 437 aa protein, contains a largely disordered N-terminal, a RNA recognition motif between residues 99 and 181, and two low complexity domains from amino acid 231 to the c-terminus (Fig 5.2B) (186). The N40K species retains all but the final low complexity domain, RD2. It has become increasingly clear that splicing perturbations associate with and may contribute to AD. The splicing factor TDP-43 was recently found to be hyperphosphorylated and accumulated in the insoluble fraction in a subset of AD (187), and RNA-Seq performed on AD whole brain, frontal, temporal, and parietal lobe tissues uncovered widespread splicing alterations in AD-affected regions, including lipid metabolism genes and APOE (188,189). Thus, our discovery of disruptions in U1-70K and other U1 snRNP subunit is of particular interest.

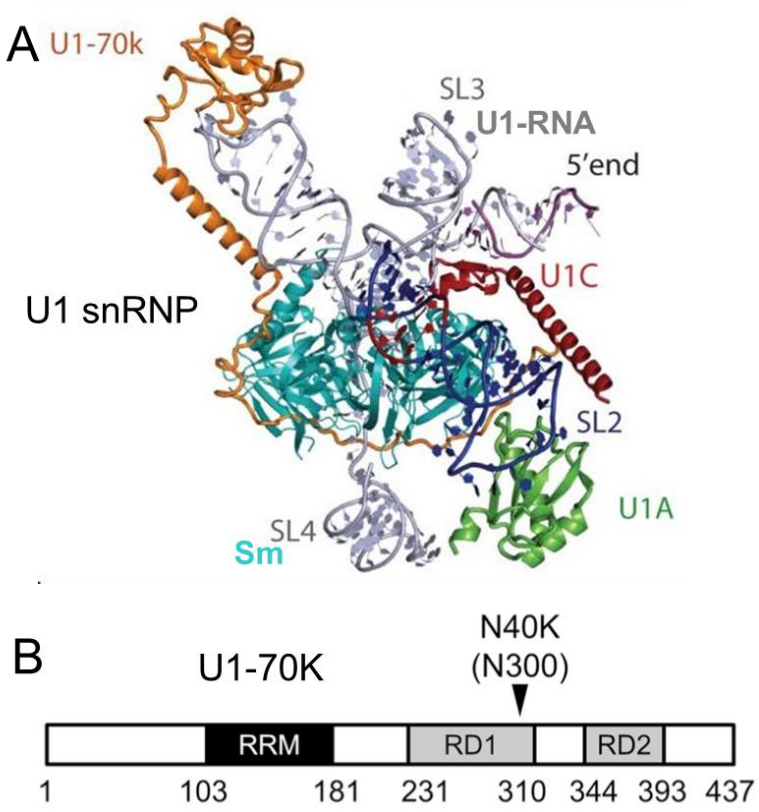


Figure 5.2. U1-70K is a subunit of the U1 snRNP complex and contains several defined domains.

(A) From Pomeranz Krummel et al. (184), showing overall structure of the U1 complex, with U1 specific proteins U1-70K, U1-A, and U1-C in addition to shared Sm proteins. (B) From Bai et al (95), depicting U1-70K primary structure, including RNA recognition motif (RRM) from residues 103-181, plus 2 low-complexity domains (RD1 and RD2) between residues 231 and 393. The N40K protein contains up to amino acid 300, lacking the final RD2 domain.

To further address the pathogenic properties of the N40K fragment, our lab has since developed a transgenic mouse expressing it under the control of the CaMKIIα promoter (Fig 5.3A). The CaMKIIα promoter was selected because it provides nearly complete high level neuronal expression in cortex and hippocampus, the regions most directly affected in AD. This mouse not only exhibits strong expression of the N40K fragment in forebrain regions, but also significant downregulation of the full length U1-70K, a result that directly matches

observations in AD brain tissue (Fig 5.3B-C). Further, behavioral analysis of the transgenic has shown that both spatial learning and working memory are significantly disrupted in aged animals (Fig 5.3D-E), again closely matching the properties of AD and AD mouse models (163).

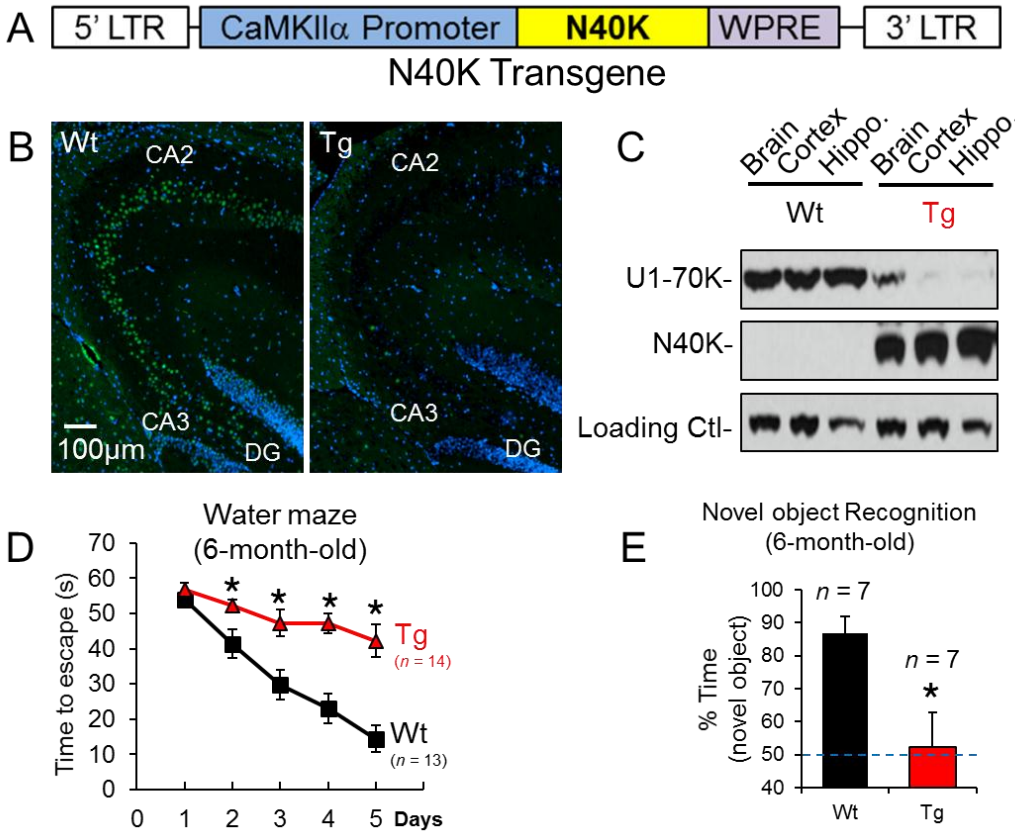


Figure 5.3. The N40K transgenic mouse exhibits decreased full length U1-70K as well as deficits in spatial learning and working memory.

(A) Transgene design for the N40K transgenic mouse including the CaMKiiα promoter and woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), which both lead to abundant gene expression.
(B) Immunohistochemistry probing the C-terminal portion of U1-70K in animals 6 mos of age, demonstrating decreased full length U1-70K in the Tg animals.
(C) Western blot at 6 mo, showing N40K expression and further demonstrating decreases in the full length protein in total brain, cortex, and hippocampus.
(D) Morris water maze results demonstrating spatial learning deficiency in 6 mo N40K animals as compared to age-matched controls.
(E) Novel object recognition test results demonstrating working memory deficits in 6 mo N40K animals.

As part of the thorough characterization of the N40K transgenic mouse, we have also undertaken several systems-based analyses including RNA-Seq, and LC-MS/MS analysis of the total and phosphoproteomes. For whole proteome

analysis we utilized 3 different ages: 3, 6, and 12 mos, and 2 different brain regions: hippocampus and cortex (Fig 5.4) in order to characterize age-related and region-specific proteomic responses to N40K expression. The core of our analysis, the 3 age-specific comparisons at 3, 6, and 12 mos of age uncovered 492 total proteins with either significant increase or decrease in the N40K animals, with 19 of these enriched at all 3 time points – including strong increases in all three U1-unique subunits: U1-70K, U1-A, and U1-C. Overall comparison of all Tg and WT samples as well as age-related changes from 3 to 12 months of age uncovered further alterations in the N40K proteome and potential molecular underpinnings to differences seen at the organism level.

MATERIALS AND METHODS

Mouse Brain Tissue Lysate Preparation and Western Blot

Frozen mouse brain tissue at 3 mos, 6 mos, and 12 mos of age was homogenized in in 8M Urea, 50mM HEPES,10% Acetonitrile, and 1x phosphatase inhibitor at 8.5 pH at a ratio of 10mg tissue to 100ul lysis buffer. Protein lysate samples were split into 2 or more aliquots. Protein concentration and integrity were measured by BCA and SDS-PAGE, respectively. Western blot was performed using rabbit polyclonal N120 antibody at 1:250 dilution in PBST containing 3% BSA. This antibody was developed previously by our group.

10-plex TMT-based quantitative LC-MS/MS analysis

Samples were digested by LysC and Trypsin. LysC was added to the protein samples at a 1:50 enzyme:substrate ratio and incubated at room

temperature for 2 h. After dilution by 50 mM HEPES, trypsin was added at 1:50 enzyme:substrate ratio and incubated overnight. After verifying complete digestion, the reaction was quenched by desalting using micro spin column (Harvard apparatus) and digested peptides were extracted, dried, and resuspended in 50 µl 50mM HEPES buffer, pH 8.5. TMT 10-plex reagents (ThermoFisher Scientific) were added to each vial according to the commercial instructions and incubated for 1 hr at room temperature and equally mixed. After drying, the labeled peptide mixture was desalted on ziptip column (EMD Millipore). Phosphopeptides were extracted from the samples by 5xTiO_2 bead purification at 1:3 total sample:bead ratio and stored at -80° C until LC-MS/MS phosphoproteome analysis. Total proteome peptides were then fractionated on a XBridge C18 column (Waters, 3.5 µm resin, 4.6 mm x 250 mm), by a gradient of solvent B (90% acetonitrile in 10 mM ammonium formate, pH8.0): 5% for the first 5 min, then steadily increased to 15%, 25%, 35%, 45%, 65%, and 95% at 7, 22, 111, 155, 175, and 180 min respectively. Solvent A (10 mM ammonium formate, pH 8.0) was used to generate the solvent B gradient. In total, 100 fractions were collected and dried.

The peptides in 5% formic acid were separated by a C18 column (75 µm × 30 cm, 1.9 µm resins) at 0.22 µl/min flow rate. The main gradient increased steadily from 15% to 45% buffer B (65% acetonitrile, 0.2% formic acid, and 5% DMSO) for 210 minutes, followed by steep increase to 95% buffer B for 30 more minutes. The gradient was generated by buffer A (0.2% formic acid and 5% DMSO) and was adjusted slightly between samples according to LC

chromatography elution profile. Eluted peptides were scanned on the Orbitrap Velos MS analyzer (Thermo Fisher Scientific) by MS1 scan (60,000 resolution 1×10^6 automatic gain control, and 50 ms maximal ion time), followed by top 10 MS2 scans (HCD collision with 38 normalized collision energy, 60,000 resolution, 1×10^5 automatic gain control, 100 ms maximal ion time).

39 MS raw files were acquired and converted into mzXML format for peptide matching by the JUMP algorithms (91). A combination of target database from Uniprot and a decoy database generated from reversed peptide sequences of the target database was used. The mass (+229.162932) of TMT tags was added to peptide N-termini and Lysines as static modifications. Only a, b, and y ions were chosen for peptide spectra match (160).

Protein Quantification by TMT Labeled Peptides

Quantification of TMT labeled peptides was carried out by an in-house program in the following steps. (i) TMT reporter ion intensities of each identified PSM were extracted and recorded. (ii) The raw intensities were corrected according to isotopic distribution of each labeling reagent (iii) The average of all ten reporter ion signals was used as a reference to compute a relative intensity between each sample and the average. (iv) The relative intensities of PSMs were averaged for identified proteins. (v) The combined reporter ion intensities were used to calculate the Tg:WT $\log_2$ ratios for 3 month, 12 month, and the average for the duplicates of 6 month old hippocampus, plus 12 month old cortex. Additional comparisons were made for both overall changes by taking the $\log_2$ ratio of all 5 transgenic fractions over all 5 wt fractions and for age-related

changes by taking the $\log_2$ ratio of 12 month Tg:WT over 3 month Tg:WT hippocampus fractions. (vi) Visualizing these comparisons via histogram analysis, the data were fit to Gaussian distributions, and thereby mean and standard deviation estimated. (vii) To analyze experimental variation, the intra-sample comparisons between the input replicates were examined as null experiments, showing an average standard deviation of 0.23. (viii) Finally, Z-scores were calculated using the fitted Gaussian parameters, and we selected a Z-score cutoff of 3 for comparing different samples

Interaction Network Analysis

Enrichment of Kyoto Encyclopedia of Genes and Genomes pathways (KEGG, (92)), and Gene Ontology terms was determined by analyzing the dataset using the NIH DAVID Bioinformatics Resources 6.7 Functional Annotation Clustering tool (93). To determine enrichment in this study we established a threshold of 1 for enrichment score, which is –log of the geometric mean of p-values of clustered pathways and ontologies.

RESULTS

Western Blot and LC-MS/MS Analysis of the N40K Brain Proteome

Previous unpublished studies by our group have demonstrated molecular and behavioral perturbations in the N40K transgenic mouse. As part of an ongoing thorough characterization of this animal model, we sought to analyze its hippocampal proteome at 3, 6, and 12 months of age and its cortical proteome at 12 months. Comparisons between N40K and WT mice, N40K mice of different

ages, and different forebrain structures afforded a broad examination of the protein perturbations that underlie behavioral deficits seen in this animal and potential links to how and why this protein species is observed in human AD sufferers.

Each fraction of the tissue was first analyzed by western blot prior to LC-MS/MS analysis, thus confirming expression of the N40K transgene and that N40K levels were relatively stable across age and brain region (Fig 5.4). Further, it appears that full length U1-70K is also reduced in the transgenic animals, which has been seen previously in this mouse model, and is similar to observations in AD tissue.

