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: Date

_____ _____

Isaac Bishof July 13, 2018

RNA-binding proteins with mixed charge domains self-assemble, aggregate, and interact with core pathologies in Alzheimer's Disease

By

Isaac James Bishof

B.S., Florida State University, 2011

Advisor: Nicholas T. Seyfried, 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 Graduate Division of Biological and Biomedical Science
Biochemistry, Cell, and Developmental Biology
2018

ABSTRACT

RNA-binding proteins with mixed charge domains self-assemble, aggregate, and interact with core pathologies in Alzheimer's Disease

By Isaac James Bishof

U1-70K and other RNA binding proteins (RBPs) are mislocalized to cytoplasmic neurofibrillary Tau aggregates in Alzheimer's disease (AD), yet the mechanisms that cause their aggregation are incompletely understood. Many RBPs that aggregate in neurodegenerative diseases self-assemble into RNA granules through intrinsically disordered low complexity (LC) domains. We report here that a LC domain within U1-70K of mixed charge, containing highly repetitive complementary repeats of basic (R/K) and acidic (D/E) residues, shares many of the same properties of the Q/N-rich LC domains found in the RBPs TDP-43 and FUS. These properties include the ability to self-assemble into oligomers, and to form nuclear granules. To analyze the functional roles of the U1-70K LC domains, we performed co-immunoprecipitation and quantitative mass spectrometry analysis of recombinant U1-70K and deletions lacking the C-terminal LC domain(s). A network-driven approach resolved functional classes of U1-70K interacting proteins that showed dependency on the U1-70K LC domain(s) for their interaction. This included structurally similar RBPs, such as LUC7L3 and RBM25, which require their respective mixed charge domains for reciprocal interactions with U1-70K and for participation in nuclear RNA granules. Strikingly, a significant proportion of RBPs with mixed charge domains have elevated insolubility in the AD brain proteome compared to controls. Furthermore, we show that the mixed charge LC domain of U1-70K can interact with β-amyloid and Tau from AD brain. This supports a hypothesis that β-amyloid directly or indirectly mediates interactions between mixed charge structural motifs on U1-70K and related RBPs with pathological Tau in AD.

RNA-binding proteins with mixed charge domains self-assemble, aggregate, and interact with core pathologies in Alzheimer's Disease

By

Isaac James Bishof
B.S., Florida State University, 2011

Advisor: Nicholas T. Seyfried, 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 Biochemistry
2018

Table of Contents

List of Figure:

ABBREVATIONS

AD- Alzheimer's Disease

Aβ- Amyloid Beta

NFT- Neurofibrillary tangles

APP- Amyloid precursor protein

PET- positron emission tomography

FAD- Familial cases of Alzheimer's Disease

PS1- Presenilin 1

PS2- Presenilin 2

LOAD- late-onset AD

APOE- Apolipoprotein E

FTLD- Frontotemporal lobar degeneration

AsymAD- asymptomatic AD

MCI- mild cognitive impairment

PTM- Post-translational modification

MS- mass spectrometry

GWAS- genome wide association studies

CBD- Corticobasal degeneration

U1-70K- U1 small nuclear ribonucleoprotein 70 kDa

U1snRNP- U1 small nuclear ribonucleoprotein nuclear particle

ALS- amyotrophic lateral sclerosis

RBP- RNA-binding proteins

LLPS- Liquid–liquid phase separation

LC1- U1-70K low complexity domain one

LC2- U1-70K low complexity domain two

BN-PAGE- blue native gel polyacrylamide gel electrophoresis

LC-MS/MS- liquid chromatography coupled to tandem mass spectrometry

IgG- immunoglobulin

IP- immunoprecipitation

LFQ- label-free quantification

WGCNA- WeiGhted Co-expression Network Analysis

PPI- protein-protein interactions

GO- gene ontologies

T-SNE- T-Distributed Stochastic Neighbor Embedding

RNP- ribonuclear particle

FRAP- fluorescence recovery after photobleaching

CHAPTER 1: INTRODUCTION

1.1 A Brief History of Alzheimer's Disease

In March 1901, the husband of the 50-year old Auguste D brought his wife to the attention of the German physician Dr. Alois Alzheimer. Auguste D suffered from sleep disorders, disturbances of memory, aggression, crying, and progressive confusion (1). Alois Alzheimer went on to describe the haunting case noting her profound memory loss, unfounded suspicions about her family, and other worsening psychological changes. Dr. Alzheimer was the first to clinically describe the disease that we know today as Alzheimer's Disease (AD) (1), which is a progressive neurodegenerative disorder and the leading cause of dementia in the world with 16.1 million patients (5.4 million in the United States) currently diagnosed with AD (2, 3). Disease rates are projected to increase with the total annual number of new cases projected to double by 2050 (3).

1.1.1 Neuropathology of Alzheimer's Disease

Dr. Alzheimer’s research was unique for his time due to his willingness to combine the fields of physiology and pathology to link cognitive symptoms to microscopic brain abnormalities (1, 4). In Auguste's brain autopsy, he saw dramatic shrinkage and abnormal deposits in and around neurons. Using Bielschowsky's newly-developed silver staining method (5), Dr. Alzheimer observed degenerating neurons with bundles of fibrils (later called neurofibrillary tangles) and miliary foci of silver-staining deposits scattered over the cortex (senile plaques). Both neurofibrillary tangles and senile plaques are now known to be the core pathologies of AD (6).

For decades Alzheimer's findings went largely unnoticed. Technological advances played key roles in revitalizing research. Namely, the invention of the electron microscope and the creation of the cognitive measurement scale accelerated AD research (7, 8). For example, it was not until Max Knoll and Ernst Ruska co-invented the electron microscope that detailed imaging of human neurons became possible (9, 10). The next era of AD research started in the 1960s with

the realization that the majority of cases with senile dementia actually had AD pathology (11). This changed AD from a rare neurological curiosity to a major research priority. In a series of influential papers, Blessed, Tomlinson, and Roth assessed patients' cognitive abilities and their brains, post-mortem, for neuropathology (11) and made the important discovery that senile plaque density correlated with senile dementia and low cognitive scores (11). This study of sixty individuals added credence to Alios Alzheimer's earlier observation. It should also be noted, while there was a correlation between senile plaque density and dementia, the authors also reported a subset of individuals with high plaque burden, but no measurable losses in cognition, suggesting that senile plaques were not sufficient to cause dementia (11). This study ignited research into the neuropathology of the human brain leading to several studies describing senile plaques and tangles and their relationship to antemortem cognition (9, 10).

Using electron microscopy, neurofibrillary tangles (NFT) were described as having a paired helical filament structure (10) and, nearly in parallel, other studies focused on senile plaques describing their fibrous structure (9, 12). These structural studies where able to morphologically describe plaque and tangles, but their biochemical composition still remained unknown. Answering this question required in-depth chemical analysis and was a major focus of AD research in the 1980s. Glenner and Wong biochemically enriched the core component of senile plaques using brute-force fractionation, liquid chromatography, and cutting edge N-terminal sequencing, which enabled the determination of the partial sequence of amyloid-beta (Aβ) (13). Other researchers later confirmed their findings (14). Based on these pioneering biochemical studies, Glenner and Wong suggested that mutations in the gene coding the Aβ peptide could cause AD (13). To this end, characterization of the Aβ peptide sequence helped lead to the discovery of the amyloid precursor protein (APP) gene, located on chromosome 21 (15). The identity of Tau as the major component of NFTs came through several methods, first from immunocytochemistry of isolated NFTs also known as tangles and later via protein

purification and sequencing of Tau peptides (16-18). Recently the presence of plaques and tangles can be determined *in vivo* using Pittsburgh compound B and various Tau tracers with positron emission tomography (PET) imaging (19, 20). Nonetheless, these techniques remain experimental and prohibitively expensive for general use in the population for the diagnosis of AD (19, 20).

1.1.2 Genetics of Alzheimer's Disease

With the increase in availability of genetic sequencing and molecular cloning the genetic causes for rare forms of early onset AD began to be understood in the early 1990s. It was through the study of families with autosomal dominant forms of AD that the first genetic risk factors were discovered. Dr. Michael Mullan and others began to characterize mutations found in familial cases of Alzheimer's Disease (FAD) (21). The families examined had high AD prevalence and early onset in the patients' 40s or 50s (22). The study of these families lead to a seminal paper by Alison Goate and John Hardy, whom discovered a mutation in APP later known as the "London mutation" linked to AD (21, 22). While mutations in APP cause FAD, the majority of FAD cases are caused by mutations in the proteins that process Aβ. In 1995, St. George-Hyslop and colleagues identified presenilin 1 (*PS1*) located on chromosome 14 as the gene underlying a subtype of early onset FAD (23, 24). Then examination of ethnic Volga-Germans with FAD lead to the discovery of the related presenilin 2 (*PS2*) located on chromosome 1 (25). PS1 and PS2 are both components of γ-secretase an enzyme that cleaves APP and aids in Aβ production (26, 27). In addition, patients with Down's syndrome have a higher risk for Alzheimer's Disease due to a third copy of *APP* found on chromosome 21 (15, 28). Although mutations in *APP*, *PS1* and *PS2* are thought to cause AD by increasing Aβ production and eventually plaque burden (21, 23, 29, 30), mutations in these genes are rare and account for less than 2-3% of all AD cases (31). Other risk factors contribute more to the occurrence and prevalence of AD (32). Although, Tau tangles are a hallmark of AD, mutations in *Tau* do not cause AD, but instead cause a clinically distinct form of dementia called frontotemporal lobar degeneration (FTLD) (33, 34).

In contrast to FAD, late-onset AD (LOAD; diagnoses over the age of 65) is much more common and is considered polygenic (estimated range from 50-70% heritable) (35), with multiple loci influencing ones risk of getting LOAD (36). The most common genetic risk factor for LOAD is Apolipoprotein E (*APOE*) (37-39). APOE is a lipoprotein that aids in lipid transport (40) and is commonly found in three variants that differ in the specific amino acid residues, APOE2 (cys112, cys158), APOE3 (cys112, arg158), and APOE4 (arg112, arg158). (40, 41). Notably, the variant APOE4 decreases clearance of Aβ and is a risk factor for AD (42, 43). The APOE2 variant is thought to be protective and decreases chances for developing AD, while the AOPE3 variant is viewed as neutral (44). The APOE4 allele is present in approximately half of patients who develop AD compared to 20–25% in controls (38) and increases AD risk in a dose-dependent manner. For example, one copy of the *APOE4* allele increases AD risk approximately 3-fold while having two increases it by 12-fold (38). Evidence supports that APOE is involved in Aβ clearance, whereas mutations in *APP*, as well as presenilin 1 and 2 influence Aβ production. Collectively, the genetic and pathological evidence (Aβ senile plaques) set the foundation for the amyloid hypothesis in AD described below.

1.2 Amyloid Hypothesis

Disruptions in Aβ metabolism have been linked to AD histologically, biochemically, and genetically. The accumulation of Aβ in senile plaques is a hallmark of every AD case (11) and every mutation directly linked to FAD influences some aspect of Aβ metabolism (45, 46). Duplications of the native *APP* in Down's syndrome on chromosome 21 increases the risk of AD (14), whereas protective mutations on *APP* that decrease the chances of developing AD have also been found (47). *APOE4* is the most common genetic risk factor for AD and has established functional roles in decreasing Aβ clearance in animal models (43). Together, these data provide a strong mechanistic foundation for the amyloid hypothesis (48, 49) that asserts that accumulation of Aβ in the brain is the central force driving AD pathogenesis (48-51).

1.2.1 Limitations of the Amyloid Hypothesis

Since the formulation of the amyloid hypothesis, there have been questions regarding whether this hypothesis completely describes AD pathogenesis. There are observations to suggest that secondary mechanisms contribute to AD. For example, one such observation is that Aβ accumulation occurs in cognitively normal asymptomatic individuals (11, 52, 53). The early studies examining senile plaques noted that cognitively normal older adults often have plaques (11) and subsequent autopsy studies have found, the prevalence of amyloid burden among cognitively normal elderly persons to be more than 25% (54). This observation suggests that the presence of Aβ is necessary, but not sufficient for the development of AD. The same hypothesis suggest that amyloid is just part of the process, perhaps the trigger that sets into motion a whole cascade of events that leads to symptomatic AD. In addition, therapies aimed at altering Aβ metabolism have failed to be effective AD treatments (55). Thus, other targets and processes outside of Aβ accumulation are likely to contribute to AD pathogenesis (56). A focus of my work is uncovering and understanding aspect of AD pathogenesis outside of Aβ metabolism.

As of 2018, 56 drugs that target Aβ production or clearance have gone through clinical trials and failed (57). Immunotherapies have successfully removed Aβ plaques from patient brains, but did not alleviate symptoms or slow down the progress of the disease (58). None of the tested treatments have produced a discernible functional recovery or, altered the course of disease. In fact, inhibitors of γ-secretase (gene products of *PS1* and *PS2*), which lowers Aβ production lead to an increased decline in cognition potentially due to off target effects (59). There are caveats to the failure of Aβ drug trials, one being that perhaps therapies are given too late in disease progression to be effective (6). A second caveat is that a specific subset of Aβ conformers, not targeted by current immunotherapies, could be responsible for AD pathogenesis (60). Despite these caveats, the insufficiency of Aβ accumulation alone to cause dementia and the failures of past Aβ therapies strongly suggests a need to better understand the downstream

molecular mechanisms that cause cognitive decline in AD. With each successive failure, the validity of Aβ toxicity singularly contributing to AD pathogenesis becomes less likely.

Other challenges face AD research, including the limitations of mouse models (61). Many sophisticated *APP* and/or *PS1* mice models exist that express one or more copies of the genes that cause FAD (62), yet none of these models completely recapitulates the cellular and behavioral phenotypes of AD (61, 63). Although the morphological resemblance to AD amyloidosis is impressive, the cognitive defects in mice are much more mild then those seen in AD (63). Furthermore none of the APP mouse models develop Tau tangles, one of the two histological hallmarks of AD (61-63). Other mouse models do have Tau tangles, but this is done independently of Aβ metabolism by over-expression of human mutant Tau transgenes that cause Tauopathies like FTLD, but do not cause AD (33, 64). This suggests that the molecular mechanisms during AD pathogenesis that link Aβ and Tau tangles do not manifest in mice, which limits how we can use mouse models to understand one of the key stages of AD pathogenesis. Hence, I argue that the study of human brain is likely necessary to understand the progression and staging of AD.

1.3 Alzheimer's Disease can be viewed as a continuum

Converging neuropathological, biochemical and neuroimaging evidence indicates that the pathophysiological stage of AD begins years, if not decades, before the diagnosis of clinical dementia (53, 65). In the simplest of models, the earliest stage of AD is the accumulation of Aβ plagues following by intraneuronal aggregation of phosphorylated Tau, and finally cognitive decline (65). This progression from control to asymptomatic AD (AsymAD), to mild cognitive impairment (MCI), and finally to AD takes place over a decade or more (53). A naming scheme to describe the stages of disease has been developed (53, 65). For example, control cases are defined by normal cognitive scores and lack of pathology. (11, 65), while the next stage, AsymAD, is defined by significant deposition of senile plaques and mild/moderate Tau tangles

but a lack of clinical dementia. Once cognition starts to decline the patient transitions to MCI, and eventually AD (Fig. 1.1) (53). The processes that connect these stages are largely unknown and represents a major gap in knowledge in the field. For example, while Aβ's role in AD pathogenesis is well established the exact mechanisms by which Aβ leads to Tau aggregation is not known (Fig. 1.1). Understanding this process could illuminate the mechanisms of Aβ toxicity that contribute to dementia.

1.4 Proteomics as a discovery tool to understand the pathophysiology of Alzheimer's Disease

Proteomics is the study of a set of proteins that compose an organism, system, or biological context, referred to as the proteome (66). Proteomics strives to answer basic questions such as: what proteins are in a sample? What is their abundance? The proteome is not constant; it differs from cell to cell and changes over time. To some degree, the proteome reflects the underlying transcriptome (67). However, protein abundance is influenced by many factors in addition to the expression level of the relevant gene. Proteomics can be used to answer several types of questions including abundance, turnover rate, location, post-translational modification (PTM), protein activity, and interactions (68). While many techniques have been historically used for proteomics including two-dimensional differential gel electrophoresis, immunoassays, and Edman degradation, it is mass spectrometry (MS) that stands as the most powerful proteomic technique. Remarkably, a mass spectrometer is able to measure thousands of proteins from a single sample with concentrations as low an attomolar (10^{-18}) (69). Thus, proteomics is an ideal strategy for identifying new pathologic proteins linked to disease transition and progression. Just as the technological advances of electron microscopy and genetic sequencing advanced AD research in that last century, MS-based proteomics is now providing new advances in understanding AD pathophysiology. Proteomics has the advantage of being "unbiased", that is no particular protein target is necessary allowing for large number of proteins to be measured.

Previous unbiased genome wide association studies (GWAS) have discovered novel AD risk factors and therapeutic targets, adding credence to the effectiveness of unbiased approaches (70).

1.4.1 Proteomic analysis of detergent-insoluble proteome across neurodegenerative disease reveals U1 snRNP aggregation in AD brain

Each neurodegenerative disease has its own subset of aggregated proteins (71, 72). Historically determining the identity of these aggregated proteins has revealed substantial information regarding neurodegenerative pathogenesis. As noted above in sections 1.1.1 and 1.1.2, discovery of amyloid beta (Aβ) as the major component of insoluble plaques in AD brain launched the discovery of the genetic and biochemical mechanisms that support the etiological roles of Aβ in AD pathogenesis (6, 13, 17, 21). Aggregated proteins are enriched in the detergent in-soluble fraction of brain (73) and therefore, identifying novel insoluble proteins in AD brain could provide insight towards mechanisms linking Aβ and tau, and more generally help describe AD etiology.

To identify new and unknown aggregate-prone proteins in AD that could contribute to disease pathogenesis our group undertook a comprehensive study of the human brain detergent-insoluble proteome in AD using MS-based proteomics (74). To identify novel aggregated proteins specific to AD, the insoluble fraction of AD brain was compared to that of control, MCI, Parkinson's, FTLD-tau, FTLD-u (with TDP-43 positive inclusions), and Corticobasal degeneration (CBD) (74). This led to the discovery that a group of RNA splicing factors aggregate exclusively in AD (74). Specifically, the U1 small nuclear ribonucleoprotein 70 kDa (U1-70K) and other components of the U1snRNP form detergent-insoluble aggregates in AD (74). Consistent with these biochemical observations, U1-70K is mislocalized to the cytoplasm and associated with fibril-like aggregates exclusively in AD (Fig1.2A-B). Similar findings were found in familial cases of AD (74, 75). Moreover, U1-70K localizes to Tau tangles, but can also be found in tau free intraneuronal aggregates (Fig.1.2B) (76). U1-70K aggregates are detected in

mild cognitive impairment (MCI) cases indicating that U1-70K abnormalities occur early in AD progression perhaps before tau aggregation (74, 76). Aggregation of U1-70K in familial cases of AD and Down's Syndrome, strongly links Aβ metabolism with U1-70K aggregation (75). When examining the insolubility of Aβ and tau, our group discover that U1-70K correlates with both Aβ and Tau insolubility. Strikingly U1-70K is the most correlated protein with Aβ insolubility (76). How or if Aβ and Tau influence U1-70K aggregation remains a topic of interest (Fig. 1.3). In addition, understanding U1snRNP protein dysfunction in AD may reveal new mechanisms that contribute to AD pathogenesis and link Aβ and Tau aggregation.

1.4.2 The biological function of the U1 snRNP complex

The U1 snRNP is a core component of the spliceosome (77), which is a protein-RNA complex responsible for splicing together introns and removing exons from pre-mRNA transcripts. For genes containing two or more exons different patterns of intron retention are possible leading to alternative splicing (78). Alternative splicing allows for one gene to generate different mRNAs that encode distinct protein products (78). Recent estimates, based on sequence read analysis of exon–exon junctions, indicate that 92–94% of human genes undergo alternative splicing (79). By selectively regulating alternative splicing events, the cell can choose to express different splice variants with distinct functions at specific time in development and under certain physiological conditions.

The U1 snRNP is composed of the U1 snRNA, the proteins U1-70K, U1A, U1C, and the seven sm proteins that come together to form the sm ring (77). The U1 snRNP recognizes and defines the 5' splice site and is the most abundance snRNP of the spliceosome (80). The binding of the U1 snRNP to the 5' splice site signifies the formation of the spliceosome early-complex (e-complex) (81). The binding of the U2 snRNP which defines the 3' site brings about the a-complex (81). Interactions between the U1 snRNP and U2 snRNP regulate splice site selection and thus alternative splicing (81). These interactions are mediated by serine/arginine (SR) proteins that

simultaneously interact with U1 snRNP and U2snRNP across the intron (82, 83). U1-70K a core component of the U1 snRNP complex is hypothesized to bind SR proteins and facilitate the interactions between the U1 snRNP and U2snRNP (84). In addition to the U1 snRNP's role in splice site selection it is also involved in polyadenylation and inhibits premature cleavage of mRNA (85). Given the U1 snRNP's vital functions in mRNA processing, we would expect U1snRNP dysfunction to be detrimental to neurons in AD.

1.4.3 Other links between the U1snRNP and Alzheimer's disease

The disruption of U1 snRNA levels induces AD-like pathology in cell culture including APP over-expression, tau hyperphosphorylation, lower acetylated α-tubulin levels, and up-regulation of caspase-8 (86, 87). Furthermore, next generation RNA-seq analysis from AD and control brains revealed a significant accumulation of unspliced or pre-mRNA species in AD, consistent with a loss of U1-spliceosome function (74). The loss of U1-70K function results in increased Aβ production suggesting that U1-70K aggregation has an important, yet unclear mechanistic role in APP processing (74). To date, our understanding of the specific mechanisms underlying U1-70K aggregation in AD is limited. This has proved to be a barrier to developing cellular models that would further our understanding of U1-70K and related RNA-binding protein aggregation events in the pathogenesis of AD. Concepts gleaned by studying RNA-binding protein aggregation in other neurodegenerative disease could therefore provide insight into U1-70K aggregation.

1.5 Links between RNA granules and neurodegenerative disease

Similar to U1-70K in AD, several RNA binding proteins (RBPs) including TDP-43, FUS, TIA-1, hnRNPA1/B2, and ATXN2, aggregate in other neurodegenerative disease such as amyotrophic lateral sclerosis (ALS) and FTLD (88-92). It has been hypothesized that the prevalence of (RBPs) aggregating in neurodegenerative diseases arises from the fact that RBPs normally form large macromolecular assemblies termed RNA granules that are naturally poised

for aggregation (Fig1.4) (93-96). Evidence by the McKnight's group indicate that low complexity domains (LC domains) in RBPs are responsible for normal RNA granule assembly (94-96). LC domains are polypeptide sequences with limited amino acid diversity, and often contain imperfect tandem repeats of amino acid (96, 97). Many of the RNA binding proteins that aggregate in disease contain LC domains including FUS, TDP-43, hnRNPA1/B2 and U1-70K. The LC domains found in TDP-43 and FUS mediate self-association, are necessary for RNA granule formation, and polymerize into amyloid-like aggregates (96, 98, 99). Mutations harbored within the LC domains of TDP-43 and FUS cause ALS and FTD (100) The ability of TDP-43 to co-immunoprecipitate other copies of TDP-43, or "self-associate" increases with ALS causing mutations (101). These mutations also increase TDP-43 RNA granule stability and removal of this domains ablates TDP-43 granule localization (99). The assembly of RBPs has been linked to neurodegeneration for example, the self-assembly of FUS contributes to neurodegeneration in *Drosophila* (102). Pathological mutations in hnRNPA1 linked to ALS increase the speed in which hnRNPA1 transition into a solid aggregate like state, connecting RBP assembly to protein aggregation (103). Notably, at high concentration *in vitro*, LC domains undergo a concentration-dependent phase transition to a gel-like state composed of uniformly polymerized amyloid-like aggregates that can sequester normal soluble forms of the same protein or other RNA binding proteins with similar, but not identical, LC domains (94, 96). Collectively, these observations have led to a hypothesis that pathological aggregation of RNA-binding proteins in disease arises from endogenous cellular pathways that regulate RNA granule assembly (93, 100, 104, 105). Therefore understanding how low complexity domains facilitate assembly of RNA granules could reveal the biochemical mechanism of protein aggregation in disease (93).