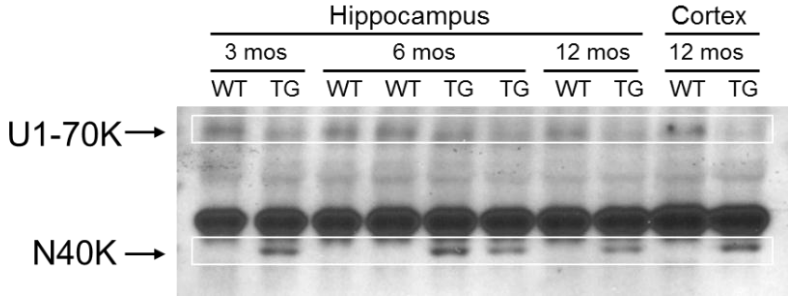


Figure 5.4. N40K mouse brain lysate samples used in this study exhibit even levels of transgene and decreased U1-70K levels.

Samples used for proteomics analysis were examined by western blot and probed using the N120 antibody, which targets an antigen from residues 99-120 of U1-70K/N40K.

LC-MS/MS examination of this tissue produced identification and quantification of 9417 proteins from mouse hippocampus and cortex at a false discovery rate less than 1%. U1-70K peptides were identified 14 times in the analysis, providing 18.97% coverage of the overall protein. As per previous findings (95), all of the sequenced peptides originated in the region also encompassed by N40K (Fig 5.5A). This result is likely due to the C-terminal

peptides being incompatible with MS analysis and/or an undersampling issue by which they were simply not selected by the machine for sequencing, as opposed to biological phenomena. It should be noted that the N40K species was originally discovered via SDS-PAGE separation followed by LC-MS/MS analysis (173), a technique that provides the approximate molecular mass of the protein from which MS sequenced tryptic peptides originated. Identification of U1-70K peptides in the 40 kD region led to the discovery of N40K. We did not utilize this method for this study, and it is therefore impossible to determine whether sequenced peptides originated from WT U1-70K or the N40K transgene.

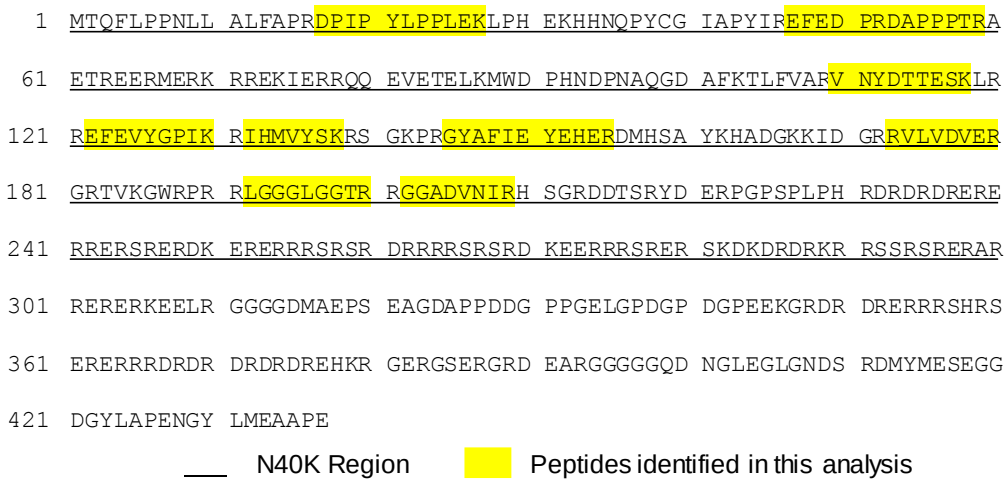


Figure 5.5. LC-MS/MS analysis detected 9 peptides from U1-70K and all originated from the N40K region.

U1-70K is a 437 aa protein, with N40K (underlined) encompassing the first 300 aa. C-terminal regions are often undetected by LC-MS/MS, and as such all peptides identified in this study originate from the N-terminal half – and N40K region.

Quantitative Analysis of the N40K Proteome

By quantification of the reporter ions of the TMT isobaric tags, we were able to determine relative abundances of each protein identified. Log_2 ratios were

calculated for each protein and histogram analysis of these values allowed a Gaussian curve to be fit to the data and sample variance determined in order to establish thresholds for statistical significance. First, null comparisons were examined for the duplicate samples from 6 mo old mice. Log_2 ratios were calculated between WT duplicates and Tg duplicates, which should be relatively unchanged from each other. We found that both comparisons were tightly clustered around a Log_2 value of zero, which indeed represents no change, and SDs were found to be 0.13 and 0.15 for WT and Tg, respectively (Fig 5.6). This is a very tight distribution, especially between biological replicates, as changes at the tissue and organism levels can produce large changes at the molecular level.

We next performed a number of test comparisons to examine changes in the Tg animals versus control. We compared protein changes overall by testing all Tg samples versus all WT samples, we tested 3 age-specific hippocampus comparisons at 3, 6, and 12 mos of age, changes in 12 mo cortex, and age-related changes by examining Tg to WT changes in 12 mo hippocampal versus 3 mo hippocampus (Fig 5.6). Each of these showed tight clustering around zero change, with the SD for overall changes at 0.145, 3 mos at 0.225, 6 mos at 0.150, 12 mos at 0.180, and age-related changes at 0.150. Taken collectively, while each exhibited rather low variance, comparisons with increased numbers of fractions – e.g. the overall and 6 mo comparisons - demonstrated relatively suppressed SD values, likely representing regression to the mean with increased *n*. Further, the fact that all comparisons closely resemble the null comparison and

technical variance depicts the relatively subtle effect of this transgene on the overall proteome.

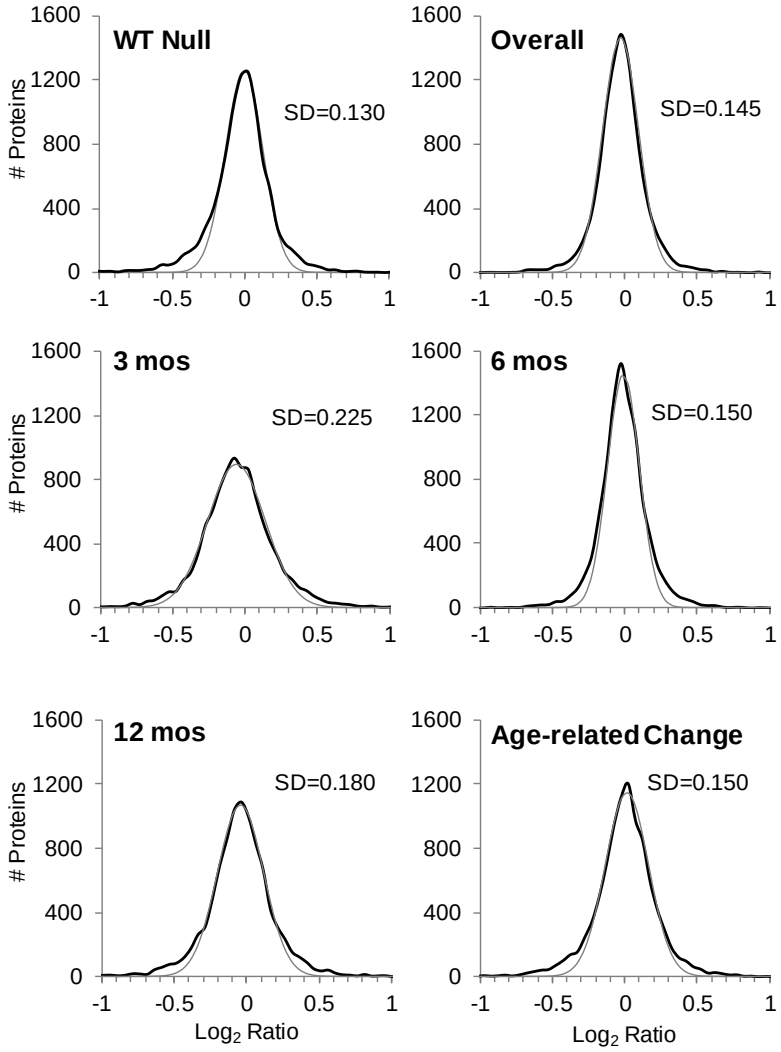


Figure 5.6. Overview of the comparisons in this study demonstrate most proteins cluster around zero change in abundance.

Histogram analysis and Gaussian fitting of the comparisons in our study allowed for determination of SD values to be used to establish thresholds for significant changes.

Using these standard deviations, we calculated z-scores and established threshold z-scores of -3 and 3 for significant change. This translates to threshold Log_2 ratios of ±0.435 for overall changes, ±0.675 for changes at 3 mos, ±0.45 at

6 mos, ±0.54 at 12 mos, ±0.715 in 12 mo cortex, and ±0.45 for age-related changes. Thus, in percentage change, significance thresholds were ±35.2% for overall changes, ±59.7% for changes at 3 mos, ±36.6% at 6 mos, ±45.4% at 12 mos, ±64.1% in 12 mo cortex, and 36.6% change at 12 mos over the change at 3 mos for age-related. Overall 185 proteins were significantly decreased and 137 increased 131 decreased and 135 increased at 3 mos saw, 142 decreased and 129 increased at 6 mos, 278 decreased and 180 increased at 12 mos saw, 197 decreased and 210 increased in 12 mos cortex, and 330 increased while 217 decreased in the age-related comparison.

Examination of Spliceosome Proteins and Others Altered in Multiple Comparisons

Our examination of proteins exhibiting significant rises and significant drops began at U1-70K, other subunits of the U1 SnRNP, and members of the spliceosome as a whole. We found that U1-70K levels are increased at roughly 2-fold over WT in the transgenic animals at all time points, in both hippocampus and cortex (Fig 5.7). This is not surprising as transcription of the N40K transgene is directed by a CaMKIIα promoter, which provides for very strong expression. What is more interesting, and perhaps surprising, is that abundance of all other U1-specific subunits are also increased to similar levels – all roughly 2 fold over WT, which is well beyond the significance threshold. Meanwhile, extending the analysis to members of other spliceosomal SnRNP complexes found no statistically significant increase or decrease in these proteins.

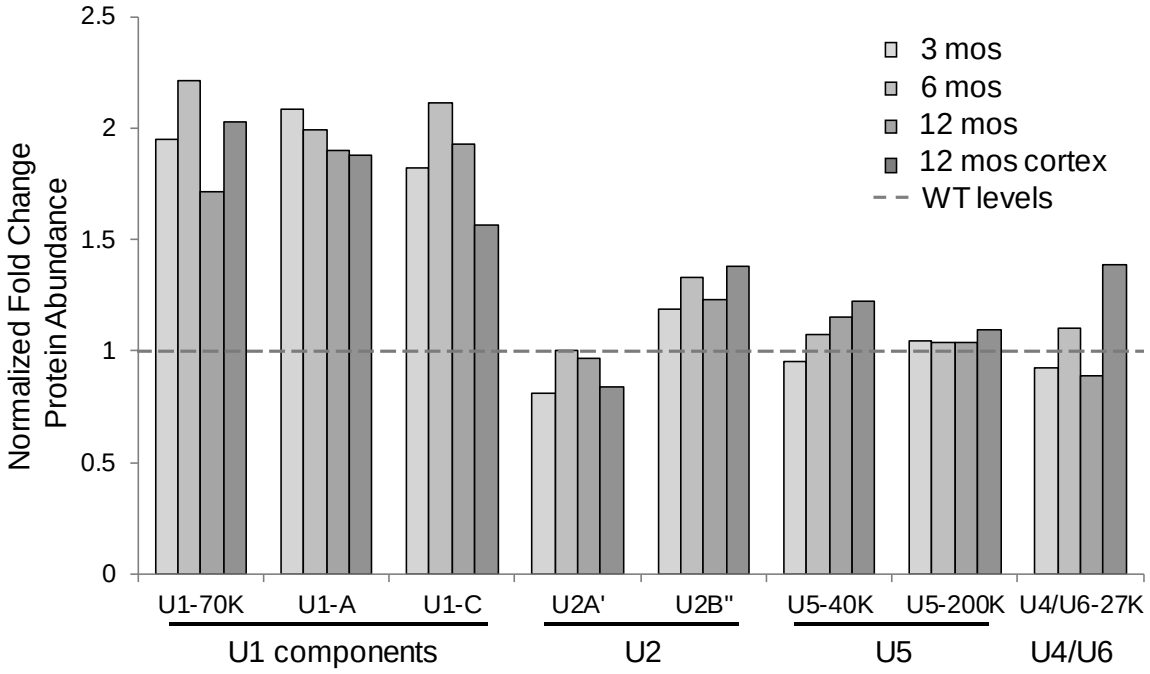


Figure 5.7. Proteins specific to U1, and no other spliceosome subunits exhibit roughly 2 fold increases in Tg over WT.

Components of the U1 snRNP were all enriched by 1.6 to 2.2 fold in all age-specific comparisons in this study. Components unique to other spliceosome snRNPs, however, demonstrated no significant alterations, up or down. Note these values are normalized raw values, as opposed to $\log_2$ ratios.

The core of the analysis rests in the most closely related – and therefore most directly comparable – fractions: 3 mo, 6 mo, and 12 mo hippocampus. In total, these fractions contained 492 proteins with significant changes, exhibiting significant, but certainly not complete, overlap between ages (Fig 5.7A). 19 proteins were significantly increased at all three ages, dominated by the three previously mentioned subunits of the U1 snRNP complex, as well as a fourth member of the spliceosome, Sm-D3, which is one of the subunits shared between the different “U” complexes. Also included in this subset is Nop2 and Thumpd1, both RNA-binding proteins, and Cplx3, a synaptic vesicle protein. 13

proteins were significantly decreased at all three ages, including Chordc1, a protein associated with Alzheimer's Disease pathways (190).

Further examination of perturbations in the N40K proteome can be accomplished by comparing all transgenic animals to all WT animals, as well as comparing difference at 12 mos of age to difference at 3 mos of age, thus discovering proteins which are particularly altered by expression of the N40K transgene, and proteins that change especially as age progresses, respectively (Fig 5.7B). There is, unsurprisingly, a fair amount of overlap between the proteins significantly altered overall and those altered in the age-specific comparisons, including the U1 subunits, as well as the mitochondrial protein Cox5b, and Foxg1 a widely studied protein important for brain development (191). Proteins displaying overall decrease included Synaptophysin, a widely studied synaptic marker that decreases in AD (192), CaMK1, a neuronal morphology regulator, Snx16, an endosomal protein, and GABRA2, a GABA receptor subunit. Proteins that are significantly altered in the age-related comparison exhibit either a positive or negative trend from 3 to 12 mos of age as compared to the WT trend. Significantly changed proteins in this comparison included Hist1h3b, a histone protein, and Rab3b, a protein involved in vesicular trafficking. These may represent age-related responses to the presence of the N40K species, or proteins altered as a result of long term cellular stress and are of particular interest as the behavioral differences in the mouse become more severe with age.

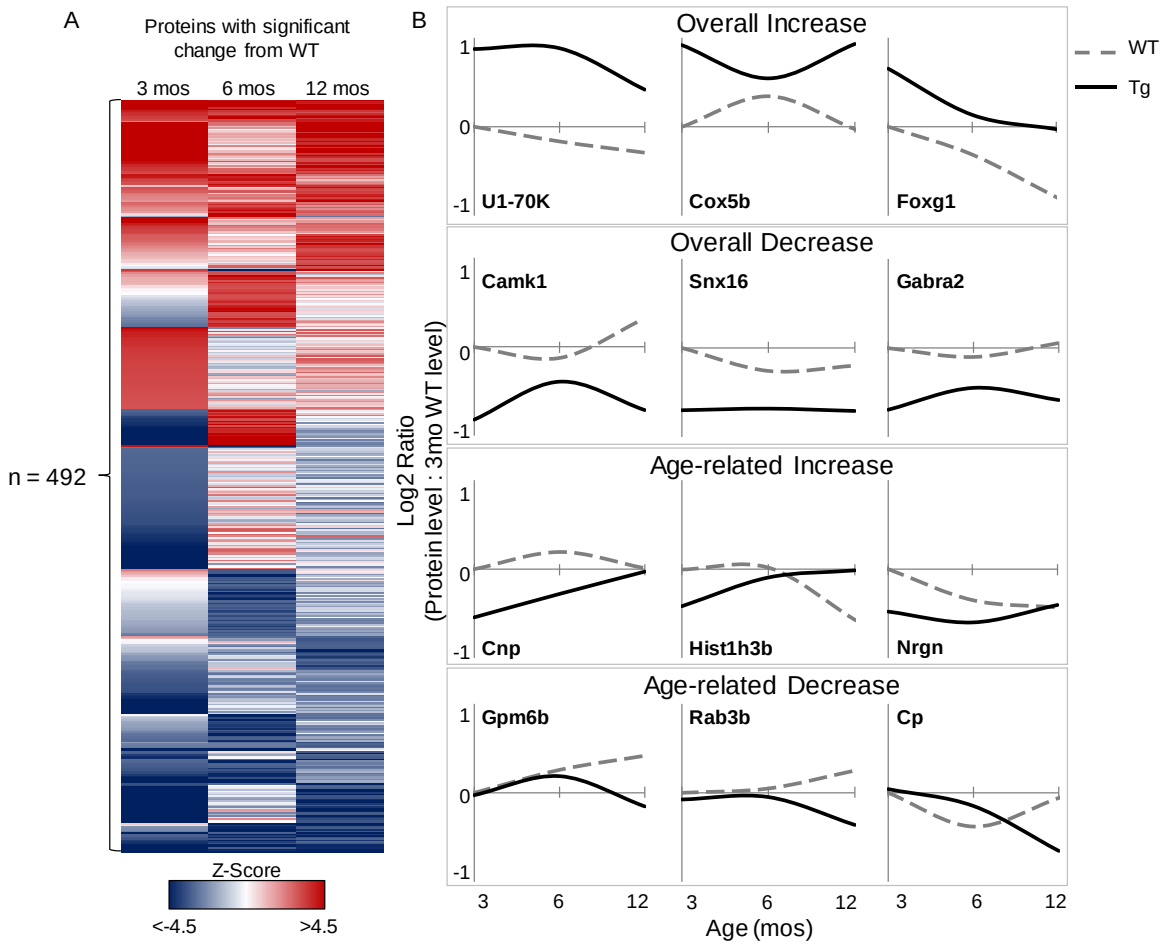


Figure 5.8. 492 proteins were significantly altered in our core age-specific analyses and many others altered in overall and age-related comparisons.

(A) Heatmap of the 492 proteins exhibiting at least one significant change (z-score > 3 or < -3) in the core age-specific hippocampal comparisons.

(B) Charts depicting protein abundance trajectories for select proteins that were enriched in the overall and age-related comparisons. All values depicted are $\log_2$ ratios calculated the raw protein abundance value for the given age and genotype over raw WT 3 mo values. This allows examination of changes at the 3 mo time point, as well as change over time.

Study of overlap between the two different brain regions at 12 mos provided a more rigorous survey of protein alterations, and potentially loosened some of the restraint of tissue specificity. Overall, 81 proteins overlapped with significant increases – out of 180 from hippocampus and 204 from cortex (Fig

5.8). These included several spliceosome proteins, as well as Nmt1, a protein involved in G-protein coupled receptor signaling, and Mfsd6, major histocompatibility complex receptor. 95 proteins were decreased in both 12 mo hippocampus and cortex, and include the aforementioned Synaptophysin and Slc12a2, a Sodium/Potassium/Chloride transporter.

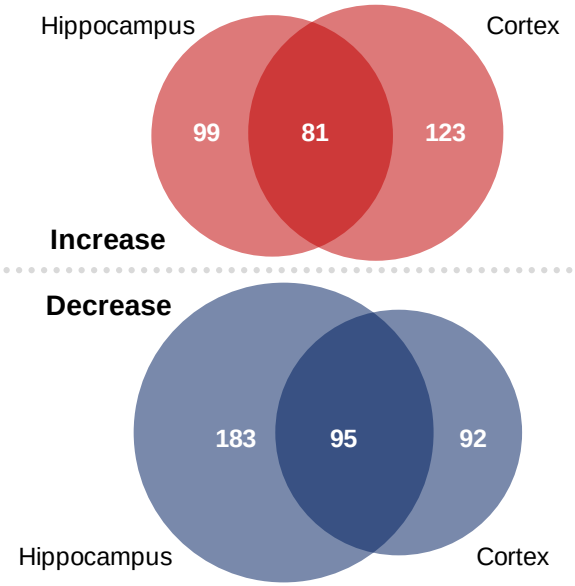


Figure 5.9. There is significant, but incomplete overlap of proteins altered in both hippocampus and cortex at 12 mos

Increases and decreases in these regions both show similar proportions of overlap, which may represent less tissue-specific functions.

Cytoskeleton, Mitochondria, Ciliogenesis, and Other Pathways Important to Neuronal Function Exhibit Perturbations in N40K Mice

To determine molecular pathways and gene ontologies enriched in each comparison, we utilized DAVID online functional analysis software (102). The algorithms utilized by this software employ a modified Fisher's exact test to determine the probability that pathways and ontologies found in the queried dataset at the observed rates versus the overall genome/proteome could be due

to random chance. It then clusters similar pathways and ontologies based on their gene members. We examined pathways enriched in all three age-specific comparisons, the overall comparison, and the age-related comparison, and they are summarized in Fig 5.9.

Taken altogether, the most prevalent enriched biological functions include the spliceosome, the cytoskeleton, mitochondria, and ciliogenesis with each enriched in at least 3 of the aforementioned comparisons. Spliceosome proteins, including the previously mentioned U1 subunits but also several Sm proteins and associated proteins such as Prpf40b, increased in all but age-related comparison. Mitochondrial proteins including Cox5b, Mthfd2, and Romo1 increased overall, at 3 mos, and 12 mos, Cytoskeletal proteins – and actin cytoskeletal proteins in particular – such as Gphn, Pdlim3 & 7, Arhgap6, and 3 members of the neurofilament family NEF H,M, and L decreased in an age-related fashion, at 6 mos, and 12 mos. Ciliogenesis proteins, many of which were also grouped as cytoskeletal proteins, included several members of the BBS family, Ift20, and Ttc8 decreased overall, and at 6 and 12 mos.

In comparison, there were very few enriched pathways and functions at 3 mos: only cell adhesion was decreased at this time point, and spliceosome, mitochondria, and ribosomal proteins were the few clear enriched pathways among increased proteins. Age-related alterations may at least provide a molecular parallel to the behavioral and organism-wide perturbations in the transgenic animal. Decreases in cytoskeletal proteins and perturbations in the complement cascade certainly point to age-related decline in neuronal

morphology and general health. While a *decrease* in complement cascade proteins over time is slightly puzzling in this context, it is possible they were already upregulated in the mutant at 3 mos, and therefore the 12 mo abundance was comparatively suppressed.

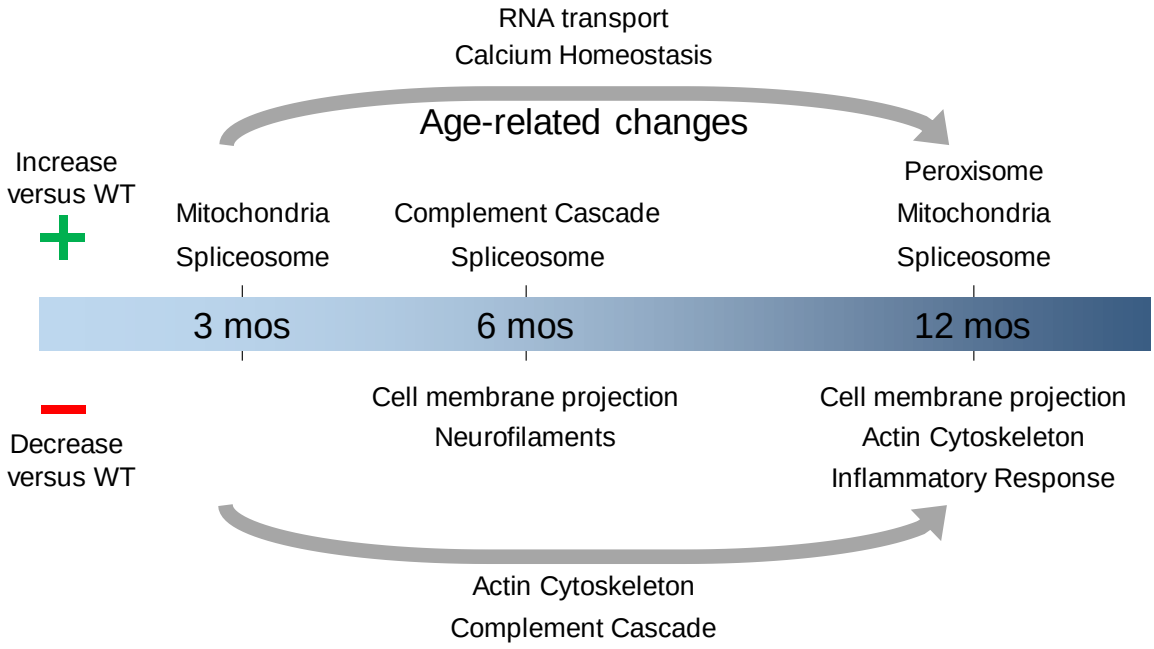


Figure 5.10. Several biological functions are enriched in comparisons from multiple time points and similar functions exhibit age-related changes.
Spliceosome, mitochondria, and calcium homeostasis pathways are all increased in transgenic animals, while cytoskeleton, ciliogenesis/membrane projection, and immune response proteins all decrease.

DISCUSSION

The proteomics study of the N40K transgenic mouse reported here is part of a larger, more comprehensive characterization of this novel mouse model. In our examination of this mouse, this proteome data through three ages and two brain regions adds to multiple behavior analyses, extensive immunohistochemical and biochemical examination, as well as additional

systems-based studies RNA-Seq and examination of the phosphoproteome – and all of which but the currently unfinished RNA-Seq have already produced promising and intriguing results.

Overall, we observed very low experimental variance, even comparisons across genotypes exhibited mean differences very close to $\log_2$ values of zero, and standard deviations of less than $\log_2$ 0.25. This points to two aspects of the analysis: first, technical variance was very low, a welcome result in any systems-based methodology; second, the perturbations of the transgene are relatively subtle and likely leave a majority of the proteome intact. Each comparison returned similar numbers of protein perturbations, further demonstrating little technical variance. The strong increase to all of the other subunits of the U1 snRNP, and several from other portions of the spliceosome were slightly unexpected but not entirely surprising, 19 proteins were increased in Tg animals at all ages measured, and 13 proteins decreased in all three. Increased proteins included Thumpd1, yet another RNA binding protein, and Complexin III, which is involved in presynaptic vesicle cycling and is perturbed in the 3xTG-AD mouse model (193). Meanwhile proteins decreased in all three included Slc12a2, a cation transporter that resides on the cell membrane, Gpcpd1, which is thought to contribute to visual cortex development (194), and Chordc1, a zinc and HSP90 binding protein (195) that is – albeit somewhat tangentially – involved in AD-related pathways(190). A particularly exciting finding was the strong decrease in Syp, or Synaptophysin, which was decreased in all comparisons besides the 6 mos comparison. This protein is a ubiquitous synaptic vesicle marker which is

connected to AD in several ways including decreased transcription in human AD tissue (192) and decreased protein levels in a human APP (hAPP) expressing mouse model (196).

As U1-70K undergoes extreme modification during human AD pathogenesis, and behavioral analysis closely resembles that of AD mouse models, we expected perturbations to AD-related proteins. While few directly AD-related proteins were detected, such as Synaptophysin and Chordc1, many of the altered protein pathways in this mouse are also altered in human AD cases. The strong perturbations of both the actin and microtubule cytoskeletal systems are primary examples of this is. It is no recent finding that neuronal cytoskeleton is disrupted in AD (197). Indeed, one of the key aberrant proteins in AD, tau is considered a key stabilizer of microtubule cytoskeletons. Recent examination of the human tau interactome discovered a strong enrichment of ribonucleoproteins and many proteins containing the RNA recognition motif (RRM), which is also contained in U1-70K/N40K (198). Further, perturbations to mitochondria (169) and the complement cascade have been widely demonstrated in AD patients, with complement cascade component C1R observed to confer risk in GWAS studies (68).

The pathway alterations we observed are also closely related to those seen in more traditional AD model mouse models – of which there are many. Previous studies of the many AD mouse models have returned varied aspects of AD pathology, with no one model fully recapitulating the human phenotype. Preeminent models involve human disease-related mutant forms of APP, Abeta,

presenilin, tau, and APOE alone or in combination and generally exhibit plaques and cognitive deficits late in life (199). Alterations in protein pathways beyond the amyloid pathway are seen in mitochondria, and synaptic function (200,201). A very recent paper described a smaller scale proteomic study examining APPswe/PS1dE9 transgenic mice found changes in cytoskeletal structure, energy metabolism, synaptic components, and protein degradation (202), while a more targeted proteomics analysis of hippocampal synaptosomes from the same mice at three months of age found significant upregulation of several protein components of the extracellular matrix (203). Each of these are mirrored in our results, suggesting strongly similar pathogenic events occur in the N40K transgenic.