1.5.1 RNA Granules

Cells are compartmentalized, allowing for spatial separation and concentration of components (106). This separation and concentration is a strategy that cells use to regulate interactions and

biochemical reactions (107). Compartmentalization is often achieved via membrane bound organelles such as the mitochondria and the nucleus (108). These organelles are surround by lipid membranes which form a boundary between the organelle and its' surroundings. It is well understood that membranes allow separate compartments to coexist (109, 110). However, less intuitively, many compartments do not have membranes (111). Examples are nucleoli, which make ribosomes inside the nucleus; centrosomes, which nucleate microtubules; Cajal bodies, which aid spliceosome assembly; splicing speckles, which likely play a role in mRNA maturation; and stress granules, which take various forms under different stress conditions (106, 112-117). Collectively these membrane-free organelles composed of RNA and RBPs are called RNA granules (112). Recently there has been a heightened interest in RNA granules due to their possible role in neurodegenerative disease (93, 118).

1.5.2 Biophysical properties of RNA granules

Supporting evidence indicates that a select group of RBPs are poised for aggregation because they self-assemble to form structures, including RNA granules (96, 119, 120). It has been proposed that RNA granules form via liquid–liquid phase separation (LLPS) (118). LLPS is the process where compounds in solution come together and fall out of solution to form a liquid droplets (106). The components and biophysical properties within the droplet are distinct from those of the surrounds. The concentration within these compartments can be as high as 200mg/ml (121). Evidence supporting RNA granules liquid state include their ability to; fuse with one another, deform by flows, and the ability to spread out and "wet" a surface (122, 123). These liquid like properties are in contrast to McKnight's "hydrogels" which are solid (96, 119, 124). The exact biochemical and biophysical mechanisms that attribute to these properties is an active area of research.

Other observations suggest that RNA granule are composed of highly dynamic sets of fibrils (125). Current theories propose at least two different layers within a granule, a stable core

and a highly dynamic surface layer (126). Understanding the biophysical properties of RNA granules is still a new and active field. While the exact nature of RNA granules is still debated, a consensus has formed that granules represent a distinct phase, separate from the cytoplasm that forms through the assembly of RNA and RBPs (106, 107).

The prevailing theory in the field is that disordered LC domains are vital in facilitating LLPS (103, 127). In this case, "disorder" refers to lack of temporally stable structures (127). The repetitive nature of LC domains could contribute to multivalent interactions and ultimately LLPS. Proteins lacking LC domains and with stable structures can also participate in LLPS (121, 128). Biophysical and biochemical studies have demonstrated that multivalency, that is the ability to interact with a substrate through several binding sites facilitates LLPS (121, 129, 130). Classic concepts in polymer science indicate that multivalent molecules naturally assemble into large oligomers or polymers when mixed (107). The assembly will inherently decrease the solubility of the molecules due to entropy-driven effects, thus promoting their phase separation (131). Thus, it is not surprising that oligomerization and phase separation are linked (131). In summary, while LC disordered domain often participate in LLPS many other type of domains can too.

1.5.3 Tau, LLPS and RNA binding proteins

In accordance with biophysical properties of RBPs, recent evidence now indicates that Tau undergoes LLPS *in vitro* (132, 133). This process is enhanced by polyanions, such as heparin (132) and RNA (134), as well as phosphorylation on Tau (132).

Recent findings from our group and others suggest that aggregation of RBPs is a pathological feature of diseases that are characterized by tau aggregation (75, 135, 136). Various RBPs, including the stress granule nucleating protein TIA1 and the splicing speckle protein U1-70K, co-localize with hyperphosphorylated tau in patient tissues (75, 135). Pathological tau promotes the aggregation of RBPs, including TIA1, to form stress granules (135). Reducing TIA1 *in vivo* protects against neurodegeneration, improved memory, and prolonged the lifespan in

PS19 transgenic tau mice, a FTLD mouse model (137). This suggests that the toxicity of Tau is dependent on interactions of RBPs with tau and adds tauopathies such as AD and FTLD to the growing list of diseases linked to the dysfunction of RNA metabolism (137). If U1-70K interacts with Tau, and how that interaction influence AD pathogenesis is a question examined in the following chapters (Fig.1.3).

1.6 Properties of U1-70K aggregation

Previous work done by our lab provided insight into U1-70K aggregation mechanisms. A key observation was that nearly all detectable U1-70K was found in the sarkosyl detergent-insoluble fraction of AD brain tissue following biochemical fractionation, despite the fact that most neurons and glia harbor nonfilamentous nuclear U1-70K. (Fig.1.2A) (91).

The brain is complex tissue composed largely of neurons and glia (138). Glial cell subtypes are varied and include astrocytes, oligodendrocytes, and microglial (138). Neurons are directly responsible for memory, cognition, and learning, but glial cells are more numerous in the frontal cortex making up majority of the volume (139). In AD, only some neurons harbor filamentous U1-70K (Fig.1.2A), while glial cells contain normal nuclear staining of U1-70K (74, 91). Paradoxically, nearly all detectable U1-70K was found in the sarkosyl-insoluble fraction of AD brain following biochemical fractionation (74). This conflicts with the observation that the vast majority of U1-70K is likely in soluble pools within the glial and neuronal nuclei. This observation led to the hypothesis that soluble U1-70K was being induced to aggregate by some biomolecule released during homogenization (Fig. 1.3).

This hypothesis was tested by adding soluble forms of U1-70K to AD homogenate and assaying for induced insolubility. It was found that human AD brain homogenates induced the aggregation of normal soluble U1-70K from control brain and cell culture (91). The insoluble fraction of AD brain, but not control or CBD induced the insolubility of U1-70K, demonstrating disease specificity (91). This induced insolubility is not influenced by RNAase treatment, yet is

sensitive to protease treatment suggesting that it is facilitated by protein-protein interactions (91). U1-70K fragments directly interact with aggregated U1-70K in AD brain (91). These data together lead to the hypothesis that following tissue homogenization, aggregated U1-70K and possibly other proteins sequester natively folded soluble U1-70K, causing it to aggregate in AD (Fig. 1.3). A similar mechanism could occur during the progression of AD whereby U1-70K seeds its own aggregation.

Further studies identified determinates of aggregation in the primary structure of U1-70K (91). U1-70K harbors two LC domains, LC1 (amino acids 231-308) and LC2 (amino acids 317-407), which together comprise a majority of the C-terminal primary sequence. The C-terminus of U1-70K was necessary for AD brain homogenate induced aggregation (91). Furthermore, the LC1 domain (residues 231-308) of U1-70K is necessary and sufficient for robust aggregation in AD brain homogenates, and directly interacts with insoluble U1-70K (91). These data together suggest that the LC1 domain of U1-70K is responsible for U1-70K pathological aggregation in AD brain. The role of the LC1 domain in aggregation was interesting, in the context that other LC domains of other RNA binding proteins have been found to be important in RNA granule formation and aggregation during neurodegenerative disease (93).

1.6.1 Special Characteristics of the U1-70K LC1 domain

The LC domains of U1-70K are different than those found in other RBPs that aggregate in neurodegenerative disease. The LC domain of TDP-43, FUS, TIA-1 and related proteins are enriched with polar non-charged residues such as glycine, glutamine, and asparagine also referred to as Q/N rich domains (99, 120, 140). In contrast the LC domains of U1-70K are enriched with charged amino acids, arginine, lysine, glutamic and aspartic acid (Fig. 1.5). The LC1 domain has an unusually high charge density with 86.2% of amino acids carrying charge. Due to the presence of positive and negative charges I refer to the LC1 and similar domains as mixed charge (MC). The main difference between the LC1 and LC2 domains, is that the LC2 domain is slightly more

diverse with the addition of glycine and proline. The prolines in the LC2 domain likely limit flexibility (141). The LC domains of U1-70K are predicted to be disordered and are not resolved in the U1snRNP crystal structure (77, 91). Due to this and other limitations, the function of the LC1 domain and MC domains in general was unknown at the start of this thesis research (Fig.1.3). It was unknown if the functions of these domains are related to their aggregate prone nature (Fig1.3).

1.7 Research focus and Innovation

The mechanism of U1-70K aggregation in AD in unknown. The self-assembly of RNA binding proteins happens along mechanisms parallel to that of pathological aggregation (section1.5). My thesis work seeks to determine whether the MC domain of U1-70K is required for self-assembly. In addition, I aim to define the biological function of the U1-70K MC domain and other similar MC domains. The mechanism of Aβ toxicity and its' relationship to Tau aggregation is not understood. Evidence suggest (section 1.4.1) that U1-70K could act a bridge between Aβ and Tau. I examine the interactions between Aβ, Tau, and U1-70K to gain insight into the mechanism linking Aβ accumulation and tau aggregation (Fig.1.3). This study explores the hypothesis that U1-70K links these two pathological events though physical protein-protein interactions (Fig.1.3). The long-term goal is to define the mechanism of U1-70K aggregation in the greater context of AD pathogenesis. We performed immunohistochemistry, oligomerization assays, immunoprecipitation, mass spectrometry, network correlation analysis, and various other data analysis to define the function of the U1-70K LC1 domains and to examine aggregation prone properties of U1-70K and other related proteins. This work has led to many discoveries regarding the role that U1-70K and other similar proteins play in molecular assembly, RNA biology, and AD pathogenesis. In the upcoming section, I will discuss the results from the various experiments listed above.

Progression of AD and Gaps in Knowledge

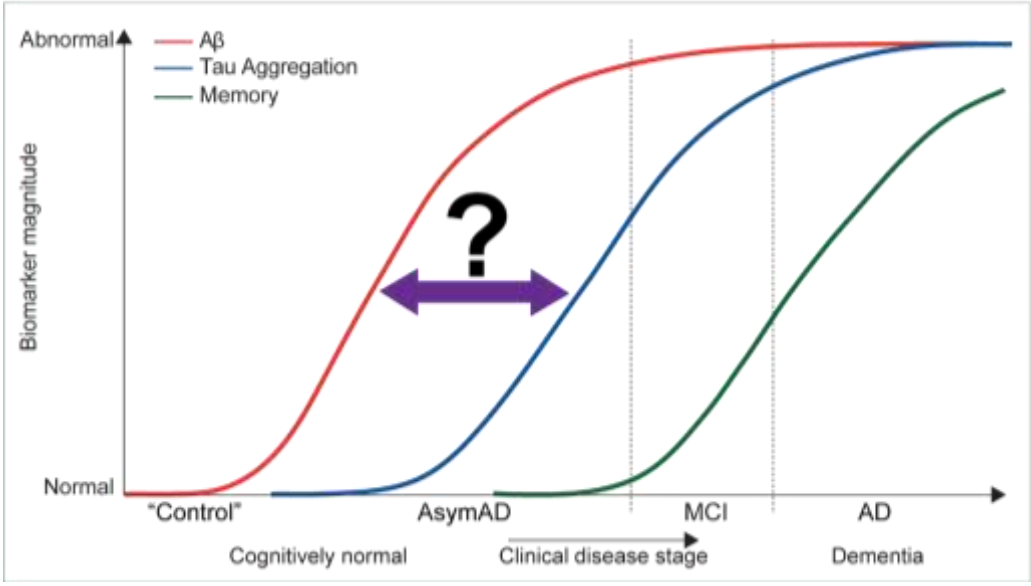


Figure 1.1 Hypothetical model of dynamic biomarkers of the AD. Biomarkers change from normal to maximally abnormal (y-axis) as a function of disease stage (x-axis). Disease progresses from control to AsymAD, to MCI, and finally AD. The molecular processes that connect Aβ accumulation and Tau aggregation are largely unknown and represent an area of research I focus on in my studies. Adapted from (53).

Immunohistochemistry of AD Tissue

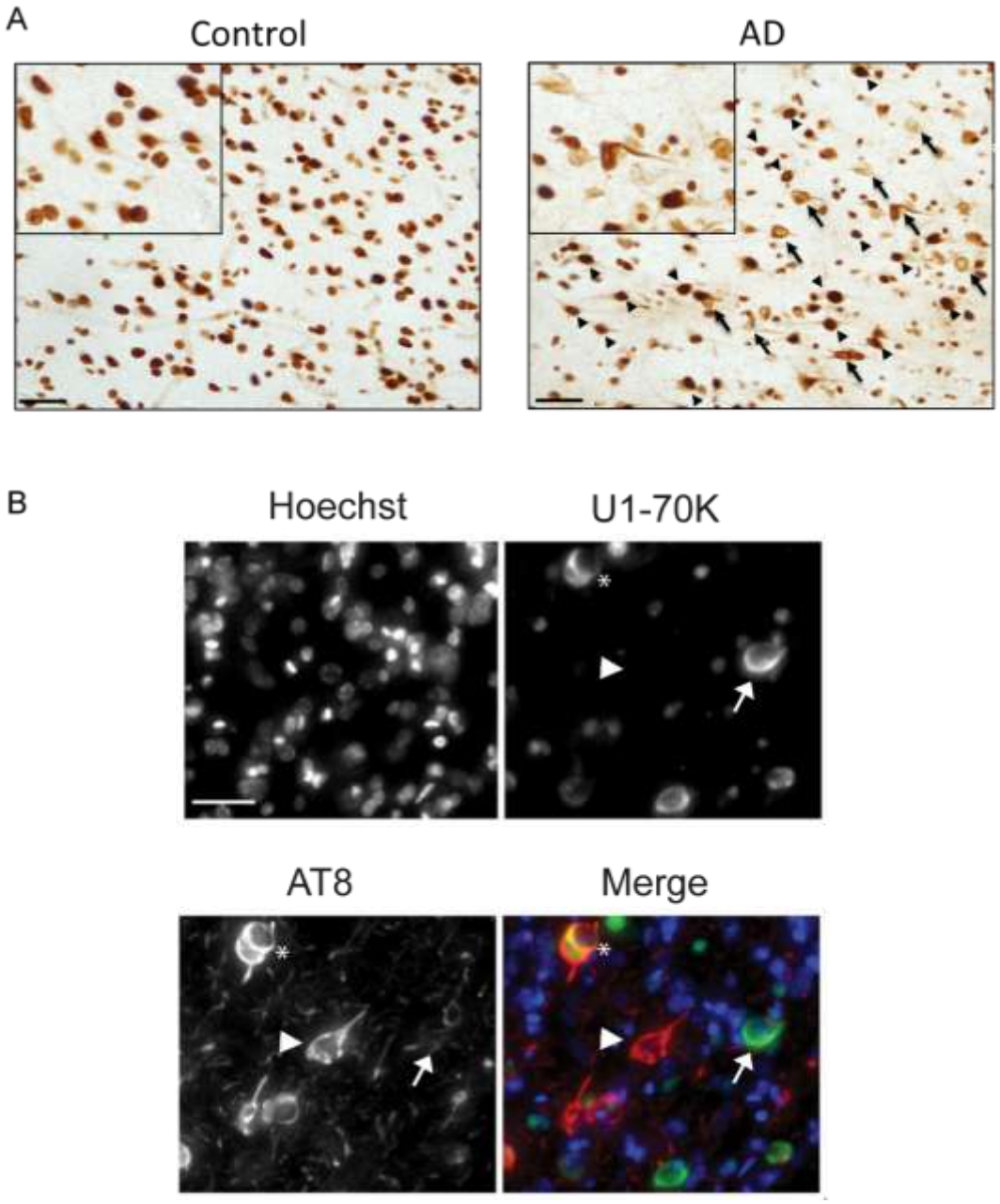


Figure 1.2 A. Immunohistochemistry using antibodies against U1-70K in both control (left panel) and AD brain (right panel) tissue (scale bare 50 μm and higher magnification shown in the insets). Arrows highlight cytoplasmic U1-70K tangles. Notably, the overwhelming majority of U1-70K exhibits normal nuclear localization in AD brain. (arrow heads). **B.** Double-labeled immunofluorescence of cells in AD prefrontal cortex displaying cytoplasmic aggregates of U1-70K alone (green, arrow), tau alone (red, arrowhead), and both (asterisk). Hoescht dye was used to stain nuclei (blue). Scale bar is 20 μm. Adapted from (76, 91).

Model for U1-70K Aggregation and Gaps in Knowledge

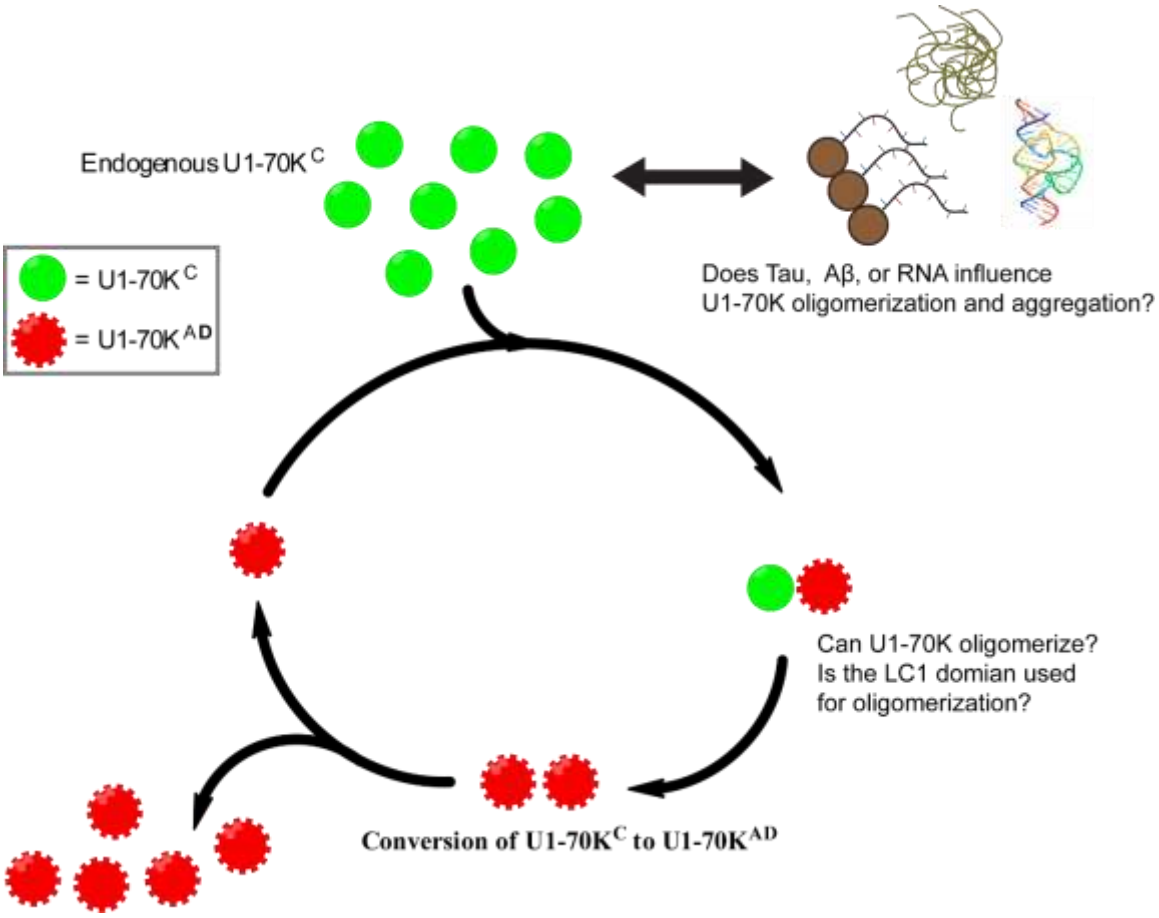


Figure 1.3. The previously proposed model for U1-70K aggregation, and the gaps in knowledge that will be focused on in chapter two.

Ribostasis Theory of RBP Aggregation

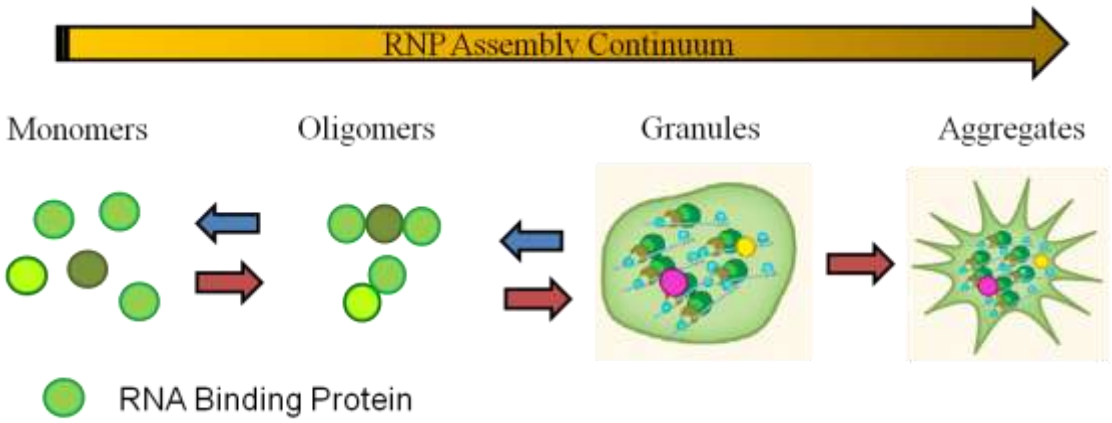


Figure 1.4 Cytoplasmic RNPs diffuse though the cyto or nucleoplasm, transiently interacting. Through multivalent LC domains RNPs form oligomers then condensates that go through LLPS to form RNA granules. Under normal metabolism, these granules reversibly assemble and disassemble. Through mechanism incompletely understood RNA granules can transition to aggregates. Adapted from (93).