In depth proteome analysis as part of the initial characterization of an original transgenic mouse is a powerful and somewhat unique luxury of our group. It provides an extremely thorough analysis of perturbations caused by introduction of the transgene, which is particularly useful in a new transgenic. While previous evidence and hypotheses can direct study toward the examination of meaningful and significant molecular alterations, the almost immeasurable complexity of an animal will inevitably lead to unexpected changes. An unbiased and system-wide analysis of a transgenic animal, especially when it has only recently been developed and is basically unstudied allows for the generation of hypotheses for future examination with greater accuracy. Expected changes may certainly be observed, and pursued further.

Meanwhile unexpected changes may arise and could possibly lead in entirely new directions.

CHAPTER SIX

General Discussion

A colleague of mine recently disparaged the “-omics revolution” by saying we are simply adding ‘-omics’ to everything and calling it a new field. My response is that this is exactly the way it should be. There are over 20,000 protein-coding genes in the human genome (204), evidence that over 8,000 are ubiquitously expressed (205), and varied ranges of over 10,000 expressed in the average cell in total – without even accounting for alternative splicing. It takes no major leap in imagination to think that even the most minute process involves vast arrays of interacting proteins, protein complexes, and even networks. We are constraining the advance of our understanding if we don’t harness the capabilities available to us. It is almost an affront to scientific endeavor – and the biology itself – to restrict our evaluation to only what more traditional targeted methods – and we ourselves can easily grasp. Analytical methods designed to identify and measure large numbers of analytes *en masse*, coupled with high-powered computer driven assays against comprehensive databases provide

information completely unattainable by more traditional methods. The spread of the –ome and –omics suffixes denotes an appropriate shift in our approach to understand the astounding complexity of biological systems.

Advances in technology have consistently driven advances in scientific discovery. The advent of new ionization methods such as electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI) protein mass spectrometry was a major advance in technology that extended the long-established analytical method of mass spectrometry to molecules the size of proteins and peptides, and greatly broadened our capability to study these major actors in the cell (206-208). Further major developments in the field have been fueled by advances in other technologies as well, such as genomics, computing, and chemistry. Complete genome sequences and the inferred amino acid sequences of every potential protein in a given organism brought protein mass spectrometry to a global scale and paved the way for true proteome analysis (209). Ever-advancing computational power, and more efficient algorithm design have made it possible to probe extensive proteomic databases. These advances simultaneously evaluate the accuracy of analysis while generally making experiments speedier, more thorough, and more cost effective. Creation of new isotope containing tags designed to label peptide samples with efficiency approaching 100%, fragment at specific covalent bonds during MS analysis, and distinguish intermixed and simultaneously run samples have greatly expanded the quantitative abilities of this methodology (11).

Interactomics

The examination of protein-protein interaction is a key field of research as the interfaces between proteins provide prime targets for small molecule intervention, and therefore disease treatment. The long-entrenched standard of assembling series of binary interactions has produced limited results, and incomplete understanding. The simple fact that individual protein copy numbers range up to 10^8 per cell (5), and that it is estimated that proteins have on average 5 interactions with other proteins (210) signals the need for high-throughput approaches in this arena. It has become increasingly clear that disease-causing mutations and environmental insults perturb interacting networks that involve compensatory mechanisms such as redundancy and feedback loops which take thorough understanding beyond simple binary interactions between a handful of proteins (6). Alzheimer's disease is a prime example: Aβ and Tau aggregations were discovered to be key hallmarks of the disease roughly 30 years ago, but no clinical treatment has gone beyond Phase III for this highly prevalent disease, and the central hypothesis around AD pathogenesis is being called into question (211).

The prevalence and biological significance of protein complexes alone suggests a need for more broad analyses. In yeast, at least 45% of the proteome enter into protein complexes, or the *complexome* (212), and recent incomplete literature-based estimates by the comprehensive resource of mammalian protein complexes (CORUM) dataset have shown that a growing value of 16% mammalian proteins are confirmed to participate in protein complexes (213). Deeper analysis of this data found mammalian protein complexes most

frequently consist of 3 to 4 different subunits and that proteins can be reused in up to 53 different protein complexes (214). Thus, many functional protein-protein interactions can take place through complexes with many of the members of these complexes never coming into direct physical contact with each other.

Even systems-based approaches must undergo significant refinement in order to fully understand interaction networks. For instance, many biologically relevant interactions are weak and last only seconds, making them very difficult to capture and examine (215). While a proteomics approach can improve sensitivity in this fashion by more thoroughly analyzing the system, and therefore improving the chance of recovering short-lived and weak interactions, improvements on study design around utilizing such strong analytical methods must be made to fully take advantage. It was a major goal of the projects reported in Chapters Two and Four to address this issue. Each used multiple levels of elution or wash stringency to capture weakly bound proteins but separate them from those more strongly bound. The effectiveness of these approaches is promising for the field moving forward.

Mounting Problems with Antibody-Based Assays

Events not directly related to the proteomics field are also propelling the advancement of proteomics methods such as the building frustration with antibody-based assays. Non-specific background and difficulty of discerning true target specificity are widely known to those who use antibodies for their research, but the problem may be worse than commonly thought. The Human Protein Atlas project recently examined more than 20,000 commercially available antibodies

and estimated that slightly more than half were not reliable for detection of their target molecule in tissue staining (the caveat being that many had not been supplier-approved for this application) (216). More disconcerting is the fact that within this study antibody reliability ranged greatly across suppliers, with many suppliers not reaching 50% success rate in either the mono- or polyclonal antibodies they offer. Further, variability between lots means even a perfectly acceptable antibody can lose its usefulness once a lot is used up. This lot-to-lot inconsistency can be difficult to estimate – and difficult to learn from suppliers – and can clearly distort results (158). Issues like these have driven a small but growing movement in the field to set strict standards for antibody production, including reducing them down to their genetic sequences and producing them in a uniform cell type (217). Even then, the bottom line remains that every time an antibody is purchased, it should be fairly extensively verified in the hands of the experimenters doing the work. Protein mass spectrometry circumvents many of the specificity issues of antibodies by simple high resolution and highly accurate mass measurements. The sensitivity, resolution, and large dynamic range of MS instruments, coupled with powerful and ever-improving computational evaluation makes this analysis method a true rival to detection and quantification via antibody.

General Conclusions and Looking Forward

Taken together, the methods and data presented herein depict small steps forward in understanding neurological disease. We connected Mib1, a widely powerful E3 ligase with many new interaction partners, and were able to partition

these both in relative binding strength and confidence of interaction. Further, we broadened the known functional role for Mib1 to include regulation of dendritic spine outgrowth, and through the degradation of another powerful regulator of neurodevelopment, CDKL5. We very nearly implemented and refined a innovative systems-based method for examining E3 ligase substrates. A method that still holds some promise given the right controls and manipulations, and the fact that the potential reward is great – as it could be applied to basically any E3 ligase. We broadened the understanding of AD pathogenesis through highly in-depth analysis of two critical proteins: APP and U1-70K. Our DEEP analysis of APP displayed the use of multiple dimensions of analysis in affinity-based interactome study, as overlapping results were much less abundant than expected. We expanded the interactome of the promiscuous APP to many proteins, and linked it to interaction networks including the cytoskeleton, mitochondria, ribosomes, and spliceosomes – of which U1-70K is a member. Our analysis of N40K transgenic mice resulted in several interesting overlaps with human AD and mouse AD models, including the synaptic marker Synaptophysin. Pathways perturbed in this animal relate strongly to neurodegenerative diseases, and support other preliminary results from this mouse well.

Looking forward longer term, the tremendous sensitivity of MS instruments and ever-growing computational power hold great promise. I expect them to provide two (among many) powerful developments in the proteomics and interactomics fields, on somewhat opposing ends of the scale from each other. First as sensitivity and instrumentation develops, I expect single cell proteomics

to hold a strong future. Online surgical MS instruments are already being used as a type of real-time biopsy, and adapting a system much like that to micromanipulators and micropipettes used for electrophysiology could allow sampling and analysis of the proteomes of individual cells. Analyzing proteomes this way would allow for comparisons between cells in different and directly observable physiological states, and would provide much cleaner results versus analysis of entire tissues or organs. On the other end of the spectrum extensive computational modeling of the physiological states of cells, tissues, organs, and organisms is another major direction already underway by biomarker researchers in the collection of -omics fields. I believe as -omics information (and information from other sources) builds and is collected in large interconnected online databases, as accessible computational power grows, and as we learn to harness the confluence of these things, powerful and reliable computer models will be established for healthy and disease states for the range from genomes to organisms. Essentially an extension of GWAS analysis in combination with other data types this ome-ome will combine into much more comprehensive phenotypes to examine and test against.

References

1. Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M. C., Tassani, S., Piva, F., Perez-Amodio, S., Strippoli, P., and Canaider, S. (2013) An estimation of the number of cells in the human body. *Annals of human biology* **40**, 463-471

2. Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., Jacob Filho, W., Lent, R., and Herculano-Houzel, S. (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of comparative neurology* **513**, 532-541

3. Murre, J. M., and Sturdy, D. P. (1995) The connectivity of the brain: multi-level quantitative analysis. *Biological cybernetics* **73**, 529-545

4. Uhlen, M., Fagerberg, L., Hallstrom, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Ponten, F. (2015) Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419

5. Wisniewski, J. R., Hein, M. Y., Cox, J., and Mann, M. (2014) A "proteomic ruler" for protein copy number and concentration estimation without spike-in standards. *Mol Cell Proteomics* **13**, 3497-3506

6. Rolland, T., Tasan, M., Charloteaux, B., Pevzner, S. J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., Kamburov, A., Ghiassian, S. D., Yang, X., Ghamsari, L., Balcha, D., Begg, B. E., Braun, P., Brehme, M., Broly, M. P., Carvunis, A. R., Convery-Zupan, D., Corominas, R., Coulombe-Huntington, J., Dann, E., Dreze, M., Dricot, A., Fan, C., Franzosa, E., Gebreab, F., Gutierrez, B. J., Hardy, M. F., Jin, M., Kang, S., Kiros, R., Lin, G. N., Luck, K., MacWilliams, A., Menche, J., Murray, R. R., Palagi, A., Poulin, M. M., Rambout, X., Rasla, J., Reichert, P., Romero, V., Ruyssinck, E., Sahalie, J. M., Scholz, A., Shah, A. A., Sharma, A., Shen, Y., Spirohn, K., Tam, S., Tejeda, A. O., Trigg, S. A., Twizere, J. C., Vega, K., Walsh, J., Cusick, M. E., Xia, Y., Barabasi, A. L., Iakoucheva, L. M., Aloy, P., De Las Rivas, J., Tavernier, J., Calderwood, M. A., Hill, D. E., Hao, T., Roth, F. P., and Vidal, M. (2014) A proteome-scale map of the human interactome network. *Cell* **159**, 1212-1226

7. Bornmann, L., and Mutz, R. (2015) Growth rates of modern science: A bibliometric analysis based on the number of publications and cited

references. *Journal of the Association for Information Science and Technology* **66**, 2215-2222

8. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, Y., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R. A., Muzny, D. M., Scherer, S. E., Bouck, J. B., Sodergren, E. J., Worley, K. C., Rives, C. M., Gorrell, J. H., Metzker, M. L., Naylor, S. L., Kucherlapati, R. S., Nelson, D. L., Weinstock, G. M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D. R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H. M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R. W., Federspiel, N. A., Abola, A. P., Proctor, M. J., Myers, R. M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D. R., Olson, M. V., Kaul, R., Raymond, C., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G. A., Athanasiou, M., Schultz, R., Roe, B. A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W. R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J. A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D. G., Burge, C. B., Cerutti, L., Chen, H. C., Church, D., Clamp, M., Copley, R. R., Doerks, T., Eddy, S. R., Eichler, E. E., Furey, T. S., Galagan, J., Gilbert, J. G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L. S., Jones, T. A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W. J., Kitts, P., Koonin, E. V., Korf, I., Kulp, D., Lancet, D., Lowe, T. M., McLysaght, A., Mikkelsen, T., Moran, J. V., Mulder, N., Pollara, V. J., Ponting, C. P., Schuler, G., Schultz, J., Slater, G., Smit, A. F., Stupka, E., Szustakowki, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y. I.,

Wolfe, K. H., Yang, S. P., Yeh, R. F., Collins, F., Guyer, M. S., Peterson, J., Felsenfeld, A., Wetterstrand, K. A., Patrinos, A., Morgan, M. J., de Jong, P., Catanese, J. J., Osoegawa, K., Shizuya, H., Choi, S., Chen, Y. J., Szustakowki, J., and International Human Genome Sequencing, C. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921

9. Wadman, M. (2013) Economic return from Human Genome Project grows. *Nature*

10. Lander, E. S. (2011) Initial impact of the sequencing of the human genome. *Nature* **470**, 187-197

11. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **17**, 994-999

12. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering* **11**, 49-79

13. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics. *Nature* **422**, 198-207

14. Chait, B. T. (2006) Chemistry. Mass spectrometry: bottom-up or top-down? *Science* **314**, 65-66

15. Liu, H. B., Sadygov, R. G., and Yates, J. R. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical Chemistry* **76**, 4193-4201

16. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., and Mann, M. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics* **4**, 1265-1272

17. Ong, S. E., and Mann, M. (2005) Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology* **1**, 252-262

18. Eberl, H. C., Spruijt, C. G., Kelstrup, C. D., Vermeulen, M., and Mann, M. (2013) A Map of General and Specialized Chromatin Readers in Mouse Tissues Generated by Label-free Interaction Proteomics. *Molecular Cell* **49**, 368-378

19. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* **1**, 376-386

20. Chen, X., Wei, S., Ji, Y., Guo, X., and Yang, F. (2015) Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **15**, 3175-3192

21. McClatchy, D. B., and Yates, J. R. (2014) Stable Isotope Labeling in Mammals (SILAM). in *Shotgun Proteomics: Methods and Protocols* (MartinsDeSouza, D. ed.), Humana Press Inc, Totowa. pp 133-146

22. Thompson, A., Schafer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K., and Hamon, C. (2003) Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**, 1895-1904

23. von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S., and Bork, P. (2002) Comparative assessment of large-scale data sets of protein-protein interactions. in *Nature*, England. pp 399-403

24. Fields, S., and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246

25. Hida, N., Awais, M., Takeuchi, M., Ueno, N., Tashiro, M., Takagi, C., Singh, T., Hayashi, M., Ohmiya, Y., and Ozawa, T. (2009) High-sensitivity real-time imaging of dual protein-protein interactions in living subjects using multicolor luciferases. *PLoS One* **4**, e5868

26. Jares-Erijman, E. A., and Jovin, T. M. (2003) FRET imaging. *Nat Biotechnol* **21**, 1387-1395

27. Sanchez, C., Lachaize, C., Janody, F., Bellon, B., Roder, L., Euzenat, J., Rechenmann, F., and Jacq, B. (1999) Grasping at molecular interactions and genetic networks in Drosophila melanogaster using FlyNets, an Internet database. *Nucleic Acids Res* **27**, 89-94

28. Bruckner, A., Polge, C., Lentze, N., Auerbach, D., and Schlattner, U. (2009) Yeast two-hybrid, a powerful tool for systems biology. *International Journal of Molecular Sciences* **10**, 2763-2788

29. Mann, M., Kulak, N. A., Nagaraj, N., and Cox, J. (2013) The Coming Age of Complete, Accurate, and Ubiquitous Proteomes. *Molecular Cell* **49**, 583-590

30. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C., and Yates, J. R., 3rd. (2013) Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* **113**, 2343-2394

31. Gstaiger, M. a. A. R. (2009) Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet* **10**, 617--627

32. Kaake, R. M. a. W. X. a. H. L. (2010) Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry. *Molecular \& cellular proteomics : MCP* **9**, 1650--1665

33. Nesvizhskii, A. I. (2012) Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments. *Proteomics* **12**, 1639--1655

34. Kaake, R. M., Wang, X. R., Burke, A., Yu, C., Kandur, W., Yang, Y. Y., Novtisky, E. J., Second, T., Duan, J. C., Kao, A., Guan, S. H., Vellucci, D., Rychnovsky, S. D., and Huang, L. (2014) A New in Vivo Cross-linking Mass Spectrometry Platform to Define Protein-Protein Interactions in Living Cells. *Molecular & Cellular Proteomics* **13**, 3533-3543

35. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nature Biotechnology* **17**, 1030-1032

36. Guerrero, C. a. T. C. a. K. P. a. H. L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Molecular \& cellular proteomics : MCP* **5**, 366--378

37. Keilhauer, E. C., Hein, M. Y., and Mann, M. (2015) Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). *Molecular & Cellular Proteomics* **14**, 120-135

38. Mann, M., and Jensen, O. N. (2003) Proteomic analysis of post-translational modifications. *Nature biotechnology* **21**, 255-261

39. Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **21**, 921-926

40. Tran, J. C., Zamdborg, L., Ahlf, D. R., Lee, J. E., Catherman, A. D., Durbin, K. R., Tipton, J. D., Vellaichamy, A., Kellie, J. F., and Li, M. (2011) Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* **480**, 254-258

41. Zhang, H., Li, X.-j., Martin, D. B., and Aebersold, R. (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nature biotechnology* **21**, 660-666

42. Kirkpatrick, D. S., Denison, C., and Gygi, S. P. (2005) Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nature Cell Biology* **7**, 750-757

43. Koo, B. K., Lim, H. S., Song, R., Yoon, M. J., Yoon, K. J., Moon, J. S., Kim, Y. W., Kwon, M. C., Yoo, K. W., Kong, M. P., Lee, J., Chitnis, A. B., Kim, C. H., and Kong, Y. Y. (2005) Mind bomb 1 is essential for generating functional Notch ligands to activate Notch. *Development* **132**, 3459-3470

44. Wang, H. B., Wang, L. J., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., and Zhang, Y. (2004) Role of histone H2A ubiquitination in polycomb silencing. *Nature* **431**, 873-878

45. Varshavsky, A. (2012) The Ubiquitin System, an Immense Realm. in *Annual Review of Biochemistry, Vol 81* (Kornberg, R. D. ed.), Annual Reviews, Palo Alto. pp 167-176

46. Sarraf, S. A., Raman, M., Guarani-Pereira, V., Sowa, M. E., Huttlin, E. L., Gygi, S. P., and Harper, J. W. (2013) Landscape of the PARKIN-dependent ubiquitylome in response to mitochondrial depolarization. *Nature* **496**, 372-376

47. Itoh, M., Kim, C.-H., Palardy, G., Oda, T., Jiang, Y.-J., Maust, D., Yeo, S.-Y., Lorick, K., Wright, G. J., Ariza-McNaughton, L., Weissman, A. M., Lewis, J., Chandrasekharappa, S. C., and Chitnis, A. B. (2003) Mind bomb is a ubiquitin ligase that is essential for efficient activation of Notch signaling by Delta. *Developmental cell* **4**, 67--82

48. Yoon, K. J., Lee, H. R., Jo, Y. S., An, K., Jung, S. Y., Jeong, M. W., Kwon, S. K., Kim, N. S., Jeong, H. W., Ahn, S. H., Kim, K. T., Lee, K., Kim, E., Kim, J. H., Choi, J. S., Kaang, B. K., and Kong, Y. Y. (2012) Mind bomb-1 is an essential modulator of long-term memory and synaptic plasticity via the Notch signaling pathway. *Mol Brain* **5**, 40

49. Choe, E.-A., Liao, L., Zhou, J.-Y., Cheng, D., Duong, D. M., Jin, P., Tsai, L.-H., and Peng, J. (2007) Neuronal morphogenesis is regulated by the interplay between cyclin-dependent kinase 5 and the ubiquitin ligase mind bomb 1. *The Journal of Neuroscience* **27**, 9503-9512

50. Tessier-Lavigne, M., and Goodman, C. S. (1996) The molecular biology of axon guidance. *Science* **274**, 1123-1133

51. Cline, H. T. (2001) Dendritic arbor development and synaptogenesis. *Current opinion in neurobiology* **11**, 118-126

52. Nimchinsky, E. A., Sabatini, B. L., and Svoboda, K. (2002) Structure and function of dendritic spines. *Annu Rev Physiol* **64**, 313-353

53. Scheiffele, P. (2003) Cell-cell signaling during synapse formation in the CNS. *Annual review of neuroscience* **26**, 485--508

54. Kwon, D. Y., Dimitriadi, M., Terzic, B., Cable, C., Hart, A. C., Chitnis, A., Fischbeck, K. H., and Burnett, B. G. (2013) The E3 ubiquitin ligase mind bomb 1 ubiquitinates and promotes the degradation of survival of motor neuron protein. *Mol Biol Cell* **24**, 1863-1871

55. Berndt, J. D., Aoyagi, A., Yang, P., Anastas, J. N., Tang, L., and Moon, R. T. (2011) Mindbomb 1, an E3 ubiquitin ligase, forms a complex with RYK to activate Wnt/β-catenin signaling. *The Journal of cell biology* **194**, 737-750

56. Murphy, S. L., Xu, J., and Kochanek, K. D. (2013) Deaths: final data for 2010. *National vital statistics reports : from the Centers for Disease Control and Prevention* **61**, 1-117

57. Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., and Ferri, C. P. (2013) The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimers & Dementia* **9**, 63-75

58. Thies, W., Bleiler, L., and Alzheimer's, A. (2013) 2013 Alzheimer's disease facts and figures Alzheimer's Association. *Alzheimers & Dementia* **9**, 208-245

59. Weuve, J., Hebert, L. E., Scherr, P. A., and Evans, D. A. (2014) Deaths in the United States among persons with Alzheimer's disease (2010-2050). *Alzheimers Dement* **10**, e40-46

60. Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Jr., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., Wagster, M. V., and Phelps, C. H. (2011) Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia* **7**, 280-292

61. Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., and Phelps, C. H. (2011) The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia* **7**, 270-279

62. McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C.,

Thies, B., Weintraub, S., and Phelps, C. H. (2011) The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers & Dementia* **7**, 263-269

63. Hugo, J., and Ganguli, M. (2014) Dementia and Cognitive Impairment Epidemiology, Diagnosis, and Treatment. *Clinics in Geriatric Medicine* **30**, 421-+

64. Gao, S., Hendrie, H. C., and Hall, K. S. (1998) The relationships between age, sex, and the incidence of dementia and Alzheimer disease - A meta-analysis. *Archives of General Psychiatry* **55**, 809-815

65. Braak, H., and Braak, E. (1991) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica* **82**, 239-259

66. Masliah, E., Mallory, M., Deerinck, T., Deteresa, R., Lamont, S., Miller, A., Terry, R. D., Carragher, B., and Ellisman, M. (1993) Re-evaluation of the structural organization of neuritic plaques in Alzheimer's disease. *Journal of Neuropathology and Experimental Neurology* **52**, 619-632

67. Brunetti, A., Postiglione, A., Tedeschi, E., Ciarmiello, A., Quarantelli, M., Covelli, E. M., Milan, G., Larobina, M., Soricelli, A., Sodano, A., and Alfano, B. (2000) Measurement of global brain atrophy in Alzheimer's disease with unsupervised segmentation of spin-echo MRI studies. *Journal of magnetic resonance imaging : JMRI* **11**, 260-266

68. Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., DeStafano, A. L., Bis, J. C., Beecham, G. W., Grenier-Boley, B., Russo, G., Thorton-Wells, T. A., Jones, N., Smith, A. V., Chouraki, V., Thomas, C., Ikram, M. A., Zelenika, D., Vardarajan, B. N., Kamatani, Y., Lin, C. F., Gerrish, A., Schmidt, H., Kunkle, B., Dunstan, M. L., Ruiz, A., Bihoreau, M. T., Choi, S. H., Reitz, C., Pasquier, F., Cruchaga, C., Craig, D., Amin, N., Berr, C., Lopez, O. L., De Jager, P. L., Deramecourt, V., Johnston, J. A., Evans, D., Lovestone, S., Letenneur, L., Moron, F. J., Rubinsztein, D. C., Eiriksdottir, G., Sleegers, K., Goate, A. M., Fievet, N., Huentelman, M. W., Gill, M., Brown, K., Kamboh, M. I., Keller, L., Barberger-Gateau, P., McGuiness, B., Larson, E. B., Green, R., Myers, A. J., Dufouil, C., Todd, S., Wallon, D., Love, S., Rogaeva, E., Gallacher, J., St George-Hyslop, P., Clarimon, J., Lleo, A., Bayer, A., Tsuang, D. W., Yu, L., Tsolaki, M., Bossu, P., Spalletta, G., Proitsi, P., Collinge, J., Sorbi, S., Sanchez-Garcia, F., Fox, N. C., Hardy, J., Deniz Naranjo, M. C., Bosco, P., Clarke, R., Brayne, C., Galimberti, D., Mancuso, M., Matthews, F., European Alzheimer's Disease, I., Genetic, Environmental Risk in Alzheimer's, D., Alzheimer's Disease Genetic, C., Cohorts for, H., Aging Research in Genomic, E., Moebus, S., Mecocci, P., Del Zompo, M., Maier, W., Hampel, H., Pilotto, A., Bullido, M., Panza, F.,

Caffarra, P., Nacmias, B., Gilbert, J. R., Mayhaus, M., Lannefelt, L., Hakonarson, H., Pichler, S., Carrasquillo, M. M., Ingelsson, M., Beekly, D., Alvarez, V., Zou, F., Valladares, O., Younkin, S. G., Coto, E., Hamilton-Nelson, K. L., Gu, W., Razquin, C., Pastor, P., Mateo, I., Owen, M. J., Faber, K. M., Jonsson, P. V., Combarros, O., O'Donovan, M. C., Cantwell, L. B., Soininen, H., Blacker, D., Mead, S., Mosley, T. H., Jr., Bennett, D. A., Harris, T. B., Fratiglioni, L., Holmes, C., de Bruijn, R. F., Passmore, P., Montine, T. J., Bettens, K., Rotter, J. I., Brice, A., Morgan, K., Foroud, T. M., Kukull, W. A., Hannequin, D., Powell, J. F., Nalls, M. A., Ritchie, K., Lunetta, K. L., Kauwe, J. S., Boerwinkle, E., Riemenschneider, M., Boada, M., Hiltuenen, M., Martin, E. R., Schmidt, R., Rujescu, D., Wang, L. S., Dartigues, J. F., Mayeux, R., Tzourio, C., Hofman, A., Nothen, M. M., Graff, C., Psaty, B. M., Jones, L., Haines, J. L., Holmans, P. A., Lathrop, M., Pericak-Vance, M. A., Launer, L. J., Farrer, L. A., van Duijn, C. M., Van Broeckhoven, C., Moskvina, V., Seshadri, S., Williams, J., Schellenberg, G. D., and Amouyel, P. (2013) Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet* **45**, 1452-1458

69. Kuhn, P. H., Wang, H., Dislich, B., Colombo, A., Zeitschel, U., Ellwart, J. W., Kremmer, E., Rossner, S., and Lichtenthaler, S. F. (2010) ADAM10 is the physiologically relevant, constitutive alpha-secretase of the amyloid precursor protein in primary neurons. *Embo Journal* **29**, 3020-3032

70. Biederer, T., Cao, X., Sudhof, T. C., and Liu, X. (2002) Regulation of APP-dependent transcription complexes by Mint/X11s: differential functions of Mint isoforms. in *J Neurosci*, United States. pp 7340-7351

71. Tarr, P. E., Roncarati, R., Pelicci, G., Pelicci, P. G., and D'Adamio, L. (2002) Tyrosine phosphorylation of the beta-amyloid precursor protein cytoplasmic tail promotes interaction with Shc. in *J Biol Chem*, United States. pp 16798-16804

72. Zhou, F., Gong, K., Song, B., Ma, T., van Laar, T., Gong, Y., and Zhang, L. (2012) The APP intracellular domain (AICD) inhibits Wnt signalling and promotes neurite outgrowth. in *Biochim Biophys Acta*, Crown 2012. Published by Elsevier B.V, Netherlands. pp 1233-1241

73. van Eeden, F. J., Granato, M., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C. P., Jiang, Y. J., Kane, D. A., Kelsh, R. N., Mullins, M. C., Odenthal, J., Warga, R. M., Allende, M. L., Weinberg, E. S., and Nusslein-Volhard, C. (1996) Mutations affecting somite formation and patterning in the zebrafish, Danio rerio. *Development* **123**, 153-164