The U1-70K LC1 domains has an unusual amino acid composition

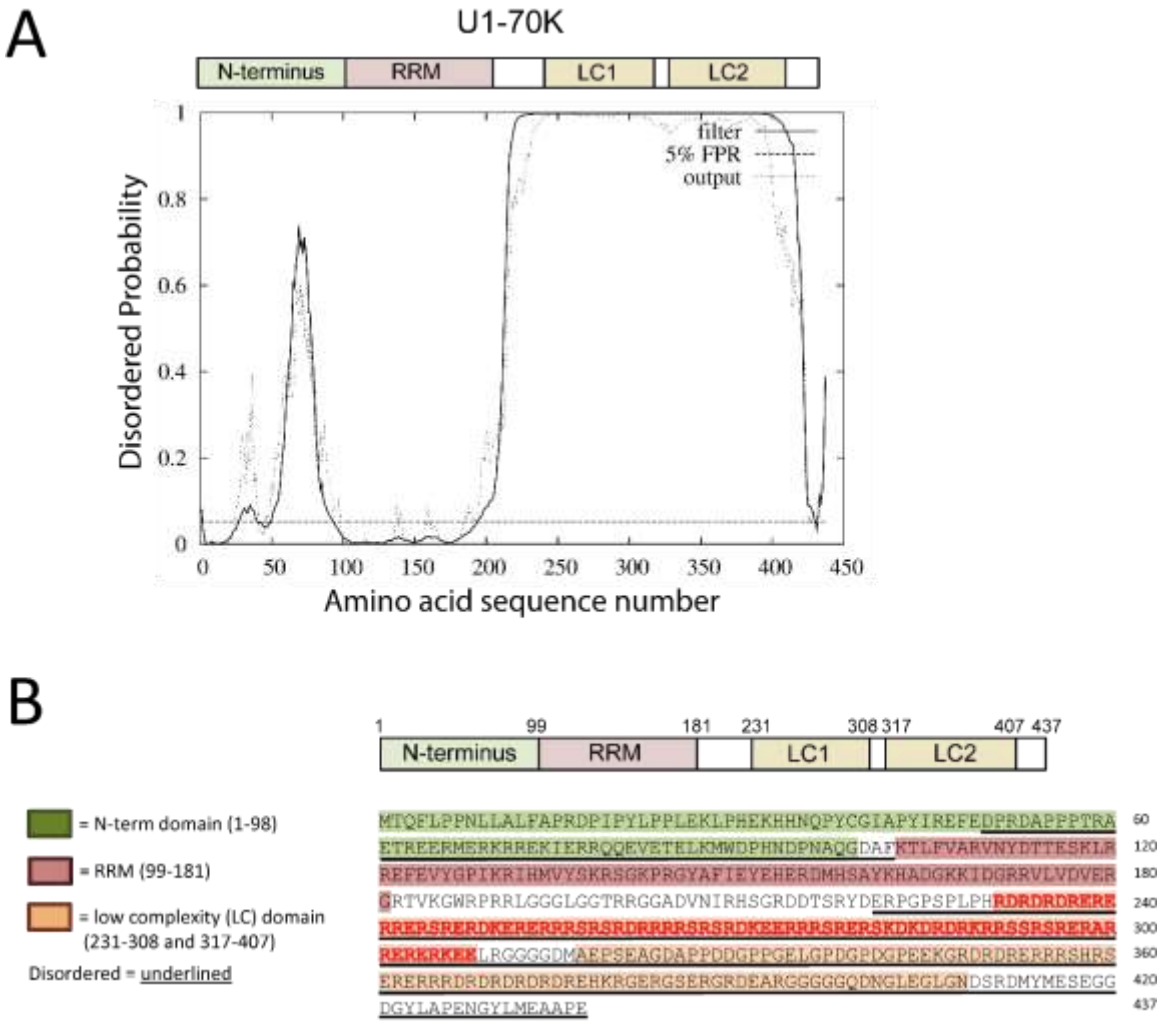


Figure 1.5 U1-70K is highly disordered and has C-terminal low complexity domains. A. DISOPRED algorithm predicts intrinsically disordered regions within the N-terminal (amino acids 50–100) and C-terminal (amino acids 220–437) portion of U1-70K, which harbors two distinct low complexity domains, LC1 (amino acids 231–308) and LC2 (amino acids 317–407). **B.** Primary amino acid sequence of U1-70K with the N-terminal (green), RNA recognition motif (RRM) domain (red), and two LC domains (orange) are highlighted. The LC1 domain is highlighted with red text. Adapted from (91).

CHAPTER 2: RNA-binding proteins with mixed charge domains self-assemble and aggregate in Alzheimer's Disease

Isaac Bishof[1,4], Eric B. Dammer[1,4], Duc M. Duong[1,4], Marla Gearing[3,4], James J. Lah[2,4], Allan I. Levey[2,4], and Nicholas T. Seyfried[1,2,4#]

[1]Department of Biochemistry, [2]Department of Neurology, [3]Department of Pathology and Laboratory Medicine, [4]Center for Neurodegenerative Diseases, Emory University School of Medicine, Atlanta, GA, 30322

2.1 ABSTRACT

ABSTRACT

U1-70K and other RNA binding proteins (RBPs) are mislocalized to cytoplasmic neurofibrillary Tau aggregates in Alzheimer's disease (AD), yet the mechanisms that cause their aggregation are incompletely understood. Many RBPs that aggregate in neurodegenerative diseases self-assemble into RNA granules through intrinsically disordered low complexity (LC) domains. We report here that a LC domain within U1-70K of mixed charge, containing highly repetitive complementary repeats of basic (R/K) and acidic (D/E) residues, shares many of the same properties of the Q/N-rich LC domains found in the RBPs TDP-43 and FUS. These properties include the ability to self-assemble into oligomers, and to form nuclear granules. To analyze the functional roles of the U1-70K LC domains, we performed co-immunoprecipitation and quantitative mass spectrometry analysis of recombinant U1-70K and deletions lacking the C-terminal LC domain(s). A network-driven approach resolved functional classes of U1-70K interacting proteins that showed dependency on the U1-70K LC domain(s) for their interaction. This included structurally similar RBPs, such as LUC7L3 and RBM25, which require their respective mixed charge domains for reciprocal interactions with U1-70K and for participation in nuclear RNA granules. Strikingly, a significant proportion of RBPs with mixed charge domains have elevated insolubility in the AD brain proteome compared to controls. Furthermore, we show that the mixed charge LC domain of U1-70K can interact with β-amyloid and Tau from AD brain. This supports a hypothesis that β-amyloid directly or indirectly mediates interactions between mixed charge structural motifs on U1-70K and related RBPs with pathological Tau in AD.

2.2 INTRODUCTION

The molecular processes that contribute to neurodegenerative diseases are not well understood. Recent observations suggest that numerous neurodegenerative diseases are promoted by the accumulation of RNA-binding protein (RBP) aggregates (1-3). This includes Alzheimer's disease (AD), where pathological RNA-protein aggregates are often, but not exclusively, associated with Tau neurofibrillary tangles in brain (4). For example, U1 small nuclear ribonucleoprotein 70 kDa (U1-70K) and other core components of the spliceosome complex form detergent-insoluble aggregates in both sporadic and familial human cases of AD (5-7). Furthermore, RNA-seq analysis from AD and control brains revealed a significant accumulation of unspliced pre-mRNA disease related transcripts in AD consistent with a loss of U1-spliceosome function (7,8). Currently, our knowledge of the specific mechanisms underlying U1-70K aggregation is limited. This has proved to be a barrier to developing cellular models that would further our understanding of U1-70K and related RBP aggregation events in the pathogenesis of AD.

Supporting evidence indicates that a select group of RBPs are poised for aggregation because they self-assemble to form structures, including RNA granules (9), which are membrane-free organelles composed of RNA and RBPs (1,9-12). It has been proposed that RNA granules form via liquid–liquid phase separation (LLPS), which is driven by a dynamic network of multivalent interactions between structurally disordered low complexity (LC) domains (13-15) that have limited diversity in their amino acid composition (16). LLPS allows specific RBPs to concentrate and separate, leading to the formation of higher-order structures including oligomers, granules, and ultimately aggregates (16-18). Notably, several RBPs that aggregate in neurodegenerative disease contain LC domains, including TDP-43 and FUS (19-21). The LC domains found in TDP-43 and FUS mediate self-association, are necessary for RNA granule formation, and polymerize into amyloid-like aggregates (9,22,23). Mutations harbored within the

LC domains of TDP-43 and FUS cause amyotrophic lateral sclerosis (ALS) and increased RNA granule stability, highlighting a critical role for LC domains in disease pathogenesis (24-26).

We recently reported that human AD brain homogenates induced the aggregation of soluble U1-70K from control brain and recombinant U1-70K, rendering it detergent-insoluble (5). The C-terminus of U1-70K, which harbors two LC domains (LC1 and LC2) was necessary for this aggregation (5). Furthermore, the LC1 domain (residues 231-308) of U1-70K was sufficient for robust aggregation, and through cross-linking studies was found to directly interact with insoluble U1-70K in AD brain homogenates (5). Collectively, these observations led to a hypothesis that pathological aggregation of RBPs in neurodegenerative diseases, including U1-70K, is driven by LC domains. However, unlike the prion-like Q/N-rich LC domains of TDP-43 and FUS, the LC1 domain of U1-70K contains highly repetitive complementary basic (R/K) and acidic (D/E) residues. These structurally unique motifs were originally described by Perutz (27), who proposed their ability to self-assemble and form higher-order structures termed polar zippers (27). Currently the physiological role of polar zipper motifs in U1-70K and other RNA binding proteins is unclear, and understanding their role in protein-protein interactions may shed light on the mechanisms underlying RBP aggregation and their association with Tau in AD (7,28).

Here we report that the mixed charge LC1 domain of U1-70K shares many of the same properties of the Q/N-rich LC domains found in TDP-43 and FUS, despite having a vastly different amino acid composition. These properties include the ability to self-assemble into high molecular weight oligomers and associate with nuclear granules in cells. To analyze the functional roles for the LC domains in U1-70K, we performed co-immunoprecipitation of recombinant U1-70K and serial deletions lacking one or both LC domains followed by quantitative proteomic analysis. Using a network-based bioinformatic approach we mapped classes of U1-70K interacting proteins that showed a dramatic reduction in their association with U1-70K in the absence of the LC1 domain. Remarkably, this revealed a group of functionally and structurally similar RBPs that also contained mixed charge domains analogous to the LC1 domain

in U1-70K. These included LUC7L3 and RBM25, which we confirm also require their respective mixed charge domains for reciprocal interactions with U1-70K and for proper nuclear RNA granule association in cells. Furthermore, global analysis of the AD detergent-insoluble proteome revealed elevated levels of mixed charge RBPs within AD brain compared to controls. Finally, we show that the LC1 domain of U1-70K specifically interact with Tau from AD brain, but not out tauopathies which supports a hypothesis that mixed charge structural motifs on U1-70K and related RBPs could mediate cooperative interactions with Tau in AD.

2.3 RESULTS

2.3.1 The LC1 domain of U1-70K is necessary and sufficient for self-association in cells

The U1-70K LC1 domain (residues 231-308) is necessary and sufficient for robust aggregation in AD brain homogenates (91), yet it remains unknown if this domain is required for endogenous U1-70K self-association under physiological conditions. To test this hypothesis, we over-expressed full-length recombinant GST-fused and Myc-tagged U1-70K (rU1-70K) in HEK293 cells with serial deletions lacking one or both LC domains followed by co-immunoprecipitation (co-IP) and western blot analysis (**Fig. 2.1A**). Full-length rU1-70K (WT) and the ΔLC2 mutant co-immunoprecipitated endogenous U1-70K (~55 kDa), while rU1-70K mutants lacking the mixed charge LC1 domain (ΔLC1 and ΔLC1+2) displayed a dramatic impairment in their ability to co-IP native U1-70K. Together, these observations indicate that the LC1 domain is necessary for U1-70K self-association. To determine if the LC1 domain is sufficient for self-association, co-IPs were performed following the over-expression of alternative rU1-70K truncations, which included the N-terminus expressed alone (residues 1-99), the N-terminus and RNA recognition motif (residues 1-181), LC1, and the LC2 domain expressed in isolation (**Fig. 2.1B**). Only the LC1 domain was sufficient for appreciable self-association with native U1-70K. Furthermore, this interaction was likely not influenced by the presence of RNA, as treatment of the lysates with RNase prior to co-IP did not impair the ability of rU1-70K to interact with native U1-70K (**Fig. 2.1C**). Thus, our findings support that the LC1 domain is necessary and sufficient for U1-70K self-association, and that this interaction is predominantly RNA-independent.

2.3.2 The U1-70K LC1 domain oligomerizes in vitro

Although the LC1 domain of rU1-70K was deemed sufficient to interact with native U1-70K in cells, it was unclear whether this interaction is direct or facilitated by indirect interactions with additional RBPs. To determine if the LC1 of U1-70K can directly self-

associate, we performed blue native gel polyacrylamide gel electrophoresis (BN-PAGE) of the GST-purified LC1 (residues 231-310) and N-terminal domain (residues 1-99) of rU1-70K; the latter was unable to interact with native U1-70K in cells (**Fig 2.2B**). In contrast to SDS-PAGE, which resolves proteins under denaturing conditions, BN-PAGE is used to determine native protein complex masses, including high molecular weight oligomeric states and to identify physiological protein–protein interactions (150). Under the denaturing conditions of SDS-PAGE (**Fig. 2A**), both the LC1 and N-terminal domain have equivalent molecular weights (~65 kDa) compared to purified GST (~20 kDa). However, under native conditions (**Fig. 2B**), the LC1 domain formed dimers, trimers, tetramers, and high-molecular weight oligomers that were in the megadalton MW range (>1,236 kDa). In contrast, the N-terminal domain mainly existed in the monomeric and dimeric state with some evidence of lower abundance high-molecular weight oligomers; GST alone was almost exclusively monomeric (**Fig. 2.2B**). These complexes became more readily visible following the transfer to a membrane and western blot analysis with Myc antibodies (**Fig. 2.2C**). Remarkably, a higher proportion of the LC1 domain formed dimers (n=2) and tetramers (n=4) compared to trimers (n=3), suggesting that dimer intermediates are favored over trimer intermediates for tetramer formation (**Fig. 2.2D**). These *in vitro* findings demonstrate that the LC1 domain can directly self-associate to form oligomers, including high molecular weight species, which implicates direct LC1-LC1 interactions as a mechanism of U1-70K self-association.

2.3.3 The U1-70K LC1 domain is necessary and sufficient for robust nuclear granule localization

Our data highlights that a key role for the U1-70K LC1 domain is self-association and oligomerization. However, it is not well established whether the LC1 domain influences nuclear localization and granule formation in cells. To address this question, full-length rU1-70K or mutants lacking one or both LC domains (**Fig. 2.1A**) were over-expressed in HEK293

cells followed by subcellular biochemical fractionation into nuclear and cytoplasmic pools (**Fig. 2.3A-B**). The rU1-70K mutants containing a LC1 domain (WT and ΔLC2) partitioned mainly to the nuclear fraction (~75% nuclear), but mutants lacking the LC1 domain (ΔLC1 and ΔLC1+2) were equally distributed between nucleus and cytoplasm, indicating a significant impairment of nuclear localization. These biochemical findings were further supported by immunocytochemistry, which showed that rU1-70K mutants lacking the LC1 domain displayed diffuse expression patterns in both the nucleus and cytoplasm compared to WT and ΔLC2 proteins. (**Fig. 2.3C**). Given that the LC1 domain in isolation can directly self-associate and oligomerize (**Fig. 2.2**), we sought to determine if the LC1 domain was necessary for RNA granule localization in cells (**Fig. 2.3C**). As expected, full-length rU1-70K protein localized to nuclear granules, in agreement with previous studies (151-153). Although the nuclear granule localization of mutants lacking the LC1 domain was diminished, the ΔLC2 mutant retained nuclear granule localization, supporting the necessity of the LC1 domain for subnuclear granule localization. Furthermore, the LC1 domain alone was found to colocalize with native U1-70K in nuclear granules (**Fig. 2.3D**), consistent with the ability of the LC1 domain to interact with native U1-70K from cell lysates (**Fig. 2.1D**). Collectively, these results demonstrate that the LC1 domain is important for U1-70K subcellular nuclear localization, and suggests a role for LC1 mediated intermolecular interactions as a mechanism of nuclear RNA granule formation.

2.3.4 Protein-Protein interaction network analysis resolves functionally distinct classes of U1-70K interacting proteins

To further assess the physiological function of the LC domains of U1-70K, we performed co-immunoprecipitations of full-length rU1-70K and various rU1-70K mutants lacking either or both LC domains from HEK293 cells followed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) to identify co-purified interacting

proteins. Each co-IP was performed in biological quadruplicate (*n*=4) and an equal number of mock IPs were performed using a non-specific immunoglobulin (IgG) as a negative control. Protein abundance was determined by peptide ion-intensity measurements across LC-MS/MS runs using the label-free quantification (LFQ) algorithm in MaxQuant (154). In total, 45,223 peptides mapping to 3,458 protein groups were identified. However, one limitation of data-dependent LFQ proteomics methods is the inherent missing data (i.e. missing protein identifications or abundance values), especially for low abundance proteins (155). Thus, proteins with ten or more missing values across the 20 samples were not included in the subsequent bioinformatic analysis. To limit the number of non-specific interactors, proteins with less than a 1.5-fold enrichment over IgG were not considered. This resulted in the final quantification of high-confidence interactors falling into 716 protein groups mapping to 713 unique gene symbols (**Table 1**).

WeiGhted Co-expression Network Analysis (WGCNA) is typically used for large-scale transcriptome and proteome datasets to categorize gene products into biologically meaningful complexes, molecular functions and cellular pathways (156).

Given the power of WGCNA to identify novel protein complex and functions, we sought to leverage co-enrichment patterns to better classify protein-protein interactions (PPIs) across WT and rU1-70K deletions to assess if specific classes of proteins selectively favor interactions with the LC domains. In WGCNA, correlation coefficients between each protein pair in the dataset is calculated and groups of highly correlated proteins are segregated into modules (157). In our dataset, a total of 7 modules were defined (**Fig. 2.4** and **Fig. 2.5A**). These modules range from 292 proteins in turquoise to 19 in black (**Table 2.1**). The premise of co-expression or in this case protein co-enrichment analysis is that the strong correlation between two or more proteins is indicative of a physical interaction, functional relationship, and/or co-regulation. We therefore hypothesized that following a co-IP for rU1-70K, specific modules would reflect biologically relevant PPIs, and thus highlight distinct complexes. As expected, modules were significantly

enriched for biologically meaningful gene ontologies (GO) as well as established cellular functions and/or organelles as determined by GO-Elite (**Table 2.1**).

Each module has an abundance profile for all member proteins across the rU1-70K co-IP conditions, termed the eigenprotein (**Fig. 2.5A-B**). Notably, six (M1-M6) of the seven modules showed a significantly higher level of co-enrichment in the WT and rU1-70K deletions compared to the IgG negative controls, indicative of specific interactions for members of these modules (**Fig 2.5A-B**). In contrast, M7 (black) had essentially equivalent levels across all conditions and was the only module with positive correlation to the non-specific IgG. Therefore, it was considered non-specific and not used for further analysis **(Fig. 2.4-5A)**. The protein interactors with reduced affinity for rU1-70K following LC1 deletion (e.g. ΔLC1 and ΔLC 1/2), included members of both the mRNA processing (blue) and snRNP assembly (green) modules (**Table 2.1**). Therefore, the LC1 dependent interactors are defined by membership to these modules. In contrast, modules enriched with large ribosomal subunit components (turquoise) and mitochondrial ribosome subunits (brown) displayed increased levels following deletion of one or both LC domains, suggesting that the LC1 domain negatively regulates their interactions with U1-70K (**Table 2.1 and Fig. 2.5A-B**). Finally, the modules enriched with proteins involved in protein folding (red) and the small ribosomal subunit (yellow) showed little difference in expression across the co-IP conditions, suggesting that these protein interactions are mainly with the N-terminus and/or RRM domain of U1-70K. The kME score quantifies how well a given protein pattern matches that of the module eigenprotein, with high scores approaching one signifying a high correlation

(158). The hub proteins, with the highest correlation to the module abundance profile (i.e., eigenprotein), are highlighted in (**Fig. 2.5B)**. These findings demonstrate that a weighted PPI network analysis of the U1-70K interactome successfully resolved biologically and structurally distinct complexes.

2.3.5 Confirmation of U1-70K LC1 dependent interacting proteins

To validate and extend the module assignments, we performed both *in silico* and biochemical analysis. First, to visualize the relationships among modules with an independent clustering method, the T-Distributed Stochastic Neighbor Embedding (tSNE) algorithm was used to map the relatedness of proteins of top module members. The tSNE analysis largely agreed with and confirmed the module assignment, whereby the majority of proteins clustered with their own module members as assigned by WGCNA (**Fig. 2.6A**). The tSNE analysis also allows for visualization of module relatedness, with distance between proteins and clusters representing similarity of co-enrichment, with similar modules in close proximity to each other and dissimilar modules further apart. For example, modules involved in translation cluster together (brown and turquoise) while those involved in mRNA processing (blue) and snRNP assembly (green) formed a separate cluster.

To experimentally validate the module assignments biochemically, western blots were performed for select rU1-70K interactors of the snRNP assembly module (SNRPD1 and U1A), mRNA processing module (SRSF1, RBM25 and LUC7L3) and the protein folding module (TDP-43) (**Fig. 2.6B**). Members of both the mRNA processing and snRNP assembly modules had higher abundance in the WT and ΔLC2 co-IPs compared to that of the IgG, ΔLC1, and ΔLC1+2 IP samples. This mirrored the pattern observed for the blue and green eigenprotein values confirming the proteomic findings (**Fig. 2.5B**). In contrast, TDP-43, showed a similar level of interaction across WT rU1-70K and the various deletions, consistent with TDP-43 being an N-terminal interactor of U1-70K and not influenced by the absence of the disordered LC domains.

2.3.6 The mRNA processing module is enriched with structurally similar RNA-binding proteins harboring mixed charge domains

U1-70K interacting proteins mapping to the mRNA processing and snRNP assembly modules are related by their affinity for the LC1 domain, yet they contain RNA binding proteins

with distinct biological functions (**Table 2.1**). For example, all U1 snRNP components and assembly factors including U1A, U1C, Sm proteins, and the SMN complex are enriched in the green module. The SMN complex is responsible for loading the Sm proteins onto the snRNA scaffold, a critical step in U1 snRNP assembly (159). This module also contains the components of the 7SK snRNP which regulates snRNA transcription and is present in Cajal bodies, a site of U1 snRNP maturation (160, 161). In contrast, the mRNA processing module (blue) is enriched with proteins associated with RNA splicing, polyadenylation, mRNA export and "nuclear specks", the latter being an analogous term for splicing speckles (162). However, it is their respective association with rU1-70K mutants that discriminates the mRNA processing and snRNP modules members. For example, proteins involved in snRNP assembly are less influenced by the loss of the LC2 domain, yet association of members of the mRNA processing module are affected, suggesting that proteins involved in granule/speckle assembly interact in part via both LC domains of U1-70K, whereas core spliceosome assembly factors do not favor interactions with the LC2 domain (77, 163).

Our observations also revealed that several members of the mRNA processing module (blue) contain stretches of highly repetitive complementary basic (R/K) and acidic (D/E) residues, analogous to the LC1 domain of U1-70K (**Fig. 2.7A**), that we refer to as mixed charge domains. To examine the relationship between this sequence similarity and U1-70K interacting proteins, a list of LC1-like (i.e., mixed charge) proteins was created using the Uniprot protein Blast feature. Many proteins (n=255) in the proteome were determined to have significant sequence overlap to the mixed charge LC1 domain of U1-70K. These included other members of the mRNA processing module such as RBM25, ZC3H13, DDX46 and LUC7L3 among others. Although not identical in length, sequence alignment highlights the similar stretches of highly repetitive complementary basic (R/K) and acidic (D/E) residues across these distinct gene products (**Fig. 2.7A**). Indeed, a one-tailed Fisher's exact test revealed that the mRNA processing module was significantly enriched with proteins harboring LC1-like mixed charge domains generated from

the Blast analysis (**Fig. 2.7B**). In contrast, a similar analysis comparing disordered RNA binding proteins harboring prion-like (Q/N-rich) domains (140), including TDP-43 and FUS, showed no enrichment in any of the modules of U1-70K interacting proteins. Notably, the members of the mRNA processing module also had a significant over-representation of nuclear proteins that selectively precipitate after treatment with biotinylated isoxazole (b-isox) (164). Many of these proteins have been shown to participate in RNA granule assembly and to form hydrogels *in vitro* (96, 119). Collectively these results suggest that structurally similar mixed charge domains, analogous to the U1-70K LC1 domain, engage in protein-protein interactions, which are essential for nuclear granule assembly.