74. Haddon, C., Jiang, Y. J., Smithers, L., and Lewis, J. (1998) Delta-Notch signalling and the patterning of sensory cell differentiation in the zebrafish ear: evidence from the mind bomb mutant. *Development* **125**, 4637-4644

75. Lawson, N. D., Scheer, N., Pham, V. N., Kim, C. H., Chitnis, A. B., Campos-Ortega, J. A., and Weinstein, B. M. (2001) Notch signaling is required for arterial-venous differentiation during embryonic vascular development. *Development* **128**, 3675-3683

76. Capoccia, B. J., Jin, R. U., Kong, Y. Y., Peek, R. M., Jr., Fassan, M., Rugge, M., and Mills, J. C. (2013) The ubiquitin ligase Mindbomb 1 coordinates gastrointestinal secretory cell maturation. *The Journal of clinical investigation* **123**, 1475-1491

77. Koo, B.-K., Yoon, M.-J., Yoon, K.-J., Im, S.-K., Kim, Y.-Y., Kim, C.-H., Suh, P.-G., Jan, Y. N., and Kong, Y.-Y. (2007) An obligatory role of mind bomb-1 in notch signaling of mammalian development. *PloS one* **2**

78. Liu, L. J., Liu, T. T., Ran, Y., Li, Y., Zhang, X. D., Shu, H. B., and Wang, Y. Y. (2012) The E3 ubiquitin ligase MIB1 negatively regulates basal IkappaBalpha level and modulates NF-kappaB activation. *Cell Res* **22**, 603-606

79. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402

80. Takeuchi, T., Heng, H. H., Ye, C. J., Liang, S. B., Iwata, J., Sonobe, H., and Ohtsuki, Y. (2003) Down-regulation of a novel actin-binding molecule, skeletrophin, in malignant melanoma. *Am J Pathol* **163**, 1395-1404

81. Tseng, L. C., Zhang, C., Cheng, C. M., Xu, H., Hsu, C. H., and Jiang, Y. J. (2014) New classes of mind bomb-interacting proteins identified from yeast two-hybrid screens. *PLoS One* **9**, e93394

82. Gstaiger, M., and Aebersold, R. (2009) Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nature Reviews Genetics* **10**, 617-627

83. Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17**, 1030-1032

84. Guerrero, C., Tagwerker, C., Kaiser, P., and Huang, L. (2006) An integrated mass spectrometry-based proteomic approach: quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26 S proteasome-interacting network. *Mol Cell Proteomics* **5**, 366-378

85. Kaake, R. M., Wang, X., and Huang, L. (2010) Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry. *Mol Cell Proteomics* **9**, 1650-1665

86. Dayon, L., Hainard, A., Licker, V., Turck, N., Kuhn, K., Hochstrasser, D. F., Burkhard, P. R., and Sanchez, J. C. (2008) Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal Chem* **80**, 2921-2931

87. Valli, E., Trazzi, S., Fuchs, C., Erriquez, D., Bartesaghi, R., Perini, G., and Ciani, E. (2012) CDKL5, a novel MYCN-repressed gene, blocks cell cycle and promotes differentiation of neuronal cells. *Biochim Biophys Acta* **1819**, 1173-1185

88. Xu, P., Duong, D. M., and Peng, J. (2009) Systematical Optimization of Reverse-Phase Chromatography for Shotgun Proteomics. *Journal of Proteome Research* **8**, 3944-3950

89. Wang, H., Yang, Y., Li, Y., Bai, B., Wang, X., Tan, H., Liu, T., Beach, T. G., Peng, J., and Wu, Z. (2015) Systematic optimization of long gradient chromatography mass spectrometry for deep analysis of brain proteome. *J Proteome Res* **14**, 829-838

90. Peng, J. M., Elias, J. E., Thoreen, C. C., Licklider, L. J., and Gygi, S. P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *Journal of Proteome Research* **2**, 43-50

91. Wang, X., Li, Y., Wu, Z., Wang, H., Tan, H., and Peng, J. (2014) JUMP: a tag-based database search tool for peptide identification with high sensitivity and accuracy. *Mol Cell Proteomics* **13**, 3663-3673

92. Wixon, J. a. K. D. (2000) The Kyoto Encyclopedia of Genes and Genomes—KEGG. *Yeast* **1**, 48-55

93. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57

94. Jensen, L. J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., and von Mering, C. (2009) STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic acids research* **37**, 412-416

95. Bai, B., Chen, P. C., Hales, C. M., Wu, Z. P., Pagala, V., High, A. A., Levey, A. I., Lah, J. J., and Peng, J. M. (2014) Integrated Approaches for Analyzing U1-70K Cleavage in Alzheimer's Disease. *Journal of Proteome Research* **13**, 4526-4534

96. Terabayashi, T., Funato, Y., Fukuda, M., and Miki, H. (2009) A coated vesicle-associated kinase of 104 kDa (CVAK104) induces lysosomal degradation of frizzled 5 (Fzd5). *J Biol Chem* **284**, 26716-26724

97. Brittain, J. M. a. W. Y. a. W. S. M. a. K. R. (2012) Regulation of CREB signaling through L-type Ca2+channels by Nipsnap-2. *Channels* **6**, 94--102

98. Han, K.-J., Foster, D. G., Zhang, N.-Y., Kanisha, K., Dzieciatkowska, M., Sclafani, R. A., Hansen, K. C., Peng, J., and Liu, C.-W. (2012) Ubiquitin-specific protease 9x deubiquitinates and stabilizes the spinal muscular atrophy protein-survival motor neuron. *The Journal of biological chemistry* **287**, 43741-43752

99. Homan, C. C., Kumar, R., Nguyen, L. S., Haan, E., Raymond, F. L., Abidi, F., Raynaud, M., Schwartz, C. E., Wood, S. A., Gecz, J., and Jolly, L. A. (2014) Mutations in USP9X are associated with X-linked intellectual disability and disrupt neuronal cell migration and growth. *Am J Hum Genet* **94**, 470-478

100. Li, Y.-f. a. C. Y.-f. a. H. Y. a. C. M. a. W. S. P. a. D. M. O. a. Z. H.-t. (2011) Phosphodiesterase-4D knockout and RNAi-mediated knockdown enhance memory and increase hippocampal neurogenesis via increased cAMP signaling. *Journal of Neuroscience* **31**, 172--183

101. Ethell, I. M., and Pasquale, E. B. (2005) Molecular mechanisms of dendritic spine development and remodeling. *Progress in neurobiology* **75**, 161-205

102. Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**, 1-13

103. Stegeman, S., Jolly, L. A., Premarathne, S., Gecz, J., Richards, L. J., Mackay-Sim, A., and Wood, S. A. (2013) Loss of Usp9x disrupts cortical architecture, hippocampal development and TGFbeta-mediated axonogenesis. *PLoS One* **8**, e68287

104. McCrea, P. D., and Gu, D. (2010) The catenin family at a glance. *J Cell Sci* **123**, 637-642

105. Wang, I. T., Allen, M., Goffin, D., Zhu, X., Fairless, A. H., Brodkin, E. S., Siegel, S. J., Marsh, E. D., Blendy, J. A., and Zhou, Z. (2012) Loss of CDKL5 disrupts kinome profile and event-related potentials leading to autistic-like phenotypes in mice. *Proc Natl Acad Sci U S A* **109**, 21516-21521

106. Tavyev Asher, Y. J., and Scaglia, F. (2012) Molecular bases and clinical spectrum of early infantile epileptic encephalopathies. *European journal of medical genetics* **55**, 299-306

107. Chen, Q., Zhu, Y.-C., Yu, J., Miao, S., Zheng, J., Xu, L., Zhou, Y., Li, D., Zhang, C., Tao, J., and Xiong, Z.-Q. (2010) CDKL5, a protein associated with rett syndrome, regulates neuronal morphogenesis via Rac1 signaling.

The Journal of neuroscience : the official journal of the Society for Neuroscience **30**, 12777-12786

108. Ricciardi, S., Ungaro, F., Hambrock, M., Rademacher, N., Stefanelli, G., Brambilla, D., Sessa, A., Magagnotti, C., Bachi, A., Giarda, E., Verpelli, C., Kilstrup-Nielsen, C., Sala, C., Kalscheuer, V. M., and Broccoli, V. (2012) CDKL5 ensures excitatory synapse stability by reinforcing NGL-1-PSD95 interaction in the postsynaptic compartment and is impaired in patient iPSC-derived neurons. *Nature cell biology* **14**, 911-923

109. Zhu, Y. C., Li, D., Wang, L., Lu, B., Zheng, J., Zhao, S. L., Zeng, R., and Xiong, Z. Q. (2013) Palmitoylation-dependent CDKL5-PSD-95 interaction regulates synaptic targeting of CDKL5 and dendritic spine development. *Proc Natl Acad Sci U S A* **110**, 9118-9123

110. Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A.-C., and Choi, H. (2014) SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *Journal of proteomics* **100**, 37-43

111. Rudiger, A. H., Rudiger, M., Carl, U. D., Chakraborty, T., Roepstorff, P., and Wehland, J. (1999) Affinity mass spectrometry-based approaches for the analysis of protein-protein interaction and complex mixtures of peptide-ligands. *Anal Biochem* **275**, 162-170

112. Righetti, P. G. (2014) The Monkey King: A personal view of the long journey towards a proteomic Nirvana. *Journal of proteomics* **107C**, 39--49

113. Song, R., Koo, B. K., Yoon, K. J., Yoon, M. J., Yoo, K. W., Kim, H. T., Oh, H. J., Kim, Y. Y., Han, J. K., Kim, C. H., and Kong, Y. Y. (2006) Neuralized-2 regulates a Notch ligand in cooperation with Mind bomb-1. *J Biol Chem* **281**, 36391-36400

114. Barsi, J. C., Rajendra, R., Wu, J. I., and Artzt, K. (2005) Mind bomb1 is a ubiquitin ligase essential for mouse embryonic development and Notch signaling. *Mech Dev* **122**, 1106-1117

115. Yoon, K. J., Koo, B. K., Im, S. K., Jeong, H. W., Ghim, J., Kwon, M. C., Moon, J. S., Miyata, T., and Kong, Y. Y. (2008) Mind bomb 1-expressing intermediate progenitors generate notch signaling to maintain radial glial cells. *Neuron* **58**, 519-531

116. Yamamoto, M., Morita, R., Mizoguchi, T., Matsuo, H., Isoda, M., Ishitani, T., Chitnis, A. B., Matsumoto, K., Crump, J. G., Hozumi, K., Yonemura, S., Kawakami, K., and Itoh, M. (2010) Mib-Jag1-Notch signalling regulates patterning and structural roles of the notochord by controlling cell-fate decisions. *Development* **137**, 2527-2537

117. Daskalaki, A., Shalaby, N. A., Kux, K., Tsoumpekos, G., Tsibidis, G. D., Muskavitch, M. A., and Delidakis, C. (2011) Distinct intracellular motifs of

Delta mediate its ubiquitylation and activation by Mindbomb1 and Neuralized. *J Cell Biol* **195**, 1017-1031

118. Dong, Z., Yang, N., Yeo, S.-Y., Chitnis, A., and Guo, S. (2012) Intralineage directional Notch signaling regulates self-renewal and differentiation of asymmetrically dividing radial glia. *Neuron* **74**, 65-78

119. Meloty-Kapella, L., Shergill, B., Kuon, J., Botvinick, E., and Weinmaster, G. (2012) Notch ligand endocytosis generates mechanical pulling force dependent on dynamin, epsins, and actin. *Dev Cell* **22**, 1299-1312

120. Blitzer, J. T., and Nusse, R. (2006) A critical role for endocytosis in Wnt signaling. *BMC Cell Biol* **7**, 28

121. Pak, D. T., Yang, S., Rudolph-Correia, S., Kim, E., and Sheng, M. (2001) Regulation of dendritic spine morphology by SPAR, a PSD-95-associated RapGAP. *Neuron* **31**, 289-303

122. Borgesius, N. Z., van Woerden, G. M., Buitendijk, G. H., Keijzer, N., Jaarsma, D., Hoogenraad, C. C., and Elgersma, Y. (2011) betaCaMKII plays a nonenzymatic role in hippocampal synaptic plasticity and learning by targeting alphaCaMKII to synapses. *J Neurosci* **31**, 10141-10148

123. S\ dhof, T. C. (2008) Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* **455**, 903--911

124. Fuchs, C., Trazzi, S., Torricella, R., Viggiano, R., De Franceschi, M., Amendola, E., Gross, C., Calza, L., Bartesaghi, R., and Ciani, E. (2014) Loss of CDKL5 impairs survival and dendritic growth of newborn neurons by altering AKT/GSK-3beta signaling. *Neurobiol Dis* **70**, 53-68

125. Artuso, R., Mencarelli, M. A., Polli, R., Sartori, S., Ariani, F., Pollazzon, M., Marozza, A., Cilio, M. R., Specchio, N., Vigevano, F., Vecchi, M., Boniver, C., Dalla Bernardina, B., Parmeggiani, A., Buoni, S., Hayek, G., Mari, F., Renieri, A., and Murgia, A. (2010) Early-onset seizure variant of Rett syndrome: definition of the clinical diagnostic criteria. *Brain & development* **32**, 17-24

126. Fehr, S., Wilson, M., Downs, J., Williams, S., Murgia, A., Sartori, S., Vecchi, M., Ho, G., Polli, R., Psoni, S., Bao, X., de Klerk, N., Leonard, H., and Christodoulou, J. (2013) The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *European journal of human genetics : EJHG* **21**, 266-273

127. Lecker, S. H., Goldberg, A. L., and Mitch, W. E. (2006) Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *Journal of the American Society of Nephrology* **17**, 1807-1819

128. Schwartz, A. L., and Ciechanover, A. (1999) The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annual Review of Medicine* **50**, 57-74

129. de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F., and Strouboulis, J. (2003) Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 7480-7485

130. Fernandez-Suarez, M., Chen, T. S., and Ting, A. Y. (2008) Protein-protein interaction detection in vitro and in cells by proximity biotinylation. *J Am Chem Soc* **130**, 9251-9253