2.3.7 Mixed charge domains in LUC7L3 and RBM25 are necessary for reciprocal interactions with U1-70K and nuclear RNA granule assembly

Based on their related structural and functional roles in RNA speckle assembly and affinity for the LC domains of U1-70K we asked whether members of the mRNA processing module co-localize with U1-70K in cells. Both RBM25 and LUC7L3 have mixed charge domains analogous to the LC1 domain of U1-70K with similarity E-values of 5.5E-27 and 1.8E-26, respectively, and amino acid overlap of 55.1% and 44.2%, respectively (**Fig 2.7A**). As expected, all three proteins were observed in nuclear granules (151, 165, 166), where U1-70K showed strong co-localization with LUC7L3 and RBM25 (**Fig. 2.8A**). Our current findings support that the LC1 domain is necessary and sufficient for U1-70K self-association and nuclear granule assembly. By extension we hypothesized that the mixed charge domains in LUC7L3 and RBM25 would similarly be important in mediating interactions with U1-70K and other structurally similar proteins. To test this possibility, we over-expressed full-length recombinant GST-fused and Myc-tagged rRBM25 or rLUC7L3 in HEK293 cells and their respective deletions lacking the mixed charge domains (ΔMC) and the mixed charge (MC) domains alone followed by IP and western blot analysis (**Fig. 2.8B**). The full-length rLUC7L3 and the MC domain were each able

to co-IP endogenous LUC7L3, mirroring the self-association observed for U1-70K (**Fig. 2.1A**). Unfortunately, the rLUC7L3-ΔMC protein migrated at a similar molecular weight to endogenous LUC7L3, thus, we were unable to determine if the MC domain was necessary for self-association in cells. Full-length rLUC7L3 and the MC domain could also interact with endogenous U1-70K and RBM25 (**Fig. 2.8B**), whereas the rLUC7L3-ΔMC mutant was unable to support these interactions **(Fig. 2.8B)**. Similarly, full-length rRBM25 could interact with both endogenous U1-70K and LUC7L3, while the rRBM25-ΔMC mutant was unable to support these interactions (**Fig. 2.8C**). In contrast, the MC domain of rRBM25 was not sufficient to interact with U1-70K or LUC7L3, perhaps due to misprocessing, post-translation modifications, or size.

Given the role of the U1-70K LC1 domain in RNA granule formation, we sought to determine if the MC domains of LUC7L3 and RBM25 influenced nuclear localization to granules (**Fig. 2.8D and E**). Both full-length rLUC7L3 and rRBM25 localized to nuclear granules by immunocytochemistry, consistent with their endogenous LUC7L3 and RBM25 expression pattern in cells (**Fig. 2.8**). However, the rLUC7L3-ΔMC was observed diffusely localized to the cytoplasm, whereas rRBM25-ΔMC remained mainly in the nucleus, but did not localize to nuclear granules. Consistent with U1-70K, the MC domain of rLUC7L3 was sufficient to localize to nuclear granules, likely due to its interactions with U1-70K and other mixed charge RBPs. However, the MC domain of rRBM25, which was unable to interact by co-IP with mixed charge RBPs, did not form nuclear granules in the cells. Taken together, our findings suggest a shared functional role for mixed-charge domains in stabilizing protein- protein interactions, that likely play a role in nuclear granule assembly.

2.3.8 RNA binding proteins with mixed charge domains have enhanced insolubility in AD brain.

Based on the ability of U1-70K to aggregate in AD brain homogenate, and the key role of the LC1 domain in U1-70K oligomerization *in vitro*, we hypothesized that proteins harboring

similar MC domains would preferentially aggregate in AD brain. To test this hypothesis, we assessed the distribution of insoluble proteins with MC domains in a recently published comprehensive analysis of the sarkosyl-insoluble proteome (*n*=4,643 proteins quantified) from individual control and AD cases (167). Protein ratios for all pairwise comparisons (i.e., control vs. AD) were converted into $\log_2$ values, and the resulting histogram fit to a normal Gaussian distribution (**Fig. 2.9A**). Compared to the normal distribution of all proteins in the AD insoluble proteome (purple histogram), quantified mixed charge proteins (yellow histogram) showed a global shift towards insolubility in AD (**Fig. 2.9A**). This increase was found to be significant using one-tailed Fisher exact test (p-value=2.028868e-09). Consistently, mixed charge proteins that fell within the top 10^{th} percentile (*n*=28) were significantly elevated in AD cases compared to controls, similar to Aβ and tau levels (**Fig. 2.9B and C**). Strikingly, 68% of the AD enriched mixed charge proteins, within the top 10^{th} percentile, including LUC7L3, were members of the blue module from rU1-70K interactome studies **(Fig. 2.5**), with shared functions in RNA binding, splicing and processing (**Fig. 2.9D and E**). In contrast, RBM25 insolubility was not significantly elevated in AD brain despite having a MC domain. Thus, while RBPs with MC domains clearly have a higher likelihood of insolubility and aggregation in AD brain, the presence of an MC domain alone is not sufficient for aggregation.

2.3.9 The LC1 domain of U1-70K interacts with pathological Tau from AD brain

We have previously reported an association of aggregated U1-70K with Tau neurofibrillary tangles in both sporadic and familial cases of AD, but not in other tauopathies (74, 75, 91, 168). However, mechanisms underlying the specificity Tau and U1-70K co-aggregation in AD are poorly understood. Similar to the biophysical properties of RBPs, recent evidence now indicates that Tau undergoes LLPS *in vitro* (132). This process is enhanced by polyanions, such as heparin (132) and RNA (134), as well as phosphorylation on Tau (132). Based on these observations, we sought to assess if the LC1 domain of U1-70K could interact with pathological

Tau from human AD brain. Equivalent amounts of GST purified LC1 domain or the N-terminal domain of rU1-70K were added to AD brain homogenates and immunoprecipitated with anti-myc antibodies followed by a western blot for Tau (**Fig. 2.10A**). Compared to the N-terminal domain, the LC1 domain of rU1-70K co-immunoprecipitated significantly more Tau, including modified Tau species of altered molecular weights (**Fig. 2.10B**).

To test whether the interaction between the LC1 domain of rU1-70K and Tau was specific to AD brain, GST purified LC1 or N-terminal rU1-70K proteins (negative control) were added to homogenates generated from control (n=6), AD (n=6) and tauopathy (n=6) brain tissue. The latter group included progressive supranuclear palsy (*n*=1) and corticobasal degeneration (*n*=5) cases. Following immunoprecipitation with anti-myc antibodies, samples were analyzed by mass spectrometry (MS). Consistent with western blot results, significantly more Tau was identified and quantified by MS in LC1 domain immunoprecipitates compared to the N-terminal domain in AD brain homogenates (**Fig. 2.10B-C**). Furthermore, MS analysis revealed that the LC1 domain selectively enriched Tau from AD compared to tauopathies (**Fig. 2.10C**), indicating that the Tau-LC1 interaction is specific to AD. Moreover, the LC1, but not the N-terminal domain, of rU1-70K significantly interacted with Aβ in AD (**Fig. 2.10D**). Together, these data suggest that the association of U1-70K and other mixed charge RBPs with Tau are specific in AD and are mediated by direct or indirect interactions (via Aβ) with MC domains.

2.4 DISCUSSION

In this study we show that the mixed charge LC1 domain of rU1-70K can directly self-interacts *in vitro* to form high molecular weight oligomers and that this domain is also necessary and sufficient for U1-70K self-association in cells. This suggests that U1-70K oligomerizes through the LC1 domain (**Fig. 1.3**). To learn more about the biological function of the LC domains in U1-70K, I used a quantitative proteomics and network based approaches (WGCNA) to define functional classes of U1-70K interacting proteins that favored interactions with the LC1 domain. This revealed a class of structurally similar RBPs that also contained analogous mixed charge LC domains. We show that for at least two other RBPs, LUC7L3 and RBM25, their respective mixed charge domains are required for reciprocal interactions with U1-70K and for proper localization to nuclear granules. Furthermore, global analysis of the detergent-insoluble proteome in human brain revealed elevated levels of mixed charge RBPs in AD suggesting a shared mechanism of protein aggregation across this class of RBPs. Finally, I show that the LC1 domain of U1-70K interacts with Aβ and Tau derived from human AD brain tissue, but not other tauopathies. This supports the hypothesis that mixed charge structural motifs on U1-70K and related RBPs could mediate cooperative protein-protein interactions with Aβ and Tau in AD. These observations for U1-70K and more generally MC proteins are consistence with research relating to FUS, TDP-43, and other RBPs that have Q/N prion-like domains and aggregate in ALS and FTLD. However, in contrast to the oligomerization and aggregation of Q/N-rich prion-like domains, the MC domains likely oligomerize via electrostatic interactions that drive protein-protein interactions, granule, and aggregate assembly. Thus, I propose that physical interactions with Aβ and or Tau are plausible modulators of U1-70K and MC protein aggregation as observed in AD (**Fig 1.3**), which I further discuss in the next chapter.

2.5 EXPERIMENTAL PROCEDURES

Materials

Primary antibodies used in these studies include an in-house rabbit polyclonal antibody raised against a synthetic KLH-conjugated peptide corresponding to a C-terminal epitope of U1-70K (EM439) (74), an anti-Myc-tag (clone 9B11, Cell Signaling), an anti-GST (ab6613 Abcam) antibody, an anti-LUC7L3 (HPA018484-100UL, Sigma), an anti-RBM25 (ab72237, Abcam), an anti-U1-70K monoclonal (05-1588, Millipore), an anti-tau (ab54193, Abcam), IgG mouse control (550339 BD Pharmigen). Secondary antibodies were conjugated to either Alexa Fluor 680 (Invitrogen) or IRDye800 (Rockland) fluorophores.

Plasmids and Cloning

The original cDNA of U1-70K containing C-terminal myc and DDK tags was cloned from pCMV6- Entry vector (Origene) and inserted into the HindIII/BamHI sites in the pcDNA3.1 vector (91). Full-length and U1-70K deletions sequences were subsequently cloned into the EcoRV/XhoI sites in the pLEXM-GST vector for the expression of N-terminal and C-terminal GST-tagged proteins. Similar cloning strategies were performed using LUC7L3 (Origene RG214406) and RBM25 (Origene RC212256) plasmids. All cloning was performed by the Emory Custom Cloning Core Facility and plasmids confirmed by DNA sequencing.

Immunoprecipitation

Human embryonic kidney (HEK) 293T cells (ATCC CRL-3216) were cultured in Dulbecco's Modified Eagle Medium (DMEM, high glucose (Gibco)) supplemented with 10% (v/v) fetal bovine serum (Gibco) and penicillin-streptomycin (Gibco) and maintained at 37 °C under a humidified atmosphere of 5% (v/v) CO_2 in air. For transient transfection, the cells were grown to 80-90% confluency in 10 cm^2 culture dishes and transfected with 10 μg expression plasmid and 30 μg linear polyethylenimine. Cells were homogenized in ice-cold immunoprecipitation (IP) buffer containing (50 mM HEPES pH 7.4, 150 mM NaCl, 5% glycerol, 1 mM EDTA, 0.5% (v/v)

NP-40, 0.5% (v/v) CHAPS, Halt phosphatase inhibitor cocktail (1:100, Thermo Fisher). Samples were sonicated for 5 seconds on 5 seconds off at 30% amplitude for a total of 1.5 minutes (13 cycles). The samples were cleared (14,000 x g for 10 minutes) and protein concentrations determined using a standard bicinchoninic acid (BCA) assay (Pierce). Protein A Sepharose 4B beads (Invitrogen 101042; 20 uL per IP), were washed twice in IP buffer and then blocked with 0.1 mg/ml bovine serum albumin (Thermo #23209) and washed three additional times in IP buffer. Then anti-Myc (4µg) mouse monoclonal antibody (Cell Signaling 2276) or 4 µg IgG control (550339 BD Pharmingen) was allowed to incubate rotating with the bead slurry in IP buffer (500µL) for a minimum of 90 minutes to allow antibody conjugation to beads. Beads were washed 3 times in IP buffer. Lysates where pre-cleared by centrifugation at 14,000 x g at 4^{o}C for 10 minutes. The pre-cleared protein lysates were added to beads (1.5 mg per IP) and incubated rotating overnight at 4^{o}C. The beads were washed 3 times in IP wash buffer (IP buffer without glycerol or CHAPS) by centrifugation at 500 x g for 5 min at 4^{o}C then resuspended in IP wash buffer. Following the last wash, the bead suspension was transferred to a new Eppendorf tube to minimize contamination. The bound protein was eluted with 8M urea buffered in 10 mM Tris pH 8.0. For proteomics assays, 4 independent biological replicates were performed for each condition. For protein digestion, 50% of the eluted protein samples were reduced with 1 mM dithiothreitol (DTT) at 25°C for 30 minutes, followed by 5 mM iodoacetamide (IAA) at 25°C for 30 minutes in the dark. Protein was digested with 1:100 (w/w) lysyl endopeptidase (Wako) at 25°C for 2 hours and diluted with 50 mM NH_4HCO_3 to a final concentration of less than 2M urea. Samples were further digested overnight with 1:50 (w/w) trypsin (Promega) at 25°C. Resulting peptides were desalted with in-house stagetips and dried under vacuum.

RNAase A Treatment

Cells were lysed in IP buffer with the addition of 5 mM $MgCl_2$. Following sonication and centrifugation as described above, the lysates were split and treated with RNAase A or buffer

alone (control) to a final concentration of 50 µg/ml of RNAase A. The RNAase A treated and control samples were incubated for 30 minutes at room temperature followed by centrifugation at 10,000 x g for 10 minutes at 4^{o}C. The supernatant was added to beads and the IP was completed as detailed above.

Blue Native Gel Electrophoresis

Recombinant N-terminal (residues 1-99) and LC U1-70K fragments (residues 231-308) were purified and their concentrations determined as described previously (91). The purified GST used as control was kindly gifted by the Dr. Richard Kahn (Emory University, Department of Biochemistry). Prior to analysis purified rU1-70K fragments were cleared by centrifugation at 20,000 x g for 15 minutes at 4^{o}C to remove any insoluble precipitates. Each protein (0.8µg) was added to blue native gel loading buffer (5% glycerol, 50mM TCEP, 0.02% (w/v) G250 coomassie, 1x Native Page Running Buffer [Invitrogen BN2001]) and allowed to incubate at room temperature for 30 minutes. Samples were loaded onto a 3-12% NativePAGE Bis-Tris Gel (Invitrogen BN2011BX10) in addition to a native gel molecular weight marker (Thermo LC0725). Samples were resolved by electrophoresis at 150 volts for 1.5 hours in anode Native PAGE Running Buffer (Invitrogen BN2001) and cathode buffer with additive (Invitrogen BN2002). Gels were de-stained overnight in a solution of 15% (v/v) methanol and 5% (v/v) acetic acid and protein visualized on the Odyssey Infrared Imaging System (Li-Cor Biosciences). For western blot analysis, native gels were prepped with a 30 minutes incubation at room temperature in 1% (v/v) SDS and then transferred using the semidry iblot transfer system (Invitrogen) onto nitrocellulose (IB23001).

Immunocytochemistry

Cells were plated on Matrigel (Corning #356234) coated coverslips and prepared for transfection using lipofectamine (ThermoFisher) according to manufacturer's protocol. Immunocytochemistry was performed 48 to 72 hours after transfection essentially as described (169). After the blocking

step, slides were dabbed to remove excess liquid and incubated in primary antibody overnight at 4^{o}C. Primary antibodies included: rabbit anti-U1-70K (EM439), mouse anti-Myc-tag, mouse anti-LUC7L3, mouse anti-RBM25, mouse anti-U1-70K. The slides were washed 3 times with PBS 0.05% (v/v) saponin then incubated with secondary antibody (Dylight 549, Alexa 488) for one hour shaking at room temperature. Again, slides were washed 3 times with PBS with 0.05% saponin. DAPI diluted in PBS was added to each slide and incubated for at least 30 minutes rotating at room temperature. Following additional rinses in PBS, cells were mounted in Vectashield (Vector Laboratories, Burlingame, CA) and sealed with nail polish. Images were captured on an FLUOVIEW FV1000 confocal laser scanning microscope (Olympus).

Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)

Tryptic peptides were analyzed by LC-MS/MS essentially as described (170). Peptides were resuspended in loading buffer (0.1% formic acid, 0.03% trifluoroacetic acid, 1% acetonitrile) and separated on a self-packed C18 (1.9 um Dr. Maisch, Germany) fused silica column (20 cm x 75 μM internal diameter; New Objective, Woburn, MA) by a NanoAcquity UHPLC (Waters, Milford, FA) and monitored on a Q-Exactive Plus mass spectrometer (ThermoFisher Scientific, San Jose, CA). Elution was performed over a 140-minute gradient at a rate of 300 nL/min with buffer B ranging from 3% to 80% (buffer A: 0.1% formic acid and 5% DMSO in water, buffer B: 0.1 % formic and 5% DMSO in acetonitrile). The mass spectrometer cycle was programmed to collect one full MS scan followed by 10 data dependent MS/MS scans. The MS scans (300-1800 m/z range, 1,000,000 AGC, 100 ms maximum ion time) were collected at a resolution of 70,000 at m/z 200 in profile mode and the MS/MS spectra (2 m/z isolation width, 28 normalized collision energy (NCE), 50,000 AGC target, 50 ms maximum ion time) were acquired at a resolution of 17,500 at m/z 200. Dynamic exclusion was set to exclude previous sequenced precursor ions for 30 seconds. Precursor ions with +1, and +6 or higher charge states were excluded from sequencing. The mass spectrometry proteomics data have been deposited to the

ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008260.

Database Search

Raw data files were analyzed using MaxQuant v1.5.2.8 with Thermo Foundation 2.0 for RAW file reading capability (171). The search engine Andromeda was used to build and search a concatenated target-decoy UniProt Knowledgebase (UniProtKB) containing both Swiss-Prot and TrEMBL human reference protein sequences (90,411 target sequences downloaded April 21, 2015), plus 245 contaminant proteins included as a parameter for Andromeda search within MaxQuant (172). Methionine oxidation (+15.9949 Da), asparagine and glutamine deamidation (+0.9840 Da), and protein N-terminal acetylation (+42.0106 Da) were variable modifications (up to 5 allowed per peptide); cysteine was assigned a fixed carbamidomethyl modification (+57.0215 Da). Only fully tryptic peptides were considered with up to 2 miscleavages in the database search. A precursor mass tolerance of ±20 ppm was applied prior to mass accuracy calibration and ±4.5 ppm after internal MaxQuant calibration. Other search settings included a maximum peptide mass of 6,000 Da, a minimum peptide length of 6 residues, 0.05 Da tolerance for high resolution MS/MS scans. The false discovery rate (FDR) for peptide spectral matches, proteins, and site decoy fraction were all set to 1%. The label free quantitation (LFQ) algorithm in MaxQuant (154, 173) was used for protein quantitation.

Protein-protein interaction network analysis

The R package Weighted Gene Correlation Network Analysis (WGCNA) was used to sort proteins into functional groups by examining relative levels of co-enrichment (158). In WGCNA, correlation coefficients between each protein pair in the dataset are first calculated and transformed continuously with the power adjacency function to generate an adjacency matrix that defines the connection strength between protein pairs. This adjacency matrix is then used to calculate a topological matrix (TO), which measures the interconnectedness or correlation

between two proteins and all other proteins in the matrix. All proteins are then hierarchically clustered (e.g. average linkage) using 1-TO as a distance measure and module assignments are subsequently determined by dynamic tree cutting (158). Threshold power Beta for reduction of false positive correlations (i.e. the beneficial effect of enforcing scale free topology) was sampled in increments of 0.5, and as the target scale free topology R^2 was approached, 0.1. The power selected was the lowest power at which scale free topology R^2 was approximately 0.80, or in the case of not reaching 0.80, the power at which a horizontal asymptote (plateau) was nearly approached before further increasing the power had a negative effect on scale free topology R^2. Other parameters were selected as previously optimized for protein abundance networks (170). Thus, for the signed network built on protein LFQ abundances obtained from IP-LC-MS/MS, parameters were input into the WGCNA::blockwiseModules() function as follows: Beta 10.9, mergeCutHeight 0.07, pamStage TRUE, pamRespectsDendro TRUE, reassignThreshold $p<0.05$, deepSplit 4, minModuleSize 15, corType bicor, and maxBlockSize greater than the total number of proteins. T-Distributed Stochastic Neighbor Embedding (tSNE) analysis was performed as described (174)**.** Proteins with WGCNA intramodular kME>=0.50 were retained, and all duplicated values removed, as well as proteins with any missing values for the 16 non-IgG measurements. Then Rtsne R package Barnes-Hut-Stochastic Neighbor Embedding (SNE) Rtsne function was run on the LFQ expression matrix to reduce dimensionality from 16 to 2. The remaining points or proteins (*n*=375) were colored according to WGCNA module membership. Gene Ontology (GO) Elite analysis on each module was performed as described previously (170).

Bioinformatic analysis of mixed charge proteins in the detergent insoluble proteome in AD brain

Quantitative proteomic analysis using isobaric tagging of sarkosyl-insoluble fractions (frontal cortex) from eight AD and six control cases was previously performed as described Supplementary proteomic data was downloaded from (167) and R was used to generate

histograms, Fisher exact p-value, box plots, and the clustered heat map with the Nonnegative Matrix Factorization (NMF) package.

Nuclear and Cytoplasmic Fractionation

The fractionation protocol was performed as essentially described in (175) with slight modifications. Briefly, after transfections with full-length rU1-70K plasmids or respective mutants, HEK cells were harvested by scraping and washed with PBS including 1x Protease Inhibitor Cocktail (Sigma). Cells were then spun down at 1000 x g at 4°C for 5 minutes and carefully resuspended in hypotonic buffer (10mM HEPES pH 7.9, 20mM KCl, 0.1mM EDTA, 1mM DTT, 5% glycerol, 0.5 mM PMSF and Halt phosphatase inhibitor cocktail [1:100, Thermo Fisher]) and incubated on ice for 15 mins. The detergent NP-40 was then added to a 0.1% (v/v) final concentration and cells were briefly agitated by vortex and left on ice for 5 additional minutes followed by centrifugation for 10 minutes at 4°C at 15,600 x g. affording the supernatant (S1) as the cytoplasmic fraction and the pellet (P1) as the nuclear fraction. To determine if difference in nuclear and cytoplasmic distribution were significant across conditions repeated measures ANOVA with post-hoc Tukey was performed in GraphPad Prism.