131. Mertz, J., Tan, H., Pagala, V., Bai, B., Chen, P. C., Li, Y., Cho, J. H., Shaw, T., Wang, X., and Peng, J. (2015) Sequential Elution Interactome Analysis of the Mind Bomb 1 Ubiquitin Ligase Reveals a Novel Role in Dendritic Spine Outgrowth. *Mol Cell Proteomics* **14**, 1898-1910

132. Reyes-Turcu, F. E., Ventii, K. H., and Wilkinson, K. D. (2009) Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. in *Annual Review of Biochemistry*, Annual Reviews, Palo Alto. pp 363-397

133. Han, K. J., Foster, D. G., Zhang, N. Y., Kanisha, K., Dzieciatkowska, M., Sclafani, R. A., Hansen, K. C., Peng, J., and Liu, C. W. (2012) Ubiquitin-specific protease 9x deubiquitinates and stabilizes the spinal muscular atrophy protein-survival motor neuron. *J Biol Chem* **287**, 43741-43752

134. Finley, D. (2009) Recognition and Processing of Ubiquitin-Protein Conjugates by the Proteasome. in *Annual Review of Biochemistry*, Annual Reviews, Palo Alto. pp 477-513

135. Chen, I., Howarth, M., Lin, W. Y., and Ting, A. Y. (2005) Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nature Methods* **2**, 99-104

136. Thyagarajan, A., and Ting, A. Y. (2010) Imaging Activity-Dependent Regulation of Neurexin-Neuroligin Interactions Using trans-Synaptic Enzymatic Biotinylation. *Cell* **143**, 456-469

137. Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *Journal of Cell Biology* **196**, 801-810

138. Coyaud, E., Mis, M., Laurent, E. M. N., Dunham, W. H., Couzens, A. L., Robitaille, M., Gingras, A. C., Angers, S., and Raught, B. (2015) BioID-based Identification of Skp Cullin F-box (SCF)(beta-TrCP1/2) E3 Ligase Substrates. *Molecular & Cellular Proteomics* **14**, 1781-1795

139. Alzheimer, A. (1907) Über eine eigenartige Erkrankung der Hirnrinde [About a peculiar disease of the cerebral cortex]. *Allgemeine Zeitschrift für Psychiatrie und psychisch-gerichtliche Medizin [General Journal of Psychiatry and psycho-forensic Medicine]* **64**, 146-148

140. Glenner, G. G., and Wong, C. W. (1984) Alzheimers Disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and Biophysical Research Communications* **120**, 885-890

141. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. (1985) Amyloid plaque core protein in Alzheimer-Disease and Down Syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 4245-4249

142. Hardy, J. A., and Higgins, G. A. (1992) Alzheimer's disease: the amyloid cascade hypothesis. *Science* **256**, 184-185

143. Bettens, K., Sleegers, K., and Van Broeckhoven, C. (2013) Genetic insights in Alzheimer's disease. *Lancet Neurology* **12**, 92-104

144. Muller, U. C., and Zheng, H. (2012) Physiological Functions of APP Family Proteins. *Cold Spring Harbor Perspectives in Medicine* **2**, 17

145. Gervais, F. G., Xu, D., Robertson, G. S., Vaillancourt, J. P., Zhu, Y., Huang, J., LeBlanc, A., Smith, D., Rigby, M., Shearman, M. S., Clarke, E. E., Zheng, H., Van Der Ploeg, L. H., Ruffolo, S. C., Thornberry, N. A., Xanthoudakis, S., Zamboni, R. J., Roy, S., and Nicholson, D. W. (1999) Involvement of caspases in proteolytic cleavage of Alzheimer's amyloid-beta precursor protein and amyloidogenic A beta peptide formation. in *Cell*, United States. pp 395-406

146. Lu, D. C., Soriano, S., Bredesen, D. E., and Koo, E. H. (2003) Caspase cleavage of the amyloid precursor protein modulates amyloid beta-protein toxicity. *J Neurochem* **87**, 733-741

147. Guenette, S. Y., Chen, J., Ferland, A., Haass, C., Capell, A., and Tanzi, R. E. (1999) hFE65L influences amyloid precursor protein maturation and secretion. *J Neurochem* **73**, 985-993

148. Miller, C. C., McLoughlin, D. M., Lau, K. F., Tennant, M. E., and Rogelj, B. (2006) The X11 proteins, Abeta production and Alzheimer's disease. in *Trends Neurosci*, England. pp 280-285

149. Scheinfeld, M. H., Ghersi, E., Davies, P., and D'Adamio, L. (2003) Amyloid beta protein precursor is phosphorylated by JNK-1 independent of, yet facilitated by, JNK-interacting protein (JIP)-1. *J Biol Chem* **278**, 42058-42063

150. Muresan, Z., and Muresan, V. (2005) Coordinated transport of phosphorylated amyloid-beta precursor protein and c-Jun NH2-terminal kinase-interacting protein-1. in *J Cell Biol*, United States. pp 615-625

151. Cao, X., and Sudhof, T. C. (2001) A transcriptionally active complex of APP with Fe65 and histone acetyltransferase Tip60. in *Science*, United States. pp 115-120

152. von Rotz, R. C., Kohli, B. M., Bosset, J., Meier, M., Suzuki, T., Nitsch, R. M., and Konietzko, U. (2004) The APP intracellular domain forms nuclear multiprotein complexes and regulates the transcription of its own precursor. in *J Cell Sci*, England. pp 4435-4448

153. Briand, S., Facchinetti, P., Clamagirand, C., Madeira, A., Pommet, J. M., Pimplikar, S. W., and Allinquant, B. (2011) PAT1 induces cell death signal and SET mislocalization into the cytoplasm by increasing APP/APLP2 at the cell surface. in *Neurobiol Aging*, 2009 Elsevier Inc, United States. pp 1099-1113

154. Tsigelny, I. F., Crews, L., Desplats, P., Shaked, G. M., Sharikov, Y., Mizuno, H., Spencer, B., Rockenstein, E., Trejo, M., Platoshyn, O., Yuan, J. X., and Masliah, E. (2008) Mechanisms of hybrid oligomer formation in the pathogenesis of combined Alzheimer's and Parkinson's diseases. *PLoS One* **3**, e3135

155. Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T., Jr. (2002) Neurodegenerative disease: amyloid pores from pathogenic mutations. in *Nature*, England. pp 291

156. Olah, J., Vincze, O., Virok, D., Simon, D., Bozso, Z., Tokesi, N., Horvath, I., Hlavanda, E., Kovacs, J., Magyar, A., Szucs, M., Orosz, F., Penke, B., and Ovadi, J. (2011) Interactions of pathological hallmark proteins: tubulin polymerization promoting protein/p25, beta-amyloid, and alpha-synuclein. *J Biol Chem* **286**, 34088-34100

157. Cottrell, B. A., Galvan, V., Banwait, S., Gorostiza, O., Lombardo, C. R., Williams, T., Schilling, B., Peel, A., Gibson, B., Koo, E. H., Link, C. D., and Bredesen, D. E. (2005) A pilot proteomic study of amyloid precursor interactors in Alzheimer's disease. *Annals of Neurology* **58**, 277-289

158. Baker, M. (2015) Reproducibility crisis: Blame it on the antibodies. *Nature* **521**, 274-276

159. Varjosalo, M., Sacco, R., Stukalov, A., van Drogen, A., Planyavsky, M., Hauri, S., Aebersold, R., Bennett, K. L., Colinge, J., Gstaiger, M., and Superti-Furga, G. (2013) Interlaboratory reproducibility of large-scale human protein-complex analysis by standardized AP-MSMS. *Nature Methods* **10**, 307-+

160. Pagala, V. R., High, A. A., Wang, X., Tan, H., Kodali, K., Mishra, A., Kavdia, K., Xu, Y., Wu, Z., and Peng, J. (2015) Quantitative protein analysis by mass spectrometry. *Methods Mol Biol* **1278**, 281-305

161. Mi, H., Muruganujan, A., Casagrande, J. T., and Thomas, P. D. (2013) Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols* **8**, 1551-1566

162. Mi, H., Muruganujan, A., and Thomas, P. D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* **41**, D377-386

163. Elder, G. A., Sosa, M. A. G., and De Gasperi, R. (2010) Transgenic Mouse Models of Alzheimer's Disease. *Mount Sinai Journal of Medicine* **77**, 69-81

164. Hardy, J. (1997) Amyloid, the presenilins and Alzheimer's disease. *Trends in Neurosciences* **20**, 154-159

165. Gotz, J., and Ittner, L. M. (2008) Animal models of Alzheimer's disease and frontotemporal dementia. *Nature Reviews Neuroscience* **9**, 532-544

166. Prassas, I., and Diamandis, E. P. (2014) Translational researchers beware! Unreliable commercial immunoassays (ELISAs) can jeopardize your research. *Clinical chemistry and laboratory medicine : CCLM / FESCC* **52**, 765-766

167. Williams, R., Chung, J. Y., Ylaya, K., Whiteley, G., and Hewitt, S. M. (2010) Characterizations and validations of novel antibodies toward translational research. *Proteomics Clinical Applications* **4**, 618-625

168. Pianu, B., Lefort, R., Thuiliere, L., Tabourier, E., and Bartolini, F. (2014) The A beta(1-42) peptide regulates microtubule stability independently of tau. *Journal of Cell Science* **127**, 1117-1127

169. Lin, M. T., and Beal, M. F. (2006) Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. *Nature* **443**, 787-795

170. Del Prete, D., Lombino, F., Liu, X. R., and D'Adamio, L. (2014) APP Is Cleaved by Bace1 in Pre-Synaptic Vesicles and Establishes a Pre-Synaptic Interactome, via Its Intracellular Domain, with Molecular Complexes that Regulate Pre-Synaptic Vesicles Functions. *Plos One* **9**, 9

171. Konietzko, U., Goodger, Z. V., Meyer, M., Kohli, B. M., Bosset, J., Lahiri, D. K., and Nitsch, R. M. (2010) Co-localization of the amyloid precursor protein and Notch intracellular domains in nuclear transcription factories. *Neurobiology of Aging* **31**, 58-73

172. Agostinho, P., Cunha, R. A., and Oliveira, C. (2010) Neuroinflammation, Oxidative Stress and the Pathogenesis of Alzheimer's Disease. *Current Pharmaceutical Design* **16**, 2766-2778

173. Bai, B., Hales, C. M., Chen, P.-C., Gozal, Y., Dammer, E. B., Fritz, J. J., Wang, X., Xia, Q., Duong, D. M., Street, C., Cantero, G., Cheng, D., Jones, D. R., Wu, Z., Li, Y., Diner, I., Heilman, C. J., Rees, H. D., Wu, H., Lin, L., Szulwach, K. E., Gearing, M., Mufson, E. J., Bennett, D. A., Montine, T. J., Seyfried, N. T., Wingo, T. S., Sun, Y. E., Jin, P., Hanfelt, J., Willcock, D. M., Levey, A., Lah, J. J., and Peng, J. (2013) U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 16562-16567

174. Amador-Ortiz, C., Lin, W. L., Ahmed, Z., Personett, D., Davies, P., Dara, R., Graff-Radford, N. R., Hutton, M. L., and Dickson, D. W. (2007) TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. *Annals of Neurology* **61**, 435-445

175. Morel, E., Chamoun, Z., Lasiecka, Z. M., Chan, R. B., Williamson, R. L., Vetanovetz, C., Dall'Armi, C., Simoes, S., Du Jour, K. S. P., McCabe, B. D., Small, S. A., and Di Paolo, G. (2013) Phosphatidylinositol-3-phosphate regulates sorting and processing of amyloid precursor protein through the endosomal system. *Nature Communications* **4**, 13

176. Lee, V. M. Y., Balin, B. J., Otvos, L., and Trojanowski, J. Q. (1991) A68 - A MAJOR SUBUNIT OF PAIRED HELICAL FILAMENTS AND DERIVATIZED FORMS OF NORMAL-TAU. *Science* **251**, 675-678

177. Zhou, J., Jones, D. R., Duong, D. M., Levey, A. I., Lah, J. J., and Peng, J. (2013) Proteomic analysis of postsynaptic density in Alzheimer's Disease. *Clinica Chimica Acta* **420**, 62-68

178. Gozal, Y. M., Dammer, E. B., Duong, D. M., Cheng, D., Gearing, M., Rees, H. D., Peng, J., Lah, J. J., and Levey, A. I. (2011) Proteomic analysis of hippocampal dentate granule cells in frontotemporal lobar degeneration: application of laser capture technology. *Frontiers in neurology* **2**, 24-24

179. Gozal, Y. M., Seyfried, N. T., Gearing, M., Glass, J. D., Heilman, C. J., Wuu, J., Duong, D. M., Cheng, D., Xia, Q., Rees, H. D., Fritz, J. J., Cooper, D. S., Peng, J., Levey, A. I., and Lah, J. J. (2011) Aberrant septin 11 is associated with sporadic frontotemporal lobar degeneration. *Molecular Neurodegeneration* **6**

180. Gozal, Y. M., Duong, D. M., Gearing, M., Cheng, D., Hanfelt, J. J., Funderburk, C., Peng, J., Lah, J. J., and Levey, A. I. (2009) Proteomics Analysis Reveals Novel Components in the Detergent-Insoluble

Subproteome in Alzheimer's Disease. *Journal of Proteome Research* **8**, 5069-5079

181. Dammer, E. B., Na, C. H., Xu, P., Seyfried, N. T., Duong, D. M., Cheng, D., Gearing, M., Rees, H., Lah, J. J., Levey, A. I., Rush, J., and Peng, J. (2011) Polyubiquitin Linkage Profiles in Three Models of Proteolytic Stress Suggest the Etiology of Alzheimer Disease. *Journal of Biological Chemistry* **286**, 10457-10465

182. Liao, L., Cheng, D., Wang, J., Duong, D. M., Losik, T. G., Gearing, M., Rees, H. D., Lah, J. J., Levey, A. I., and Peng, J. (2004) Proteomic characterization of postmortem amyloid plaques isolated by laser capture microdissection. *Journal of Biological Chemistry* **279**, 37061-37068

183. Taylor, J. P., Hardy, J., and Fischbeck, K. H. (2002) Biomedicine - Toxic proteins in neurodegenerative disease. *Science* **296**, 1991-1995

184. Pomeranz Krummel, D. A., Oubridge, C., Leung, A. K., Li, J., and Nagai, K. (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature* **458**, 475-480

185. Buratti, E., and Baralle, D. (2010) Novel roles of U1 snRNP in alternative splicing regulation. *Rna Biology* **7**, 412-419