Immunoprecipitation of rU1-70K fragments from human brain homogenates

Post-mortem frontal cortex tissue from pathologically confirmed AD cases were provided by the Emory Alzheimer's Disease Research Center (ADRC) brain bank (**Table 2.2)**. Neuropathological evaluation of amyloid plaque distribution was performed according to the Consortium to Establish a Registry for AD (CERAD) semi-quantitative scoring criteria (176), while neurofibrillary tangle pathology was assessed in accordance with the Braak staging system (177). Tissues were homogenized in NP-40 lysis buffer (25mM Tris-HCl (pH 7.5), 150mM NaCl, 1mM EDTA, 1% NP-40, 5% Glycerol + protease + phosphatase inhibitors) using a bullet blender (170) followed by centrifugation at 10,000 x g for 10 mins at 4^{o}C to clear tissue debris. Immunoprecipitation was performed from 1 mg of brain homogenate from three independent AD

cases. Homogenates were first pre-cleared using 30µL of Protein A-Sepharose conjugated beads (Invitrogen #101041) rotating at 4^{o}C for one hour. GST purified rU1-70K fragments (4µg) were added independently to the pre-cleared homogenates. IP was performed using anti-Myc-tag (clone 9B11, Cell Signaling) in samples containing the rU1-70K proteins. An IgG control antibody (550339 BD Pharmigen) was used as negative control. Immunocomplexes were captured using Dynabeads Protein G magnetic beads (Invitrogen #1003D), which were washed 3 times using wash buffer (50mM Tris HCl, pH 8, 150 mM NaCl and 1% NP-40) followed by 5 min boiling in Laemmli sample buffer to elute bound proteins prior to western blot analysis. Samples were prepared for mass spectrometry as described above and analyzed on the Orbitrap Fusion mass spectrometer (ThermoFisher)(178). A total of 25 Tau (MAPT) and 18 amyloid precursor protein (APP) peptides were also added to an inclusion list to increase the likelihood of identification and quantification following database searching using MaxQuant as described above. To quantify Aβ levels in the samples, the ion intensities for two APP peptides, corresponding to residues 6–16 and residues 17–28 of the Aβ sequence, were summed for each individual sample(170). The intensities of Tau and APP was compared across conditions. ANOVA was used to determine significance with Tau and Aβ levels across samples in GraphPad Prism.

Western Blotting

Western blotting was performed according to standard procedures as reported previously (91, 94, 169). Samples in Laemmli sample buffer 8% glycerol, 2% SDS, 50mM tris pH 6.8, 3.25% β-mercaptoethanol were resolved by SDS-PAGE before an overnight wet transfer to 0.2 µm nitrocellulose membranes (Bio-Rad) or a semi-dry transfer using the iBlot2 system. Membranes were blocked with casein blocking buffer (Sigma B6429) and probed with primary antibodies (see materials) at a 1:1000 dilution overnight at 4 °C. Membranes were incubated with secondary antibodies conjugated to Alexa Fluor 680 (Invitrogen) or IRDye800 (Rockland) fluorophores for

one hour at room temperature. Images were captured using an Odyssey Infrared Imaging System (Li-Cor Biosciences) and band intensities were quantified using Odyssey imaging software.

Assessment of protein similarity to LC1 domain of U1-70K

The U1-70K LC1 protein similarity list was created using the Uniprot pblast feature (http://www.uniprot.org/blast/) using the following parameters: Target Database Human containing 160,363 entries (updated March 2015), E-threshold: 10, Matrix: Auto, Filtering: None, Gapped: Yes, Hits: 1000. The input blast entry was the LC1 domain of U1-70K (residues 231-308). The resulting list of proteins was subsequently filtered to remove un-reviewed entries, producing 255 proteins with E-values less than 0.005 and similarity to the LC1 domain of higher than 20 percent. The "biotin-isoxazole" list originates from a previous study (164) using the biotinylated isoxazole compound to precipitate proteins from HEK293 nuclear extracts. The "prion-like" list of RNA binding proteins originates from a previous work study using *in silico* methods to identify Q/N rich prion-like proteins (140). Finally, protein alignment was done using Clustal Omega multiple sequence alignment(https://www.ebi.ac.uk/Tools/msa/clustalo/).

2.6 ACKNOWLEDGEMENTS

Support was provided by 5R01AG053960, the Accelerating Medicine Partnership for AD (U01AG046161), the Emory Alzheimer's Disease Research Center (P50 AG025688), and the NINDS Emory Neuroscience Core (P30NS055077). N.T.S. is also supported in part by an Alzheimer's Association (ALZ), Alzheimer's Research UK (ARUK), The Michael J. Fox Foundation for Parkinson's Research (MJFF), and the Weston Brain Institute Biomarkers Across Neurodegenerative Diseases Grant (11060). I.B was funded by a pre-doctoral T32 NINDS training grant 3T32NS007480-16 and F31 NRSA grant F31NS09385902. We also acknowledge Dr. Measho Abreha for technical advice on brain homogenization and immunoprecipitation conditions.

Conflict of interest statement. None declared.

Author contributions: Conceptualization, I.B. J.J.L., A.I.L., and N.T.S; Methodology, I.B., D.M.D., E.B.D. and N.T.S.; Investigation, I.B. D.M.D., and E.B.D; Formal Analysis, I.B., D.M.D., E.B.D. and N.T.S., Writing – Original Draft, I.B., and N.T.S,; Writing – Review & Editing, I.B., E.B.D, D.M.D., M.G., J.J.L., A.I.L. and N.T.S.; Funding Acquisition, I.B., A.I.L. and N.T.S.; Resources, M.G., A.I.L. and N.T.S.; Supervision, N.T.S.

2.7 Figures and tables

Figure 2.1. The LC1 domain of U1-70K is necessary and sufficient for self-association in cells.

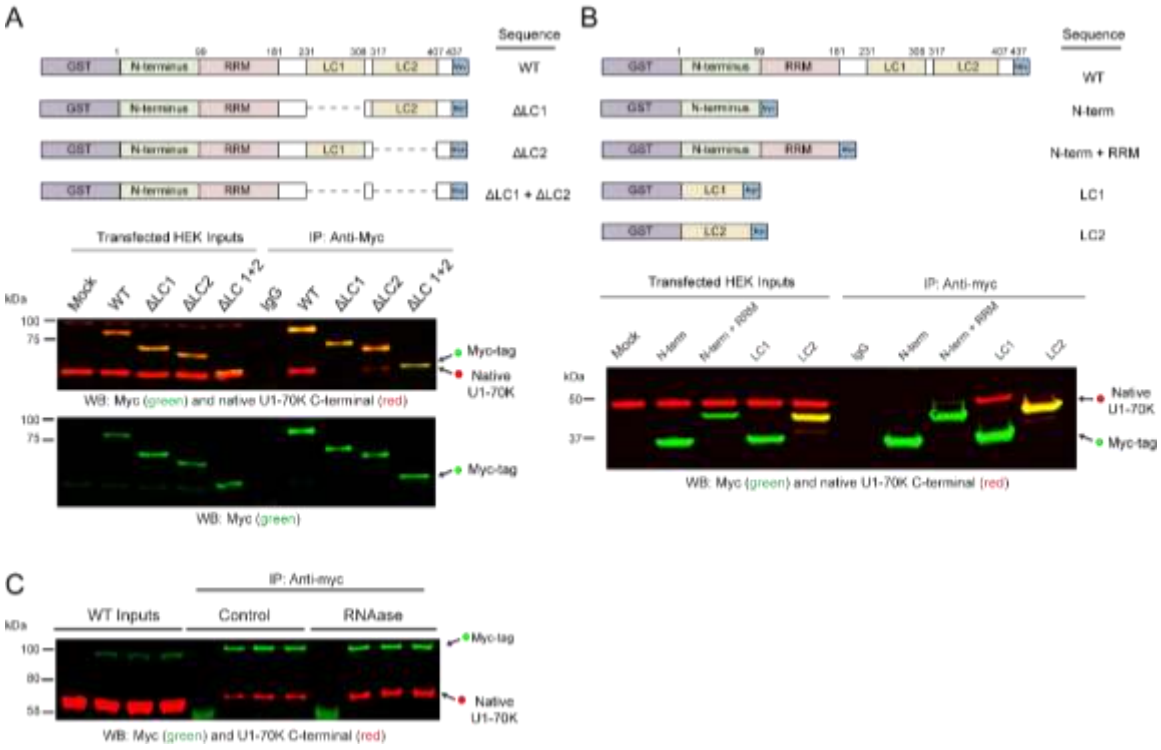


Figure 2.1. The LC1 domain of U1-70K is necessary and sufficient for self-association in cells. **A)** Full-length (WT) recombinant GST-fused and Myc-tagged U1-70K (rU1-70K) and serial deletions lacking one or both LC domains (ΔLC1 and ΔLC2, and ΔLC1+ΔLC2) were over-expressed in HEK293 cells and immunoprecipitated (IP) with anti-Myc antibodies. IP with a non-specific IgG was also performed from mock transfected cells as a negative control. Western blot for recombinant myc-tagged proteins (green) and native U1-70K (red) are shown for both the inputs and co-IPs (A, bottom panels). **B)** Full-length WT and rU1-70K truncations including the N-terminus (1-99 residues) alone, the N-terminus and RRM (1-181 residues), LC1 alone (231-308 residues) and the LC2 domain alone (317-407). IP with a non-specific IgG was also performed from mock transfected cells as a negative control. Western blot for recombinant myc-tagged proteins (green) and native U1-70K (red) are shown for both the inputs and co-IPs. **C)** WT rU1-70K was immunoprecipitated from untreated and RNAase (50ng/uL) treated lysates followed by western blot for the myc-tag recombinant protein (green) and native U1-70K (red).

Figure 2.2. The LC1 domain of U1-70K directly self-interacts and oligomerizes *in vitro*.

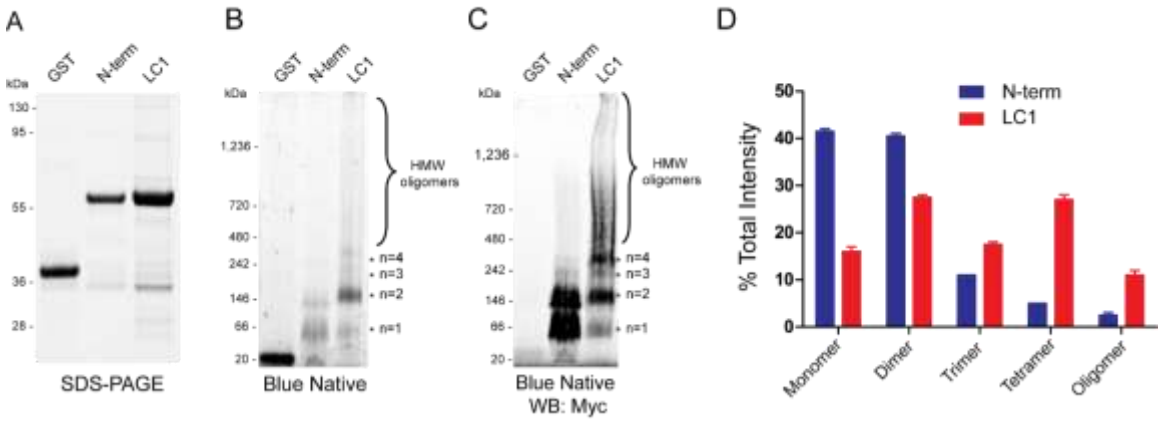


Figure 2.2 The LC1 domain of U1-70K directly self-interacts and oligomerizes *in vitro*. A) SDS-PAGE of GST alone, GST purified N-terminal domain and GST purified LC1 domain of rU1-70K. Both the LC1 and N-terminal domain have equivalent molecular weights (~65 kDa) while GST alone is (~20 kDa). **B)** Blue native gel polyacrylamide gel electrophoresis (BN-PAGE) of GST alone, the N-terminal domain and the LC1 domain of rU1-70K, respectively. The LC1 domain formed higher molecular weight species (*) consistent with dimers (~130 kDa), trimers (~195 kDa), tetramers (~260 kDa) and high-molecular weight (HMW) oligomers (>400 kDa). **C)** Western blot detection of blue native complexes using Myc antibodies. **D)** Densitometry of monomeric, dimeric, trimeric and HMW species of the N-terminal domain (blue) and LC1 domain (red) of rU1-70K. Each form is represented as the fraction of total signal in each sample analyzed in technical replicate (n=2). Error bars represent the standard deviation (s.d.).

Figure 2.3. The LC1 domain is necessary and sufficient for robust nuclear granule formation.

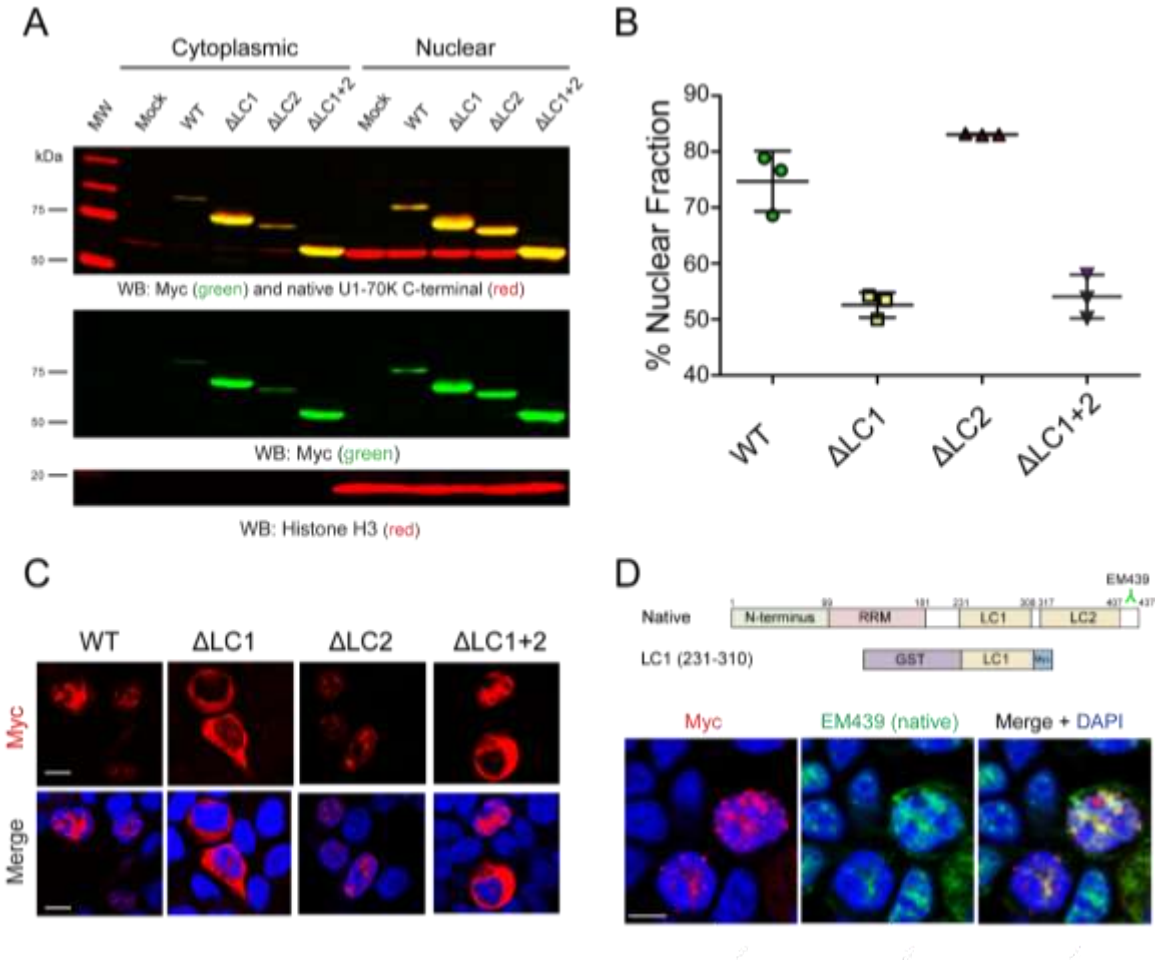


Figure 2.3. The LC1 domain is necessary and sufficient for robust nuclear granule formation. A) WT full-length rU1-70K or mutants lacking one or both LC domains were over-expressed in HEK293 cells. The cells were then fractionated into nuclear and cytoplasmic pools followed by western blot analysis for both recombinant myc-tagged proteins (green) and native U1-70K (red). Western blots for histone H3 (bottom panel) was used as a positive control in the nuclear fraction. **B)** Densitometry analysis was performed to calculate the levels of cytoplasmic and nuclear rU1-70K and mutants and % nuclear intensity for each rU1-70K protein is reported. Each experiment was performed in biological triplicate (n=3) with error bars representing the standard deviation (s.d.). Both the ΔLC1 and ΔLC1+2 rU1-70K fragments were significantly less nuclear than full length rU1-70K (** p-value<0.01). **C)** Immunocytochemistry for WT rU1-70K and mutants that lacked the LC1, LC2 or both LC domains was performed and visualized by confocal microscopy. Scale bar equates to 10 μM. **D)** Over-expression of the rU1-70K LC1 domain alone (red) resulted in robust nuclear granule formation and sequestration of native U1-70K (green). The EM439 detects and extreme C-terminal epitope not present in the LC1 rU1-70K fragment, which allows discrimination between the recombination protein and the native U1-70K. DAPI stained nuclei are shown in blue. Scale bar equates to 10 μM

Figure 2.4. Cluster dendrogram of rU1-70K interacting proteins quantified mass spectrometry analysis.

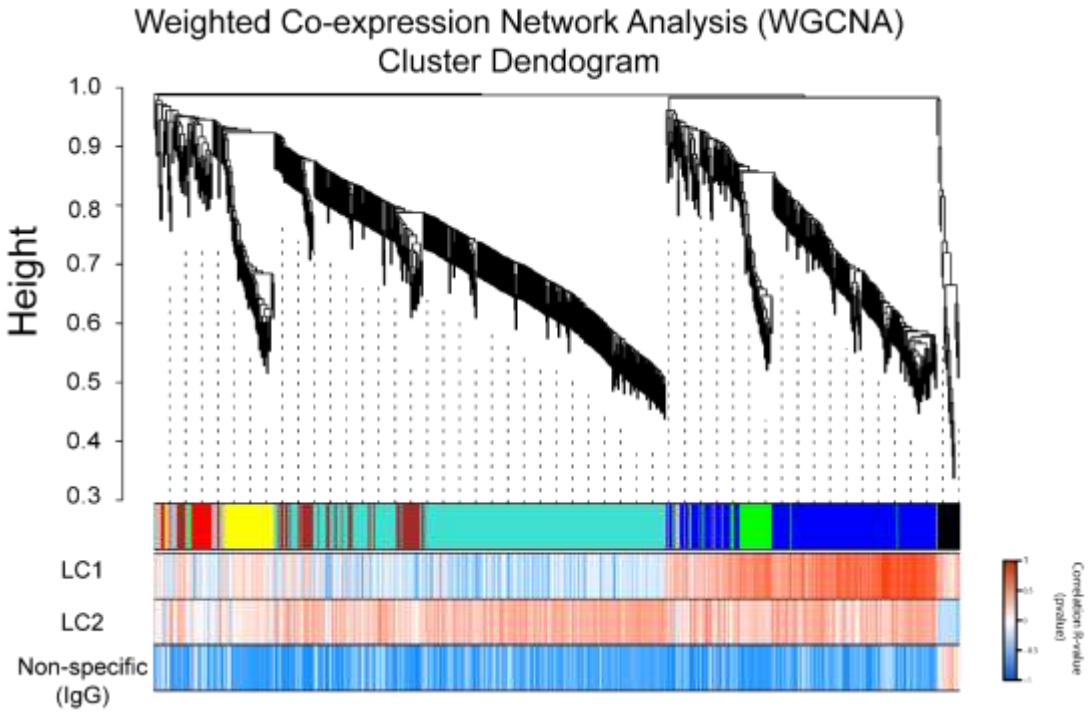


Figure 2.4 Cluster dendrogram of rU1-70K interacting proteins quantified mass spectrometry analysis. The WGCNA cluster dendrogram stratified rU1-70K interacting proteins (*n* = 716) into distinct modules (M1-7) defined by dendrogram branch cutting. The correlation of each protein to the specific traits (LC1, LC2, and IgG) are portrayed in the heat map

Figure 2.5. Correlation network analysis resolves distinct modules of U1-70K interacting proteins that differ in their association with the LC1 domain.

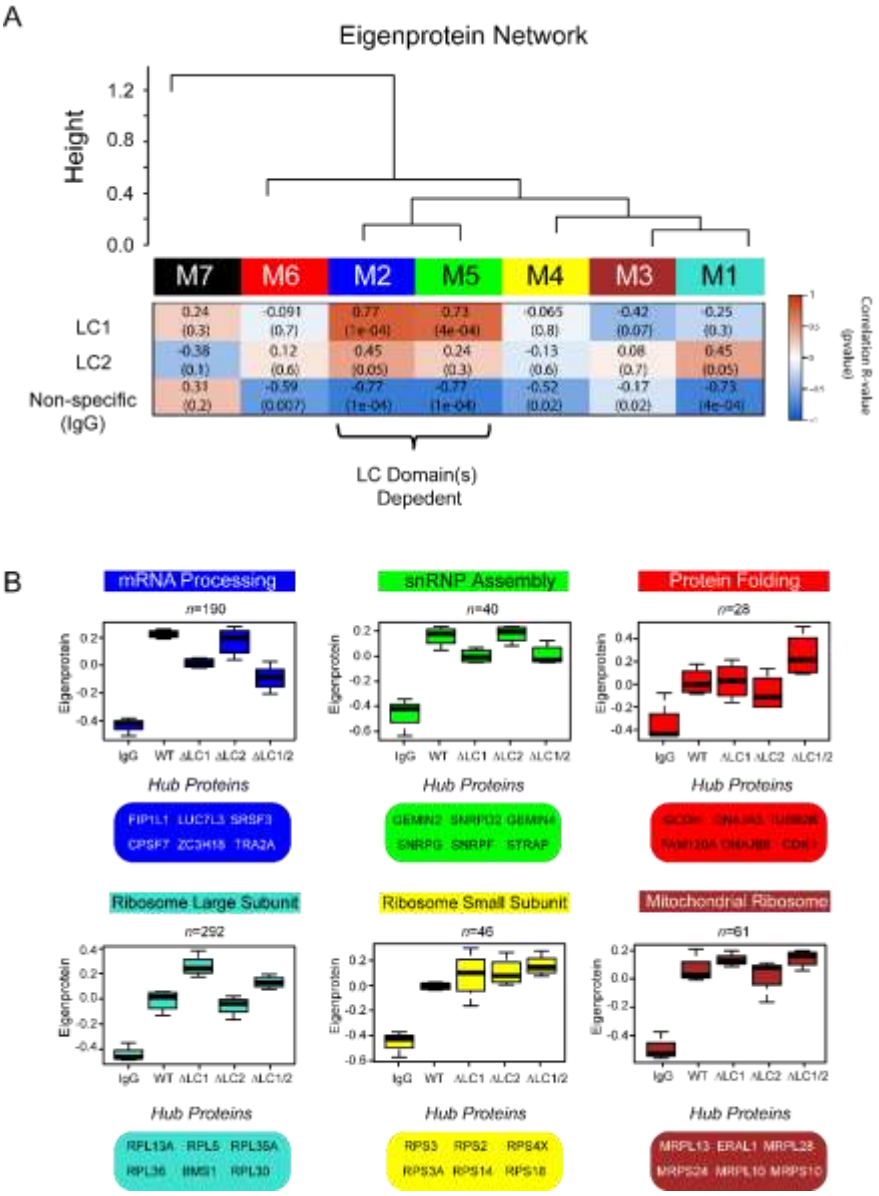


Figure 2.5. Correlation network analysis resolves distinct modules of U1-70K interacting proteins that differ in their association with the LC1 domain. A) WGCNA organized proteins defined by dendrogram branch cutting (**Fig 2.4**) measured across all co-IP samples (*n*=716) into modules (M1-M7) that represent clusters of proteins defined by their correlation to each other across the five co-IP conditions analyzed (IgG, WT rU1-70K and deletions ΔLC1, ΔLC2, and ΔLC1+ΔLC2). Listed in the heatmap are bicor correlations and p-values defining relationship between module eigenprotein level and rU1-70K protein (defined as 0-IgG, 1- ΔLC1, and 3-ΔLC2). **B)** Eigenproteins, which correspond to the first principal component of a given module and serve as a summary expression profile for all proteins within a module, are shown for 6 modules generated by WGCNA. Box plots with error bars beyond the 25th and 75th percentiles

are shown for all four groups (IgG, WT, ΔLC1, ΔLC2, and ΔLC1+2). Hub proteins for each of these modules are also highlighted below for each module.

Figure 2.6. Confirmation of U1-70K interacting partners that favor interactions via the LC1 domain.

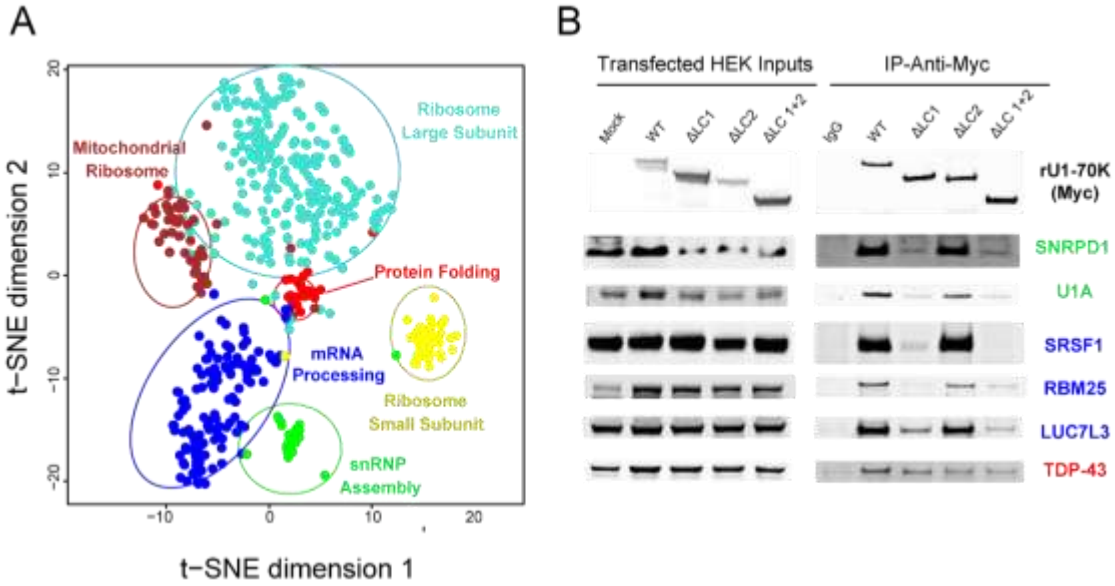
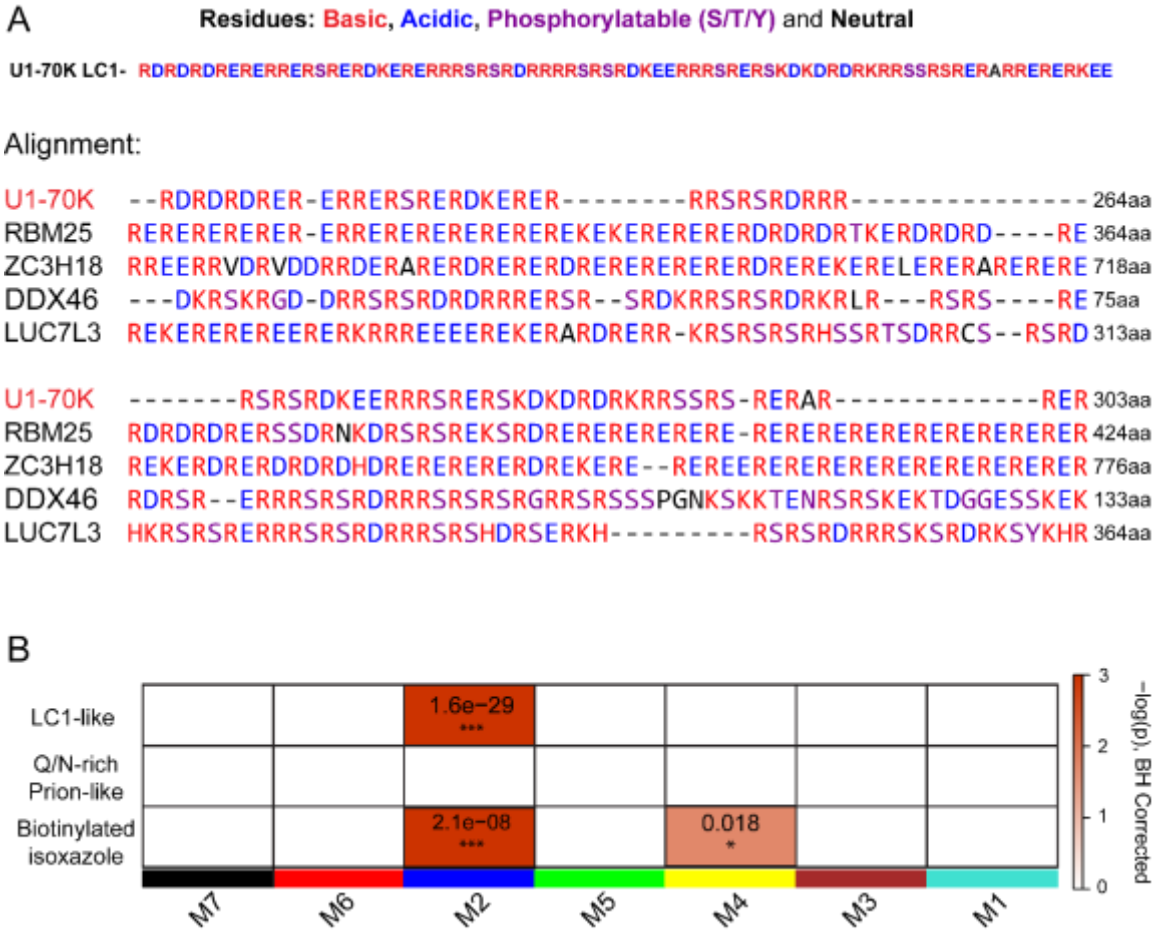


Figure 2.6. Confirmation of U1-70K interacting partners that favor interactions via the LC1 domain. A) To visualize the relationship between modules and validate the WGNCA results, the T-Distributed Stochastic Neighbor Embedding (tSNE) algorithm was used to map the relatedness of proteins with a kME score of 0.5 or greater, which is a measure of intramodular connectivity (kME), defined as the Pearson correlation between the expression pattern of a protein and the module eigenprotein. The tSNE analysis overlaid with module assignments determined by WGCNA allows for visualization of module relatedness, with distance between proteins representing similarity of co-enrichment, with the more similar clusters of proteins (related modules) in closer proximity to each other compared to dissimilar modules. Proteins with similar co-enrichment across the rU1-70K co-IPs are highly correlated to one another and are related to modules with distinct biological functions (**Table 2.1**). **B)** Western blot analysis to validate the proteomic results and module assignment were performed for select interactors of the snRNP assembly (SNRPD1 and U1A), mRNA processing (SRSF1, RBM25 and LUC7L3), and Protein folding (TDP-43) modules following co-IP across the five experimental conditions (IgG, WT rU1-70K and deletions ΔLC1, ΔLC2, and ΔLC1+ΔLC2).

Figure 2.7. The mRNA processing module is enriched with structurally similar RNA-



binding proteins harboring U1-70K LC1-like domains.

Figure 2.7. The mRNA processing module is enriched with structurally similar RNA-binding proteins harboring U1-70K LC1-like domains. A) The LC1 domain of U1-70K (residues 231-308) contains highly repetitive complementary basic (R/K) and acidic (D/E) residues. A list of 255 proteins that shared greater than 20% similarity to the LC1 domain of U1-70K (E-values less than 0.005) was created using the Uniprot protein Blast feature. Using clustal omega, an alignment was performed on the U1-70K LC1 domain and the four most structurally similar proteins to highlight the mixed charge nature of their sequence. **B)** A one-tailed Fisher's exact test (FET) was used to assess structural overlap of LC1-like mixed charge proteins from Blast analysis with module membership for U1-70K interacting partners (upper panel). FET analysis was also performed using prion-like RNA binding proteins (middle panel) or proteins that were precipitated from nuclear extracts using biotin-isoxazole compound (bottom panel). Benjamin Hochberg corrected p-values (to control FDR for multiple comparisons) for the module enrichment is highlighted. Significance is demonstrated by the color scales, which go from 0 (white) to 3 (red), representing -log(p).

Figure 2.8. The mixed charge domains in LUC7L3 and RBM25 are necessary for reciprocal interactions with U1-70K and nuclear RNA granule assembly.

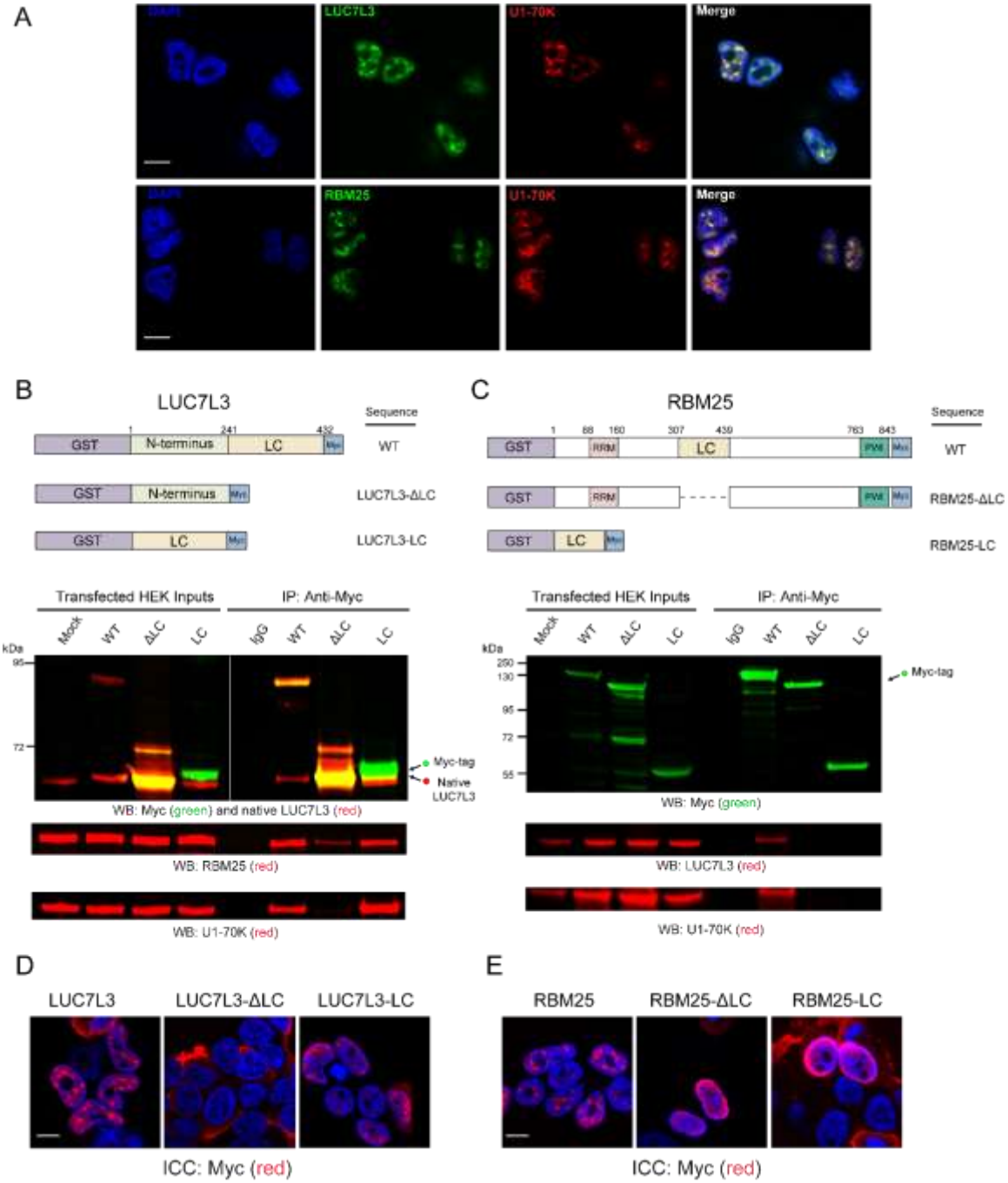


Figure 2.8. The mixed charge domains in LUC7L3 and RBM25 are necessary for reciprocal interactions with U1-70K and nuclear RNA granule assembly. A) Immunocytochemistry (ICC), to assess the co-localization of native U1-70K (green) with RBM25 (red) or LUC7L3 (red). DAPI stained nuclei are shown in blue. Scale bar equates to 10 μM. **B)** Full-length (WT) recombinant GST-fused and Myc-tagged LUC7L3 (rLUC7L3) and mutants lacking the mixed

charge (MC) domain or the MC domain alone (upper panel) were over-expressed in HEK293 cells and immunoprecipitated (IP) with anti-Myc antibodies. IP with a non-specific IgG was also performed from mock transfected cells as a negative control. Western blot for recombinant myc-tagged proteins (green) and native LUC7L3 (red) are shown for both the inputs and co-IPs (bottom panels). Membranes were also re-probed for native U1-70K (red) or RBM25 (red). **C)** Full-length (WT) recombinant GST-fused and Myc-tagged RBM25 (rRBM25) and mutants lacking the MC domain or the MC domain alone (upper panel) were over-expressed in HEK293 cells and IP with anti-Myc antibodies. IP with a non-specific IgG was also performed from mock transfected cells as a negative control. Western blot for recombinant myc-tagged proteins (green) and native U1-70K (red) or LUC7L3 (red) are shown for both the inputs and co-IPs (bottom panels). **D)** Immunocytochemistry for full-length rLUC7L3, a mutant lacking the MC domain (LCU7L3-ΔMC) and the MC domain alone (LUC7L3-MC) were expressed in HEK293 cells and visualized by confocal microscopy. **E)** Immunocytochemistry for rRBM25, a mutant lacking the MC domain (RBM25-ΔMC) and the MC domain alone (RBM25-MC) were expressed in HEK293 cells and visualized by confocal microscopy. DAPI was used to visualize nuclei (blue), scale bars = 10μM for both panels D and E.

Figure 2.9. RNA binding proteins with mixed charge domains have increased insolubility in AD brain.

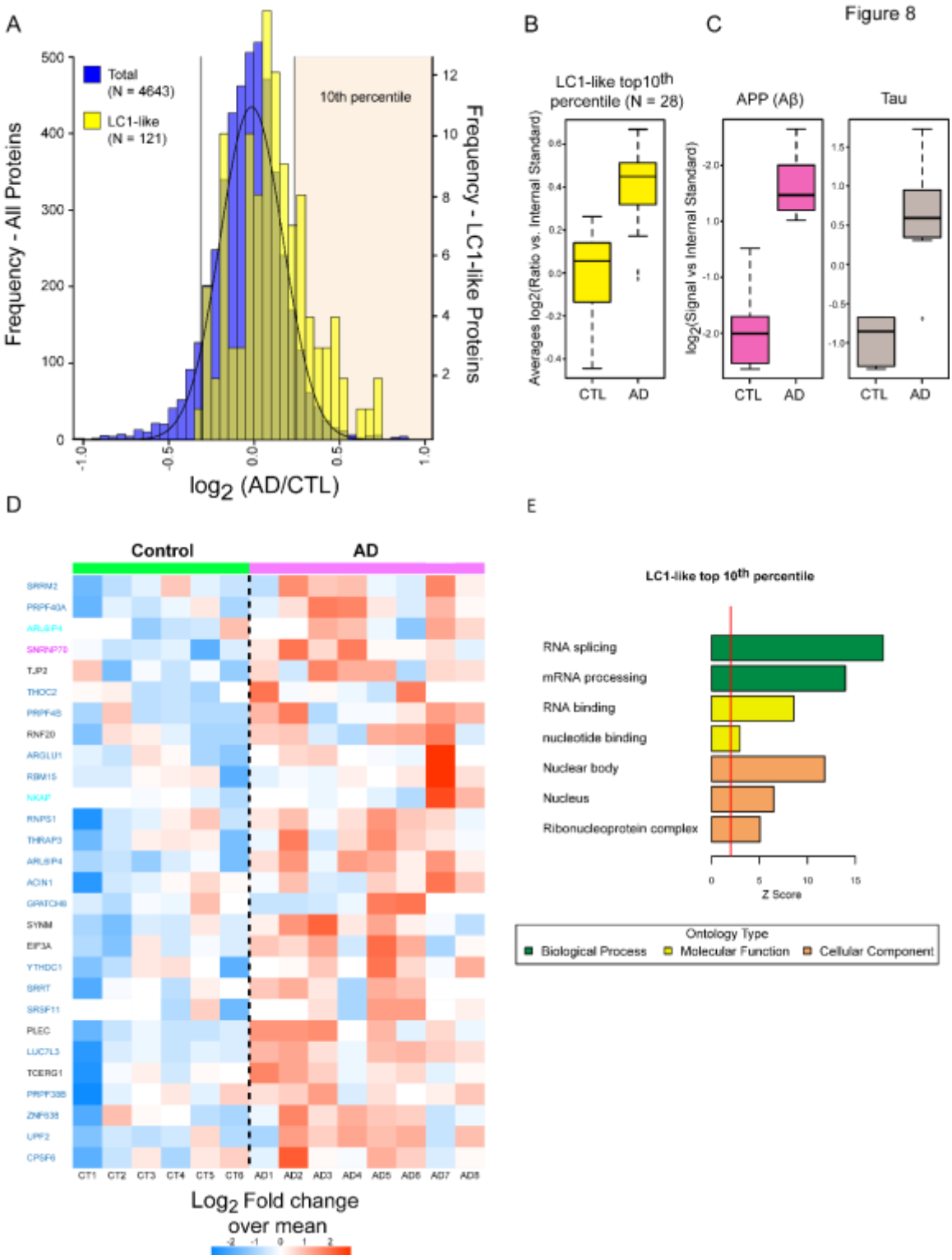


Figure 2.9. RNA binding proteins with mixed charge domains have increased insolubility in AD brain. **A)** Histogram of average $\log_2$ ratios (AD/control) from the control (*n*=6) and AD (*n*=8) brain detergent insoluble fractions. Protein ratios for all pairwise comparisons (i.e., control vs. AD) were converted into $\log_2$ values, and the resulting histogram fit to a normal Gaussian

distribution. Compared to the normal distribution of all proteins in the AD insoluble proteome (purple histogram), quantified mixed charge proteins (*n* =112 yellow histogram) showed a global shift towards insolubility in AD. The 28 proteins with MC domains that fell above the ninetieth percentile (top 10% percentile) of the total population distribution are significantly overrepresented (Fisher exact p-value 2.0e-09). **B)** Box plot of the mixed charge proteins that fell into the top tenth percentile. **C)** APP (Aβ) and MAPT (Tau) protein levels. The central bar depicts mean and box edges indicate 25th and 75th percentiles, with whiskers extending to the 5th and 95th percentiles, excluding outlier measurements. **D)** Heat map representing the fold-change over the mean of mixed charge proteins in the top tenth percentile across the control and AD cases. Gene symbols are highlighted by their module color in U1-70K interactome (Black=not in a module). **E)** Gene ontology (GO) analysis of the 28-enriched mixed charge domain proteins highlights functions in RNA binding and processing. Significant over-representation of the ontology term is reflected with a $Z>1.96$, which is equivalent to $p<0.05$ (above red line).

Figure 2.10. The mixed charge LC1 domain of U1-70K interacts specifically with Tau in AD brain

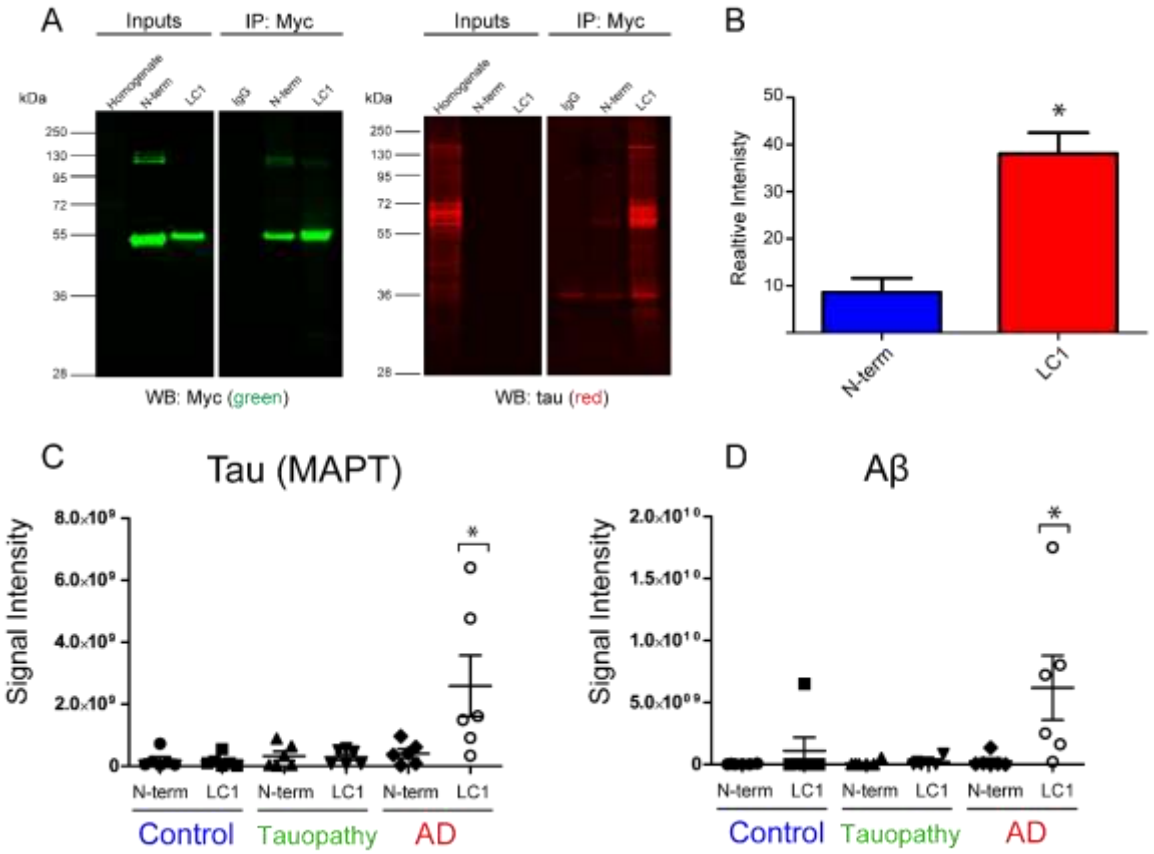


Figure 2.10. The mixed charge LC1 domain of U1-70K interacts specifically with Tau in AD brain. A) GST purified N-terminal or the mixed charge LC1 domain (4 µg) of rU1-70K were added separately to AD brain homogenates and immunoprecipitated with anti-Myc antibodies. IP with a non-specific IgG was also performed as a negative control. Inputs and immunoprecipitates were analyzed by Western blot using anti-Tau antibodies (red) and myc antibodies (green). **B)** The LC1 domain interacted with significantly higher levels of Tau from AD brain than the N-terminal domain (t-test one-tail p value = 0.0156). The experiment was done in biological triplicate (*n*=3) from independent AD cases **(Table 2.2)**. **C and D**) GST purified N-terminal or the mixed charge LC1 domain (4 µg) of rU1-70K were added separately brain homogenates from control (*n*=6), AD (*n*=6) and tauopathy (*n*=6) brain tissue and immunoprecipitated with anti-Myc antibodies followed by mass spectrometry (MS) analysis. Label free quantification was used to determine the signal intensities (y-axis) of Tau and Aβ across co-IP conditions (LC1 or N-term)) and brain homogenates. Statistical significance for Tau and Aβ interactions (LC1 and N-term) was determined by ANOVA (*p value <0.05).