186. Diner, I., Hales, C. M., Bishof, I., Rabenold, L., Duong, D. M., Yi, H., Laur, O., Gearing, M., Troncoso, J., Thambisetty, M., Lah, J. J., Levey, A. I., and Seyfried, N. T. (2014) Aggregation properties of the small nuclear ribonucleoprotein U1-70K in Alzheimer disease. *J Biol Chem* **289**, 35296-35313

187. Seyfried, N. T., Gozal, Y. M., Dammer, E. B., Xia, Q., Duong, D. M., Cheng, D., Lah, J. J., Levey, A. I., and Peng, J. (2010) Multiplex SILAC Analysis of a Cellular TDP-43 Proteinopathy Model Reveals Protein Inclusions Associated with SUMOylation and Diverse Polyubiquitin Chains. *Molecular & Cellular Proteomics* **9**, 705-718

188. Twine, N., Janitz, K., Wilkins, M. R., and Janitz, M. (2015) Whole Transcriptome Sequencing Reveals Gene Expression and Splicing Differences in Brain Regions Affected by Alzheimer's Disease. *PloS one* **6**

189. Mills, J. D., Nalpathamkalam, T., Jacobs, H. I. L., Janitz, C., Merico, D., Hu, P. Z., and Janitz, M. (2013) RNA-Seq analysis of the parietal cortex in Alzheimer's disease reveals alternatively spliced isoforms related to lipid metabolism. *Neuroscience Letters* **536**, 90-95

190. Zhao, F., Xu, G., Zhou, Y., Wang, L., Xie, J., Ren, S., Liu, S., and Zhu, Y. (2014) MicroRNA-26b inhibits hepatitis B virus transcription and replication by targeting the host factor CHORDC1 protein. *J Biol Chem* **289**, 35029-35041

191. Mencarelli, M. A., Spanhol-Rosseto, A., Artuso, R., Rondinella, D., De Filippis, R., Bahi-Buisson, N., Nectoux, J., Rubinsztajn, R., Bienvenu, T., Moncla, A., Chabrol, B., Villard, L., Krumina, Z., Armstrong, J., Roche, A., Pineda, M., Gak, E., Mari, F., Ariani, F., and Renieri, A. (2010) Novel FOXG1 mutations associated with the congenital variant of Rett syndrome. *Journal of medical genetics* **47**, 49-53

192. Heffernan, J. M., Eastwood, S. L., Nagy, Z., Sanders, M. W., McDonald, B., and Harrison, P. J. (1998) Temporal cortex synaptophysin mRNA is reduced in Alzheimer's disease and is negatively correlated with the severity of dementia. *Experimental neurology* **150**, 235-239

193. Gatta, V., D'Aurora, M., Granzotto, A., Stuppia, L., and Sensi, S. L. (2014) Early and sustained altered expression of aging-related genes in young 3xTg-AD mice. *Cell Death Dis* **5**, e1054

194. Bakken, T. E., Roddey, J. C., Djurovic, S., Akshoomoff, N., Amaral, D. G., Bloss, C. S., Casey, B. J., Chang, L., Ernst, T. M., Gruen, J. R., Jernigan, T. L., Kaufmann, W. E., Kenet, T., Kennedy, D. N., Kuperman, J. M., Murray, S. S., Sowell, E. R., Rimol, L. M., Mattingsdal, M., Melle, I., Agartz, I., Andreassen, O. A., Schork, N. J., Dale, A. M., Alzheimer's Disease Neuroimaging, I., Pediatric Imaging, N., Genetics, S., Weiner, M., Aisen, P., Petersen, R., Jack, C. R., Jr., Jagust, W., Trojanowki, J. Q., Toga, A. W., Beckett, L., Green, R. C., Saykin, A. J., Morris, J., Liu, E., Montine, T., Gamst, A., Thomas, R. G., Donohue, M., Walter, S., Gessert, D., Sather, T., Harvey, D., Kornak, J., Dale, A., Bernstein, M., Felmlee, J., Fox, N., Thompson, P., Schuff, N., Alexander, G., DeCarli, C., Bandy, D., Koeppe, R. A., Foster, N., Reiman, E. M., Chen, K., Mathis, C., Cairns, N. J., Taylor-Reinwald, L., Trojanowki, J. Q., Shaw, L., Lee, V. M., Korecka, M., Crawford, K., Neu, S., Foroud, T. M., Potkin, S., Shen, L., Kachaturian, Z., Frank, R., Snyder, P. J., Molchan, S., Kaye, J., Quinn, J., Lind, B., Dolen, S., Schneider, L. S., Pawluczyk, S., Spann, B. M., Brewer, J., Vanderswag, H., Heidebrink, J. L., Lord, J. L., Johnson, K., Doody, R. S., Villanueva-Meyer, J., Chowdhury, M., Stern, Y., Honig, L. S., Bell, K. L., Morris, J. C., Ances, B., Carroll, M., Leon, S., Mintun, M. A., Schneider, S., Marson, D., Griffith, R., Clark, D., Grossman, H., Mitsis, E., Romirowsky, A., deToledo-Morrell, L., Shah, R. C., Duara, R., Varon, D., Roberts, P., Albert, M., Onyike, C., Kielb, S., Rusinek, H., de Leon, M. J., Glodzik, L., De Santi, S., Doraiswamy, P. M., Petrella, J. R., Coleman, R. E., Arnold, S. E., Karlawish, J. H., Wolk, D., Smith, C. D., Jicha, G., Hardy, P., Lopez, O. L., Oakley, M., Simpson, D. M., Porsteinsson, A. P., Goldstein, B. S., Martin, K., Makino, K. M., Ismail, M. S., Brand, C., Mulnard, R. A., Thai, G., Mc-Adams-Ortiz, C., Womack, K., Mathews, D., Quiceno, M., Diaz-Arrastia, R., King, R., Weiner, M., Martin-Cook, K., DeVous, M., Levey, A. I., Lah, J. J., Cellar, J. S., Burns, J. M., Anderson, H. S., Swerdlow, R. H., Apostolova, L., Lu, P. H., Bartzokis, G., Silverman, D. H., Graff-Radford, N. R., Parfitt, F., Johnson, H., Farlow, M. R., Hake, A. M., Matthews, B. R.,

Herring, S., van Dyck, C. H., Carson, R. E., MacAvoy, M. G., Chertkow, H., Bergman, H., Hosein, C., Black, S., Stefanovic, B., Caldwell, C., Ging, Y., Hsiung, R., Feldman, H., Mudge, B., Assaly, M., Kertesz, A., Rogers, J., Trost, D., Bernick, C., Munic, D., Kerwin, D., Mesulam, M. M., Lipowski, K., Wu, C. K., Johnson, N., Sadowsky, C., Martinez, W., Villena, T., Turner, R. S., Johnson, K., Reynolds, B., Sperling, R. A., Johnson, K. A., Marshall, G., Frey, M., Yesavage, J., Taylor, J. L., Lane, B., Rosen, A., Tinklenberg, J., Sabbagh, M., Belden, C., Jacobson, S., Kowall, N., Killiany, R., Budson, A. E., Norbash, A., Johnson, P. L., Obisesan, T. O., Wolday, S., Bwayo, S. K., Lerner, A., Hudson, L., Ogrocki, P., Fletcher, E., Carmichael, O., Olichney, J., Kittur, S., Borrie, M., Lee, T. Y., Bartha, R., Johnson, S., Asthana, S., Carlsson, C. M., Potkin, S. G., Preda, A., Nguyen, D., Tariot, P., Fleisher, A., Reeder, S., Bates, V., Capote, H., Rainka, M., Scharre, D. W., Kataki, M., Zimmerman, E. A., Celmins, D., Brown, A. D., Pearlson, G. D., Blank, K., Anderson, K., Santulli, R. B., Schwartz, E. S., Sink, K. M., Williamson, J. D., Garg, P., Watkins, F., Ott, B. R., Querfurth, H., Tremont, G., Salloway, S., Malloy, P., Correia, S., Rosen, H. J., Miller, B. L., Mintzer, J., Longmire, C. F., Spicer, K., Finger, E., Rachinsky, I., Drost, D., Jernigan, T., McCabe, C., Grant, E., Ernst, T., Kuperman, J., Chung, Y., Murray, S., Bloss, C., Darst, B., Pritchett, L., Saito, A., Amaral, D., DiNino, M., Eyngorina, B., Sowell, E., Houston, S., Soderberg, L., Kaufmann, W., van Zijl, P., Rizzo-Busack, H., Javid, M., Mehta, N., Ruberry, E., Powers, A., Rosen, B., Gebhard, N., Manigan, H., Frazier, J., Kennedy, D., Yakutis, L., Hill, M., Gruen, J., Bosson-Heenan, J., and Carlson, H. (2012) Association of common genetic variants in GPCPD1 with scaling of visual cortical surface area in humans. *Proc Natl Acad Sci U S A* **109**, 3985-3990

195. Gano, J. J., and Simon, J. A. (2010) A proteomic investigation of ligand-dependent HSP90 complexes reveals CHORDC1 as a novel ADP-dependent HSP90-interacting protein. *Mol Cell Proteomics* **9**, 255-270

196. Mucke, L., Masliah, E., Yu, G. Q., Mallory, M., Rockenstein, E. M., Tatsuno, G., Hu, K., Kholodenko, D., Johnson-Wood, K., and McConlogue, L. (2000) High-level neuronal expression of A beta(1-42) in wild-type human amyloid protein precursor transgenic mice: Synaptotoxicity without plaque formation. *Journal of Neuroscience* **20**, 4050-4058

197. Lee, V. M., and Trojanowski, J. Q. (1992) The disordered neuronal cytoskeleton in Alzheimer's disease. *Curr Opin Neurobiol* **2**, 653-656

198. Gunawardana, C. G., Mehrabian, M., Wang, X., Mueller, I., Lubambo, I. B., Jonkman, J. E., Wang, H., and Schmitt-Ulms, G. (2015) The Human Tau Interactome: Binding to the Ribonucleoproteome, and Impaired Binding of the Proline-to-Leucine Mutant at Position 301 (P301L) to Chaperones and the Proteasome. *Mol Cell Proteomics* **14**, 3000-3014

199. Hall, A. M., and Roberson, E. D. (2012) Mouse models of Alzheimer's disease. *Brain Res Bull* **88**, 3-12

200. Chou, J. L., Shenoy, D. V., Thomas, N., Choudhary, P. K., Laferla, F. M., Goodman, S. R., and Breen, G. A. (2011) Early dysregulation of the mitochondrial proteome in a mouse model of Alzheimer's disease. *J Proteomics* **74**, 466-479

201. Hong, I., Kang, T., Yoo, Y., Park, R., Lee, J., Lee, S., Kim, J., Song, B., Kim, S. Y., Moon, M., Yun, K. N., Kim, J. Y., Mook-Jung, I., Park, Y. M., and Choi, S. (2013) Quantitative proteomic analysis of the hippocampus in the 5XFAD mouse model at early stages of Alzheimer's disease pathology. *Journal of Alzheimer's disease : JAD* **36**, 321-334

202. Fu, Y., Zhao, D., Pan, B., Wang, J., Cui, Y., Shi, F., Wang, C., Yin, X., Zhou, X., and Yang, L. (2015) Proteomic Analysis of Protein Expression Throughout Disease Progression in a Mouse Model of Alzheimer's Disease. *Journal of Alzheimer's disease : JAD* **47**, 915-926

203. Vegh, M. J., Heldring, C. M., Kamphuis, W., Hijazi, S., Timmerman, A. J., Li, K. W., van Nierop, P., Mansvelder, H. D., Hol, E. M., Smit, A. B., and van Kesteren, R. E. (2014) Reducing hippocampal extracellular matrix reverses early memory deficits in a mouse model of Alzheimer's disease. *Acta neuropathologica communications* **2**, 76

204. Consortium, E. P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74

205. Ramsköld, D., Wang, E. T., Burge, C. B., and Sandberg, R. (2009) An Abundance of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data. in *PLoS Comput Biol*. pp

206. Nobelprize.org. (2002) Press Release: The Nobel Prize in Chemistry 2002. Nobel Media AB 2014.

207. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B. (1985) Electrospray interface for liquid chromatographs and mass spectrometers. *Analytical Chemistry* **57**, 675-679

208. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up tom/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **2**, 151-153

209. Eng, J. K., McCormack, A. L., and Yates, J. R. (1994) An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5**, 976-989

210. Grigoriev, A. (2003) On the number of protein-protein interactions in the yeast proteome. *Nucleic Acids Res* **31**, 4157-4161

211. Karran, E., Mercken, M., and De Strooper, B. (2011) The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature Reviews Drug Discovery* **10**, 698-U1600

212. Guldener, U., Munsterkotter, M., Kastenmuller, G., Strack, N., van Helden, J., Lemer, C., Richelles, J., Wodak, S. J., Garcia-Martinez, J., Perez-Ortin, J. E., Michael, H., Kaps, A., Talla, E., Dujon, B., Andre, B., Souciet, J. L., De Montigny, J., Bon, E., Gaillardin, C., and Mewes, H. W. (2005) CYGD: the Comprehensive Yeast Genome Database. *Nucleic Acids Res* **33**, D364-368

213. Ruepp, A., Waegele, B., Lechner, M., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Mewes, H. W. (2010) CORUM: the comprehensive resource of mammalian protein complexes--2009. *Nucleic Acids Res* **38**, D497-501

214. Ruepp, A., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Stransky, M., Waegele, B., Schmidt, T., Doudieu, O. N., Stumpflen, V., and Mewes, H. W. (2008) CORUM: the comprehensive resource of mammalian protein complexes. *Nucleic Acids Res* **36**, D646-650

215. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G., and Orengo, C. (2010) Transient protein-protein interactions: structural, functional, and network properties. *Structure* **18**, 1233-1243

216. Berglund, L., Bjorling, E., Oksvold, P., Fagerberg, L., Asplund, A., Szigyarto, C. A., Persson, A., Ottosson, J., Wernerus, H., Nilsson, P., Lundberg, E., Sivertsson, A., Navani, S., Wester, K., Kampf, C., Hober, S., Ponten, F., and Uhlen, M. (2008) A genecentric Human Protein Atlas for expression profiles based on antibodies. *Mol Cell Proteomics* **7**, 2019-2027

217. Bradbury, A., and Pluckthun, A. (2015) Standardize antibodies used in research. *Nature* **518**, 27-29