Table 2.1. U1-70K Protein-Protein Interaction Network Generates Modules Enriched with Specific Gene Ontology (GO) terms

Module Color	Ontology type	Top GO Terms	Fisher Exact P-value
Turquoise (*n*=292)	*Biological Process*	ncRNA processing	2.07E-11
		rRNA metabolic process	4.19E-10
		endocrine pancreas development	9.99E-09
	Molecular Function	nucleic acid binding	1.8E-08
		structural constituent of ribosome	1.28E-20
		protein-DNA loading ATPase activity	0.00153
	Cellular Component	cytosolic large ribosomal subunit	2.66E-17
		Nucleolus	2.16E-10
		Intracellular	1.58E-05
Blue (*n*=190)	*Biological Process*	mRNA processing	1.01E-44
		RNA splicing	6.82E-41
		mRNA export from nucleus	2.14E-16
	Molecular Function	RNA binding	1.97E-08
		RS domain binding	6.57E-06
		nucleotide binding	2.49E-05
	Cellular Component	nuclear speck	3.55E-14
		spliceosomal complex	1.29E-12
		Nucleus	3.26E-15
Brown (*n*=61)	*Biological Process*	Translation	1.82E-15
	Molecular Function	structural constituent of ribosome	1.28E-20
		translation regulator activity	0.002275
	Cellular Component	Mitochondrion	9.67E-34
		Ribosome	1.91E-28
		mitochondrial large ribosomal subunit	1.51E-11
Yellow (*n*=46)	*Biological Process*	endocrine pancreas development	9.99E-09
		viral transcription	9.99E-09
		viral infectious cycle	4.61E-08
	Molecular Function	structural constituent of ribosome	1.28E-20
		mRNA binding	0.005748
	Cellular Component	cytosolic small ribosomal subunit	1.78E-44
		Ribosome	1.91E-28
		Intracellular	1.58E-05
Green (*n*=40)	*Biological Process*	spliceosomal snRNP assembly	2.67E-20
		regulation of cyclin-dependent protein kinase activity	1.4E-06
		spliceosome assembly	5.74E-06
	Molecular Function	snRNA binding	6.37E-08
	Cellular Component	small nuclear ribonucleoprotein complex	5.26E-09
		U12-type spliceosomal complex	4.85E-08
		Cajal body	1.05E-07
Red (*n*=20)	*Biological Process*	protein folding	3.65E-06
		response to abiotic stimulus	2.37E-05
		oxidation-reduction process	0.000248
	Molecular Function	heat shock protein binding	7.28E-07
		purine ribonucleoside triphosphate binding	3.21E-05
		catalytic activity	0.000225
	Cellular Component	Microtubule	0.000258
		Membrane	2.47E-05
		intrinsic to membrane	0.016518

Table 2.2 Case information

Case Number	Diagonis	BRAAK Score	Neocortical Neuritic Plaque Frequency	PMI (hr)	Age at Death	Race/Sex	Western	Mass Spec
E05-04	AD	VI	High	5	64	wf	x	x
E04-172	AD	V	High	6	87	wf	x	x
OS01-02	AD	VI	High	5.5	69	wf	x	x
0SO2-12	AD							x
E08-53	AD	VI	High	8	78	wf		x
E06-155	AD	VI	High	6.5	67	wm		x
0SOO-33	CBD							x
0SOO-13	CBD							x
0SO2-284	CBD							x
0S93-39	CBD							x
0SOO-35	CBD							x
0S99-12	CBD							x
A86-50	CTL							x
EI6-45	CTL							x
A86-46	CTL	0	None	N/A	65	wm		x
E08-101	CTL			11.5	78	wf		x
A93-03	CTL			4.5	70	hm		x
OSO3-299	CTL	II	None	6	65	wm		x

Table 2.2 Case information. The information for cases used in this study.

CHAPTER 3 DISCUSSION

3.0 Summary of Presented Studies and Further Perspectives About the role of Mixed Charged RNA binding Proteins in Alzheimer's Disease

These collective works aim to better understand the mechanisms of U1-70K aggregation, downstream effects of Aβ, and basic principles of ribonuclear particle assembly (RNP). In chapter two, I aimed to understand how U1-70K self-assembles *in vitro*, and forms granules *in vivo*. It has been hypothesized that endogenous processes (i.e. oligomerization and granules assembly) are precursors to pathological aggregates and occur along parallel mechanisms (93). Through understanding U1-70K self-assembly *in vitro*, and granule localization *in vivo* I have revealed mechanistic insight into U1-70K aggregation. In addition, I demonstrated a possible physical link between Aβ and tau aggregation. In summary, based on these collective findings, our understanding of U1-70K's role in AD pathogenesis has been expanded.

3.1.1 Biological function of Mixed Charge domains

I propose that mixed charge (MC) proteins be considered a proteins class, due to their related biological function, and shared primary structure. For example, I provide evidence to support the function of the LC1 of U1-70K, and more broadly other mixed charge (MC) domains,

in nuclear RNA granule formation. Thus, the reciprocal interactions of MC domains could form the “glue” that drives granule assembly. It has been previously proposed that MC domains, such as those found in U1-70K, RBM25 and LUC7L3, self-assemble through the formation of polar zippers (149). Polar zippers are a structural motif proposed by Perutz to facilitate protein-protein interactions between highly charged domains though electrostatic interactions. Perutz hypothesized that polar zippers contribute to RBP complex assembly and protein aggregation (149). The novelty of my work is that it provides biochemical evidence of this theorized "polar zipper".

The strong conservation of the dipeptide residues repeats stands in sharp contrast to the poor conservation of their codons across species, thus implying that the MC nature cannot be entirely accounted for by rounds of genetic duplications (149). This observation indicates the existence of a strong selective pressure responsible for the conservation of the MC domains indicating that the MC domains have an essential function in proteins. The interaction between MC domains could be crucial to facilitating the complicated and multifaceted function of RNA processing proteins (148, 179). The presence of distinct LC domains like the LC2 (317-407 residues) in U1-70K may further refine the dynamics of this process. For example, U1-70K interacts with both the spliceosome and members of the polyadenylation complex, including FIP1L1. Analogous to U1-70K, FIP1L1 harbors a MC domain, and is the top hub protein in the blue module (180). Thus, the presence of MC domains in U1-70K enables physical, if not also functional, crosstalk between the role of U1-70K in 5’-splice site recognition and the polyadenylation complex in mRNA processing.

Additional information was revealed via the protein-protein interaction network obtained in this study including interactions between U1-70K and 1) translation machinery, 2) mitochondrial components and 3) the exon-junction complex. Currently, U1-70K does not have a known role in translation, but due to the number of U1-70K interacting ribosome components and translation factors I would predict that U1-70K does has a role in translation. U1-70K was found

to interact with several components of the mitochondrial ribosome. A recent study found that U1-70K and other splicing factors are found in mitochondria (181). This opens up the possibility that U1-70K interactions regulate splicing inside the mitochondria. Splicing in non-canonical locations has been observed in the dendrites of neurons (182, 183). This is interesting given mitochondria dysfunction is a well documented event during AD (184). Our interaction data indicates that U1-70K interacts with several components of the exon-junction complex, which shuttles mRNA out of the nucleus (185). We found a minority pool of U1-70K in the cytoplasm (**Fig. 2.3A-B**). This suggest that U1-70K is not purely a nuclear protein but has functions outside the nucleus. This in combination with U1-70K-ribosome interactions paints a picture where U1-70K aids in splicing mRNA, then shuttles out of the nucleus with the exon-junction complex, where it associates with active ribosomes. Then shuttles back into the nucleus for a new round of splicing. It could be a disruption in the shuttling that triggers U1-70K aggregation in AD.

The normally nuclear U1-70K is found mislocalized to cytoplasmic Tau-immunoreactive neurofibrillary aggregates in AD neurons (75), which may contribute to a loss of spliceosome function given recently identified RNA splicing deficits in the disease (144). Mislocalization of other RBPs contributes to neurodegenerative disease (186, 187). Here I show that the LC1 domain is important for nuclear localization of U1-70K, supporting a link between aberrant LC1 interactions and mislocalization of U1-70K. Our findings also shed light on previous studies in which the C-terminal domain (residues 161-437) of U1-70K was found sufficient for nuclear localization (151, 188). Moreover, our ΔLC1+2 mutant protein mimicked previous findings for the 1-199 U1-70K C-terminal truncation, wherein expression of this N-terminal fragment localized to the nucleus, but not to granules. It is likely that the nuclear localization sequence within the LC1 domain was missed by earlier studies due to the selected sites of truncation, as the LC1 domain was never expressed in its entirety (188).

3.1.2 Parallels between RNA binding protein aggregation events in AD and ALS

I show that U1-70K and other MC proteins share many of the same properties as Q/N-rich prion-like proteins, despite their difference in primary sequences. These properties include the ability to self-assemble into high molecular weight oligomers, form nuclear granules in cells, and promote aggregation. The formation of RNA granules has been viewed as an intermediary step towards protein aggregation (93, 96, 136) and our observations place U1-70K and other mixed charge proteins among prion-like RBPs, such as TDP-43, FUS, hnRNPA1, and TIA-1 that form granules and aggregate in neurodegenerative disease (90, 105, 143, 189). Thus, the ability of the LC1 and other MC domains to self-interact poises them for pathological aggregation in neurodegenerative diseases, which is consistent with their increased insolubility in AD.

The theory that RNA granules proteins are poised to aggregate in neurodegenerative disease had previously not included AD (93, 190). The theory originally grew out of observations made studying ALS and the related disease like FTLD (101, 102, 189). My work suggests that U1-70K and other MC proteins in AD aggregate through a different, but parallel mechanism to that of TDP-43, FUS, and hnRNPA1. This is the first time MC domains have been biochemically linked to self-assembly, granule formation, and aggregation, mirroring what has been observed for polar non-charged LC domains. This opens up the possible of other LC domain subtypes aggregating in other neurodegenerative diseases. Despite their different amino acid composition, a uniting feature of the RBPs listed above is their aggregation in age related neurodegenerative diseases. This suggests that RBP solubility is particularly susceptible to age related defects in cellular homeostasis.

Stress granules are enriched with proteins containing homologous LC domains largely composed of polar noncharged amino acids (116, 140). It is theorized that stress granules are assembled though reciprocal interactions between these various LC domains (93, 118). I provide evidence that proteins containing MC domains co-localize to nuclear granules and that MC domains reciprocally interact to form nuclear granules mechanistically mirroring that of stress granules. Self-association of other LC domain subtypes could be a broadly used feature to

establish cellular organization. It is likely that other examples of homologous LC domains forming membrane free organelle will be uncovered in the future.

3.1.3 Implication of U1-70K interactions with Aβ and Tau

Both U1-70K and Tau co-localize to neurofibrillary tangles in late-onset sporadic and familial cases of AD, but not in other tauopathies (74, 75, 91, 168, 191). U1-70K also aggregates in preclinical or asymptomatic AD cases (191), defined by significant Aβ deposition in the absence of significant cortical tau deposition and cognitive impairment. Although mechanisms underlying the relationship between Aβ, Tau, and U1-70K aggregation are not well established our co-immunoprecipitation findings using the MC domain of U1-70K as bait indicates that Aβ either directly or indirectly influences U1-70K interactions with Tau in brain. Although Aβ is predominantly found aggregated in the extracellular space within senile plaques there is evidence of intraneuronal Aβ in both mouse models (192) and humans with AD (193), which I hypothesize could directly mediate interactions between U1-70K, Tau, and other MC RBPs leading to pathological aggregations (**Fig. 10**).

Alternatively, downstream signaling pathways specific to AD could induce Tau post-translational modifications (PTMs) and tertiary structures that favor interactions with U1-70K and MC proteins. For example, Tau undergoes LLPS via electrostatic static interactions *in vitro*, referred to as coacervation,

(133, 134, 194). Tau LLPS is mediated by the Tau microtubule binding repeats (residues 244-369) (195), which is notable as a recent study examining the physical structure of Tau filaments in AD brain revealed an exposed mixed charge region within the tertiary structure of Tau, comprised of residues 338-358 in the microtubule binding repeat domain (196). Thus, it is tempting to speculate that pathological Tau in AD may behave like other mixed charge RBPs and sequester U1-70K to neurofibrillary tangles or vice versa. Specifically, I propose that this MC surface in pathological Tau mediates physical interactions between Tau and the U1-70K LC1 domain (**Fig. 3.1**). Thus, unique structural conformations of Tau in AD may contribute to its

specific interactions with the MC domain of rU1-70K we observed. Furthermore, Tau-U1-70K hetero-oligomers may have unique aggregation propensity, though additional determinants of aggregation may reside in the cytoplasm including RNA (168).

Aberrations in Aβ metabolism is a well documented trigger of AD pathogenesis (6, 11, 14) and the subsequent aggregation of intraneuronal Tau is also a hallmark of AD (17, 18, 53). However currently the events that connect Aβ accumulation and tau tangle formation are not well understood (55). Here, I propose a hypothesis that U1-70K aggregation is an intermediary step between Aβ accumulation and tau aggregation, both temporally and mechanistically. U1-70K interacts with both Aβ and tau, and aggregates before Tau. This strongly suggest that U1-70K could act as bridge connecting these two pathologies.

In summary, I have identified novel functional roles for MC domains in protein-protein interactions, nuclear localization, granule formation, and pathological aggregation. I show similarities between MC domains and Q/N-rich domains found in RBPs that aggregate in neurodegenerative diseases. I also demonstrate how a weighted protein-protein interaction network analysis can be used to resolve biologically and structurally distinct complexes. Notably, RBPs with MC domains demonstrated elevated insolubility in AD brain and the MC domain of U1-70K was sufficient to specifically interact with both Aβ and pathological Tau specifically in AD brain homogenates. This supports a hypothesis that MC structural motifs on U1-70K and related RBPs could mediate cooperative interactions with β-amyloid and Tau in AD.

3.2 A Comprehensive Model for AD Pathogenesis: A role for mixed charge RNA binding proteins

I put forth a mechanism whereby MC domains interact though electrostatic interactions to oligomerize, to form granules, and under certain conditions aggregate (**Fig. 3.1**). I propose a hypothesis in which U1-70K and perhaps other MC RNA-binding proteins act as the bridge between Aβ and Tau. Due to a combination of genetic and environmental factors, Aβ metabolism is disrupted and Aβ levels in the brain increase (6). At some point related to concentration,

localization, or other factors, intraneuronal Aβ tightly associates, directly or indirectly, with U1-70K via the LC1 domain. This results in the dysfunctional assembly and maintenance of MC containing complexes, including the U1snRNP. The U1snRNP and other RNPs components begin to accumulate and aggregate. Due to the MC nature of the LC1 domain, U1-70K acts as a scaffold for other MC proteins to aggregate onto, further driving aggregation. This together results in RNA processing dysfunction, perhaps acting as a positive feedback loop towards Aβ accumulation (74). Structural alterations in Tau specific to AD, due to PTMs and or conformation changes allow Tau to interact with U1-70K and other MC proteins. This facilitates the aggregation of Tau and MC proteins, which finally leads to synapse loss, cell death, and ultimately dementia. See (**Fig.3.1**) for illustration of model.

3.3 Power of Network Approaches

A common strategy to assigning function to an unclassified protein is examining the function of its interactors. This approach is based on the concept that a majority of interacting protein pairs share at least one function (197). This idea of "guilt by association" can be used to predict the function of proteins. When a protein's co-enrichment correlates with that of a known protein complex it is likely a member of that complex. The power of WGCNA and other network approaches is that novel members of a protein complexes can be identified in this manner. This approach can be widely used to globally identify protein interaction networks across various systems and conditions. This technique could also be used to unbiasedly discover new complexes and how they are altered in control and disease state. In addition, I think as a scientific community we need to reexamine previously described protein interaction networks, due to the common frequency in which proteins have secondary roles that are yet to be described. For instance, at the time of our data collection, the 7SK snRNP was thought to be purely an RNA polymerase II regulator (198). However, our data placed the 7SK snRNP components in the green module with other Cajal body members. Cajal bodies are site of U1 snRNP maturation (161).

Recently, the 7SK snRNP was discovered to reside in Cajal bodies and to regulate U1 snRNA transcription (160). This example illustrates the ability of WGCNA to predict protein function.

3.4 Limitations

A general shortcoming of this research was the use of over-expressed recombinant protein for most assays. Over-expression can lead to artifacts unrelated to the endogenous biology wished to studied (199). In addition the use of GST and myc tags could alter the characteristics of the proteins studied, in unpredictable manners (200). These concerns are partially alleviated by the observation that rU1-70K, rLUC7L3 and rRBM25 maintained their endogenous localization in nuclear granules. In addition rU1-70K maintained interactions with well documented interactors, U1A, U1C, and SRSF1 (81). Together this suggest that the recombinant proteins used in this study where able to recapitulate endogenous function and properties.

The definition I used for MC proteins is imperfect and improvements can be made. I define MC by homology based strictly on the primary sequence of the LC1 domain of U1-70K. Defining MC protein by their positively and negatively charged residue density would be more direct and accurate. However, the size (number of residues) of MC domains necessary to facilitate protein-protein interactions still needs to be resolved.

Finally, a limitation of the experiments used to assess the interaction between U1-70K, Aβ and tau is that it cannot differentiate between direct and indirect interactions. Below I describe several experiments that examine this question.

3.5 Remaining Questions and Future Studies

My graduate research leaves many remaining questions. It is my recommendation that an area of future research focus be the relationship between Aβ, U1-70K, and Tau. Observations suggest that U1-70K aggregation is downstream of Aβ metabolism (76, 87). In the insoluble fraction of human brain, U1-70K is the most tightly correlated protein with Aβ across all AD disease states (76). I also show that the LC1 domain of U1-70K and Aβ interact. Thus, aggregation of U1-70K could be influenced by direct and/or indirect interactions with Aβ, the

next section focuses on experiment to probe the relationship between U1-70K and Aβ. These experiments could reveal novel mechanisms fundamental to AD progression.

3.5.1 Aβ specificity

Future studies should aim to validate U1-70K and Aβ interactions. Below is a list of experiments I think will validate and expand upon previous results. First, the co-immunoprecipitation mass spectrometry experiment could not distinguish between direct and indirect interactions, as mediator proteins could bridge interactions between the LC1 domain and Aβ. To test for direct interactions synthetic Aβ peptides could be added to a solution containing purified LC1 domain. Then a co-IP mass spectrometry experiment could be performed to test for Aβ binding. In addition, chemical cross linking before immunoprecipitation could distinguish between direct and indirect interactions in a complex mixture, such as brain homogenate (91). Using this experimental design, the LC1's specificity to Aβ subtypes, $A\beta_{40}$ and $A\beta_{42}$ could also be examined. In addition, different incubation times could be used to enrich for oligomers versus fibrils of Aβ to determine relative binding strength. These experiments could be repeated with purified WT U1-70K and ΔLC1 mutants, to determine if the LC1 is necessary for binding. Finally, purified LUC7L3, RBM25, or various other MC proteins could be used in the various experiments listed above to determine if Aβ association is a general property of mixed charge domains or specific to the LC1.

The characteristics of the theorized Aβ-U1-70K-Tau complex are largely undescribed. Currently it is unknown if Tau-U1-70K interactions are dependent on the presence of Aβ. Brain homogenate from AD brain could be immunodepleted of Aβ. From this depleted homogenate an indirect IP could be performed to determine if Tau-LC1 interactions are retained.

3.5.2 Correlating Aβ-LC1 Binding to Disease Progression

The progression of AD happens in stages with buildup of Aβ happening early on. I observed the LC1 domain binding Aβ in one of the controls (**Fig. 2.10D**). This experiment should be repeated with more high plaque burden Asymptomatic AD cases. This experiment could

provide evidence for U1-70K interactions with Aβ in the early stages of disease, and in the absence of Tau. Results should then be correlated to features such as U1-70K insolubility and total, soluble, and intraneuronal Aβ levels.

The amyloid precursor protein (APP) is localized largely to the cellular membrane (6). The APP cleavage product, Aβ, is believed to be mostly produced and released in the extracellular space (6). Yet, evidence indicates that intracellular and nuclear pools of Aβ exist in cells expressing APP (201, 202). Quantification and description of intercellular Aβ is still in the early stages. Further understanding of intercellular Aβ could aid in understanding U1-70K aggregation. Intercellular pools of Aβ in neurons could explain the cell type specificity of U1-70K aggregation.

3.5.3 Determining the Biological Purpose of U1-70K-Ribosome Interactions

Our IP mass spectrometry data suggested that U1-70K interacts with ribosomes. This is interesting for two reasons: 1) U1-70K previously had no known role in translation or ribosome biogenesis 2) There is increasing evidence indicating that Tau interacts with ribosomes (203). Immunoprecipitation of Tau enriches for several ribosome components (203). Tau has been observed to interact more closely with ribosomes in AD than control resulting in decreased translation (204). Under some conditions Tau locates to ribosomal DNA loci (205). Tau can regulate the transcription of ribosomal RNA (205). Ribosomal deficiencies appear in regions with tau pathology (206). A better understanding of the relationship between U1-70K and ribosomes could provide biological context to U1-70K tau interactions.

At this time, we cannot distinguish between U1-70K interacting with ribosomes during biogenesis or translation or both. To examine this question, active ribosomes from cells could be purified then analyzed for U1-70K co-enrichment. Association with active ribosome would suggest U1-70K is involved in translation. Conversely enrichment of pre-assembled ribosomes could be done then tested for co-enrichment of U1-70K. This could provide evidence that U1-70K plays a role in ribosome biogenesis. Since mitochondrial ribosome components were found

to interact with U1-70K. A similar question could be asked regarding the mitochondrial ribosome.

3.5.4 Examining Biophysical Properties of LC1 U1-70K LLPS

U1-70K is a strong candidate for LLPS due to its granule localization, ability to form oligomers, and containing two low complexity domains. Direct observation of U1-70K LLPS has yet to be observed. Droplets of purified LC1 could be observed under a differential interference microscope, due the refractive index of water being lower than putative LC1 droplets (118). Additives such as salts, heparin, crowding agents. and RNA could then be tested for their ability to modify LLPS, providing insight into U1-70K LLPS mechanisms (134, 195, 207). To examine the MC domains influence on granule stability fluorescence recovery after photobleaching (FRAP) miscopy could be used. Granules formed from mutants of U1-70K fluorescently tagged could be photobleached then the rate of recovery measured, to determine stability of U1-70K granules or other structures. I would expect that removal of the LC1 domain to decrease U1-70K granule stability.

3.6 Long Term Future Directions

Factors that regulate RNA Granule disassembly are viable targets for therapeutics

Interactions between various LC domains facilitate RNA granule and pathological protein aggregation. RNA granules are also known to be highly dynamic, growing, shrinking, and moving. This inherently suggests that native mechanisms in the cell must regulate RNA granule disassembly, as LC domains contribute to assembly. To summarize, if MC and other LC domains self-interact what prevents them from constantly oligomerizing, forming clumps, and aggregates in the cell? Recent findings suggests that disassembly of stress granules might be force driven mechanism to release proteins out of kinetic traps (125). I think ATP driven granule disassembly is a key component of granule dynamics, since depletion of ATP inhibits granule dynamics, and numerous granule proteins contain ATPase domains (125, 126). Two major families of ATPases can be found in granules: 1) RNA helicases which unwind RNA secondary structure and 2)

protein chaperones which aid in protein folding. Inhibition of RNA helicase DDX6 halts granule disassembly (125). Protein chaperones are involved in both maintaining stress granule fluidity and clearance (208). Phosphorylation and methylation of residues influences phase separation (209-211). Several RNA helicases, protein chaperones, and kinases were found to interact with U1-70K in this study. Future in vitro studies could examine how these various enzymes influence U1-70K oligomerization utilizing purified protein and blue native gels. In addition, in vivo studies could examine granule stability by modulating enzyme levels in the cell and using FRAP as described above. We are just now beginning to understand the mechanisms behind granule modulation. These factors that regulate RNA granule dynamics are viable and attractive target for therapeutics. Activating or inhibiting these factors could prevent or even remove RBP aggregates in AD or other neurodegenerative diseases, resulting in relief from symptoms. I think understanding these mechanisms will likely be a focus of future studies and lead to drugs that increase life span and cognition.

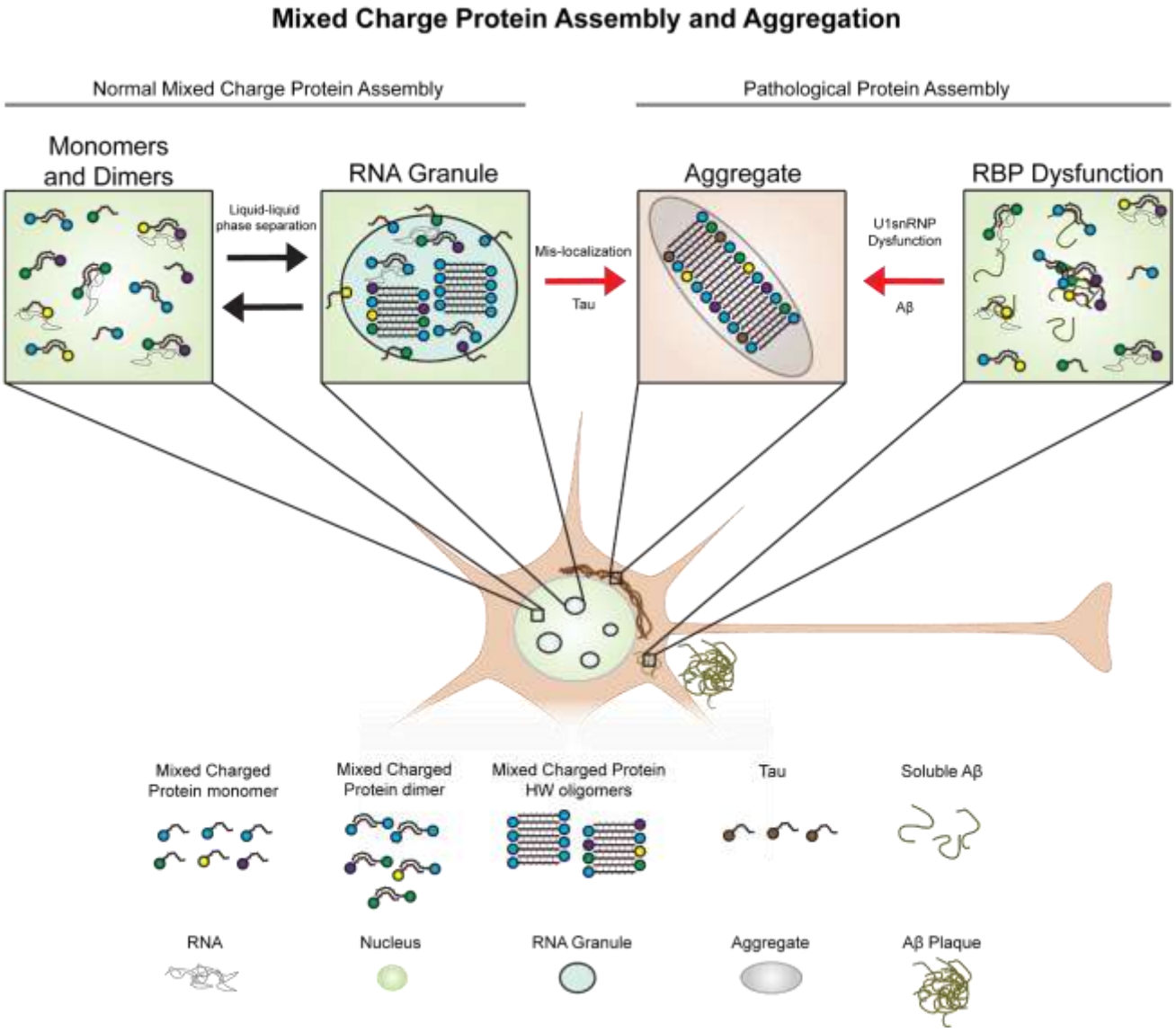


Figure 3.1. A Model for the assembly and pathological aggregation of U1-70K and other RNA-binding proteins with mixed charge domains in AD. RNA binding proteins with mixed charge domains reciprocally interact to form dimers, oligomers, and RNA granules under normal endogenous conditions. In AD, RNA binding proteins with mixed charge domains interact with Aβ and co-aggregate with tau in the cytoplasm to form aggregates.

4.0 References

1. Hippius H, Neundörfer G. The discovery of Alzheimer's disease. Dialogues in Clinical Neuroscience. 2003;5(1):101-8. PubMed PMID: PMC3181715.
2. Association As. 2018 ALZHEIMER'S DISEASE FACTS AND FIGURES: Alzheimer's Association

2018 [4/9/2018]. Available from: https://www.alz.org/facts/.
3. Hebert LE, Beckett LA, Scherr PA, Evans DA. Annual incidence of Alzheimer disease in the United States projected to the years 2000 through 2050. Alzheimer disease and associated disorders. 2001;15(4):169-73. Epub 2001/11/28. PubMed PMID: 11723367.
4. Association As. Major Milestones in Alzheimer's and Brain Research2018. Available from: https://www.alz.org/research/science/major_milestones_in_alzheimers.asp.
5. Stahnisch FW. Max Bielschowsky (1869–1940). Journal of Neurology. 2015;262(3):792-4. doi: 10.1007/s00415-014-7544-z.
6. Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Molecular Medicine. 2016;8(6):595-608. doi: 10.15252/emmm.201606210. PubMed PMID: PMC4888851.
7. Association As. Major Milestones in Alzheimer's and Brain Research 2018. Available from: https://www.alz.org/research/science/major_milestones_in_alzheimers.asp.
8. A. G.Greenwald TCB, & T. M.Ostrom. Cognitive learning, cognitive response persuasion, and attitude change. Psychological foundations of attitudes New York. 1968.
9. Miyakawa T, Uehara Y. Observations of amyloid angiopathy and senile plaques by the scanning electron microscope. Acta Neuropathol. 1979;48(2):153-6. Epub 1979/11/01. PubMed PMID: 506698.
10. Kidd M. Paired Helical Filaments in Electron Microscopy of Alzheimer's Disease. Nature. 1963;197:192. doi: 10.1038/197192b0.
11. Blessed G, Tomlinson BE, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. The British journal of psychiatry : the journal of mental science. 1968;114(512):797-811. Epub 1968/07/01. PubMed PMID: 5662937.
12. Yamaguchi H, Nakazato Y, Hirai S, Shoji M, Harigaya Y. Electron micrograph of diffuse plaques. Initial stage of senile plaque formation in the Alzheimer brain. The American Journal of Pathology. 1989;135(4):593-7. PubMed PMID: PMC1880032.
13. Glenner GG, Wong CW. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. 1984. Biochemical and biophysical research communications. 2012;425(3):534-9. Epub 2012/08/29. doi: 10.1016/j.bbrc.2012.08.020. PubMed PMID: 22925670.

14. Masters CL, Simms G, Weinman NA, Multhaup G, McDonald BL, Beyreuther K. Amyloid plaque core protein in Alzheimer disease and Down syndrome. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(12):4245-9. Epub 1985/06/01. PubMed PMID: 3159021; PubMed Central PMCID: PMCPMC397973.
15. Goldgaber D, Lerman MI, McBride OW, Saffiotti U, Gajdusek DC. Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. Science (New York, NY). 1987;235(4791):877-80. Epub 1987/02/20. PubMed PMID: 3810169.
16. Brion J-P. Immunological demonstration of tau protein in neurofibrillary tangles of Alzheimer's disease2006. 177-85 p.
17. Grundke-Iqbal I, Iqbal K, Quinlan M, Tung YC, Zaidi MS, Wisniewski HM. Microtubule-associated protein tau. A component of Alzheimer paired helical filaments. The Journal of biological chemistry. 1986;261(13):6084-9. Epub 1986/05/05. PubMed PMID: 3084478.
18. Brion J-P, Passareiro H, Nunez J, Flament-Durand J. Mise en évidence immunologique de la protéine tau au niveau des lésions de dégénérescence neurofibrillaire de la maladie d'Alzheimer. Arch Biol1985. 229-35 p.
19. Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, et al. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound-B. Annals of neurology. 2004;55(3):306-19. Epub 2004/03/03. doi: 10.1002/ana.20009. PubMed PMID: 14991808.
20. Saint-Aubert L, Lemoine L, Chiotis K, Leuzy A, Rodriguez-Vieitez E, Nordberg A. Tau PET imaging: present and future directions. Molecular Neurodegeneration. 2017;12(1):19. doi: 10.1186/s13024-017-0162-3.
21. Goate A, Chartier-Harlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. Nature. 1991;349(6311):704-6. Epub 1991/02/21. doi: 10.1038/349704a0. PubMed PMID: 1671712.
22. Mullan M, Tsuji S, Miki T, Katsuya T, Naruse S, Kaneko K, et al. Clinical comparison of Alzheimer's disease in pedigrees with the codon 717 Val→Ile mutation in the amyloid precursor protein gene. Neurobiology of Aging. 1993;14(5):407-19. doi: https://doi.org/10.1016/0197-4580(93)90099-W.
23. Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature. 1995;375(6534):754-60. Epub 1995/06/29. doi: 10.1038/375754a0. PubMed PMID: 7596406.
24. Schellenberg GD, Bird TD, Wijsman EM, Orr HT, Anderson L, Nemens E, et al. Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. Science (New York, NY). 1992;258(5082):668-71. Epub 1992/10/23. PubMed PMID: 1411576.
25. Levy-Lahad E, Wijsman EM, Nemens E, Anderson L, Goddard KA, Weber JL, et al. A familial Alzheimer's disease locus on chromosome 1. Science (New York, NY). 1995;269(5226):970-3. Epub 1995/08/18. PubMed PMID: 7638621.
26. Cai Y, An SSA, Kim S. Mutations in presenilin 2 and its implications in Alzheimer's disease and other dementia-associated disorders. Clinical Interventions in Aging. 2015;10:1163-72. doi: 10.2147/CIA.S85808. PubMed PMID: PMC4507455.
27. De Strooper B. Aph-1, Pen-2, and Nicastrin with Presenilin generate an active gamma-Secretase complex. Neuron. 2003;38(1):9-12. Epub 2003/04/15. PubMed PMID: 12691659.
28. Karlinsky H. Alzheimer's disease in Down's syndrome. A review. Journal of the American Geriatrics Society. 1986;34(10):728-34. Epub 1986/10/01. PubMed PMID: 2944942.
29. Bird TD, Sumi SM, Nemens EJ, Nochlin D, Schellenberg G, Lampe TH, et al. Phenotypic heterogeneity in familial Alzheimer's disease: a study of 24 kindreds. Annals of neurology. 1989;25(1):12-25. Epub 1989/01/01. doi: 10.1002/ana.410250104. PubMed PMID: 2913924.

30. Bird TD, Lampe TH, Nemens EJ, Sumi SM, Nochlin D, Schellenberg GD, et al. Characteristics of familial Alzheimer's disease in nine kindreds of Volga German ancestry. Progress in clinical and biological research. 1989;317:229-34. Epub 1989/01/01. PubMed PMID: 2602419.
31. Strobel G. What Is Early Onset Familial Alzheimer Disease (eFAD)? ALZFORUM2018. Available from: https://www.alzforum.org/early-onset-familial-ad/overview/what-early-onset-familial-alzheimer-disease-efad.
32. Tanzi RE. The genetics of Alzheimer disease. Cold Spring Harb Perspect Med. 2012;2(10). Epub 2012/10/03. doi: 10.1101/cshperspect.a006296. PubMed PMID: 23028126; PubMed Central PMCID: PMCPMC3475404.
33. Foster NL, Wilhelmsen K, Sima AA, Jones MZ, D'Amato CJ, Gilman S. Frontotemporal dementia and parkinsonism linked to chromosome 17: a consensus conference. Conference Participants. Annals of neurology. 1997;41(6):706-15. Epub 1997/06/01. doi: 10.1002/ana.410410606. PubMed PMID: 9189031.
34. Wolfe MS. Tau mutations in neurodegenerative diseases. The Journal of biological chemistry. 2009;284(10):6021-5. Epub 2008/10/25. doi: 10.1074/jbc.R800013200. PubMed PMID: 18948254.
35. Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, et al. The Association between a Polygenic Alzheimer Score and Cortical Thickness in Clinically Normal Subjects. Cerebral Cortex (New York, NY). 2012;22(11):2653-61. doi: 10.1093/cercor/bhr348. PubMed PMID: PMC3464416.
36. Van Cauwenberghe C, Van Broeckhoven C, Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genetics In Medicine. 2015;18:421. doi: 10.1038/gim.2015.117

https://www.nature.com/articles/gim2015117#supplementary-information.
37. Pericak-Vance MA, Bebout JL, Gaskell PC, Jr., Yamaoka LH, Hung WY, Alberts MJ, et al. Linkage studies in familial Alzheimer disease: evidence for chromosome 19 linkage. American journal of human genetics. 1991;48(6):1034-50. Epub 1991/06/01. PubMed PMID: 2035524; PubMed Central PMCID: PMCPMC1683100.
38. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science (New York, NY). 1993;261(5123):921-3. Epub 1993/08/13. PubMed PMID: 8346443.
39. Holtzman DM. In vivo effects of ApoE and clusterin on amyloid-beta metabolism and neuropathology. Journal of molecular neuroscience : MN. 2004;23(3):247-54. Epub 2004/06/08. doi: 10.1385/jmn:23:3:247. PubMed PMID: 15181253.
40. Mahley RW, Rall SC, Jr. Apolipoprotein E: far more than a lipid transport protein. Annual review of genomics and human genetics. 2000;1:507-37. Epub 2001/11/10. doi: 10.1146/annurev.genom.1.1.507. PubMed PMID: 11701639.
41. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. Jama. 1997;278(16):1349-56. Epub 1997/10/29. PubMed PMID: 9343467.
42. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(5):1977-81. Epub 1993/03/01. PubMed PMID: 8446617; PubMed Central PMCID: PMCPMC46003.
43. Castellano JM, Kim J, Stewart FR, Jiang H, DeMattos RB, Patterson BW, et al. Human apoE isoforms differentially regulate brain amyloid-beta peptide clearance. Science translational

medicine. 2011;3(89):89ra57. Epub 2011/07/01. doi: 10.1126/scitranslmed.3002156. PubMed PMID: 21715678; PubMed Central PMCID: PMCPMC3192364.
44. Chartier-Harlin MC, Parfitt M, Legrain S, Perez-Tur J, Brousseau T, Evans A, et al. Apolipoprotein E, epsilon 4 allele as a major risk factor for sporadic early and late-onset forms of Alzheimer's disease: analysis of the 19q13.2 chromosomal region. Human molecular genetics. 1994;3(4):569-74. Epub 1994/04/01. PubMed PMID: 8069300.
45. Strittmatter WJ, Saunders AM, Schmechel D, Pericak-Vance M, Enghild J, Salvesen GS, et al. Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. Proceedings of the National Academy of Sciences. 1993;90(5):1977-81. doi: 10.1073/pnas.90.5.1977.
46. Selkoe DJ. Alzheimer's disease: genotypes, phenotypes, and treatments. Science. 1997;275(5300):630-1. Epub 1997/01/31. PubMed PMID: 9019820.
47. Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012;488(7409):96-9. Epub 2012/07/18. doi: 10.1038/nature11283. PubMed PMID: 22801501.
48. Hardy J, Selkoe DJ. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. Science. 2002;297(5580):353-6. doi: 10.1126/science.1072994.
49. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. Science. 1992;256(5054):184-5. Epub 1992/04/10. PubMed PMID: 1566067.
50. Cleary JP, Walsh DM, Hofmeister JJ, Shankar GM, Kuskowski MA, Selkoe DJ, et al. Natural oligomers of the amyloid-beta protein specifically disrupt cognitive function. Nature neuroscience. 2005;8(1):79-84. Epub 2004/12/21. doi: 10.1038/nn1372. PubMed PMID: 15608634.
51. Tanzi RE, Bertram L. Twenty Years of the Alzheimer’s Disease Amyloid Hypothesis: A Genetic Perspective. Cell. 2005;120(4):545-55. doi: http://dx.doi.org/10.1016/j.cell.2005.02.008.
52. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. Annals of neurology. 1999;45(3):358-68. Epub 1999/03/11. PubMed PMID: 10072051.
53. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. 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. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011;7(3):280-92. doi: 10.1016/j.jalz.2011.03.003. PubMed PMID: PMC3220946.
54. Aizenstein HJ, Nebes RD, Saxton JA, Price JC, Mathis CA, Tsopelas ND, et al. Frequent Amyloid Deposition Without Significant Cognitive Impairment Among the Elderly. Archives of neurology. 2008;65(11):1509-17. doi: 10.1001/archneur.65.11.1509. PubMed PMID: PMC2636844.
55. Morris GP, Clark IA, Vissel B. Inconsistencies and Controversies Surrounding the Amyloid Hypothesis of Alzheimer's Disease. Acta Neuropathologica Communications. 2014;2:135. doi: 10.1186/s40478-014-0135-5. PubMed PMID: PMC4207354.
56. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. 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. Alzheimer's and Dementia. 2011;7(3):280-92.
57. THERAPEUTICS 2018. Available from: https://www.alzforum.org/therapeutics/search?target_types%5B0%5D=170.
58. Holmes C, Boche D, Wilkinson D, Yadegarfar G, Hopkins V, Bayer A, et al. Long-term effects of Abeta42 immunisation in Alzheimer's disease: follow-up of a randomised, placebo-controlled phase I trial. Lancet (London, England). 2008;372(9634):216-23. Epub 2008/07/22. doi: 10.1016/s0140-6736(08)61075-2. PubMed PMID: 18640458.

59. Doody RS, Raman R, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. A phase 3 trial of semagacestat for treatment of Alzheimer's disease. The New England journal of medicine. 2013;369(4):341-50. Epub 2013/07/26. doi: 10.1056/NEJMoa1210951. PubMed PMID: 23883379.
60. Benilova I, Karran E, De Strooper B. The toxic Abeta oligomer and Alzheimer's disease: an emperor in need of clothes. Nat Neurosci. 2012;15(3):349-57. Epub 2012/01/31. doi: 10.1038/nn.3028. PubMed PMID: 22286176.
61. Kokjohn TA, Roher AE. Amyloid precursor protein transgenic mouse models and Alzheimer's disease: understanding the paradigms, limitations, and contributions. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2009;5(4):340-7. Epub 2009/06/30. doi: 10.1016/j.jalz.2009.03.002. PubMed PMID: 19560104; PubMed Central PMCID: PMCPMC2704491.
62. Elder GA, Gama Sosa MA, De Gasperi R. Transgenic Mouse Models of Alzheimer's Disease. The Mount Sinai journal of medicine, New York. 2010;77(1):69-81. doi: 10.1002/msj.20159. PubMed PMID: PMC2925685.
63. Saito T, Matsuba Y, Mihira N, Takano J, Nilsson P, Itohara S, et al. Single App knock-in mouse models of Alzheimer's disease. Nature Neuroscience. 2014;17:661. doi: 10.1038/nn.3697

https://www.nature.com/articles/nn.3697#supplementary-information.
64. Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, et al. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. Nature. 1998;393(6686):702-5. Epub 1998/06/26. doi: 10.1038/31508. PubMed PMID: 9641683.
65. Jack CR, Jr., Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. The Lancet Neurology. 2010;9(1):119-28. Epub 2010/01/20. doi: 10.1016/s1474-4422(09)70299-6. PubMed PMID: 20083042; PubMed Central PMCID: PMCPMC2819840.
66. Chandramouli K, Qian P-Y. Proteomics: Challenges, Techniques and Possibilities to Overcome Biological Sample Complexity. Human Genomics and Proteomics : HGP. 2009;2009:239204. doi: 10.4061/2009/239204. PubMed PMID: PMC2950283.
67. Maier T, Güell M, Serrano L. Correlation of mRNA and protein in complex biological samples. FEBS Letters. 2009;583(24):3966-73. doi: https://doi.org/10.1016/j.febslet.2009.10.036.
68. Yates JR, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. Annual review of biomedical engineering. 2009;11:49-79. Epub 2009/04/30. doi: 10.1146/annurev-bioeng-061008-124934. PubMed PMID: 19400705.
69. Bagramyan K, Barash JR, Arnon SS, Kalkum M. Attomolar Detection of Botulinum Toxin Type A in Complex Biological Matrices. PLoS ONE. 2008;3(4):e2041. doi: 10.1371/journal.pone.0002041. PubMed PMID: PMC2323579.
70. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature genetics. 2013;45(12):1452-8. doi: 10.1038/ng.2802. PubMed PMID: PMC3896259.
71. Soto C. Unfolding the role of protein misfolding in neurodegenerative diseases. Nature Reviews Neuroscience. 2003;4:49. doi: 10.1038/nrn1007.
72. Ross CA, Poirier MA. Protein aggregation and neurodegenerative disease. Nature medicine. 2004;10 Suppl:S10-7. Epub 2004/07/24. doi: 10.1038/nm1066. PubMed PMID: 15272267.
73. Diner I, Nguyen T, Seyfried NT. Enrichment of Detergent-insoluble Protein Aggregates from Human Postmortem Brain. Journal of visualized experiments : JoVE. 2017(128). Epub 2017/11/21. doi: 10.3791/55835. PubMed PMID: 29155708.
74. Bai B, Hales CM, Chen PC, Gozal Y, Dammer EB, Fritz JJ, et al. 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. 2013;110(41):16562-7. Epub 2013/09/12. doi: 10.1073/pnas.1310249110. PubMed PMID: 24023061.
75. Hales CM, Seyfried NT, Dammer EB, Duong D, Yi H, Gearing M, et al. U1 small nuclear ribonucleoproteins (snRNPs) aggregate in Alzheimer's disease due to autosomal dominant genetic mutations and trisomy 21. Mol Neurodegener. 2014;9:15. PubMed PMID: 24773620.
76. Hales CM, Dammer EB, Deng Q, Duong DM, Gearing M, Troncoso JC, et al. Changes in the detergent-insoluble brain proteome linked to amyloid and tau in Alzheimer's Disease progression. PROTEOMICS. 2016;16(23):3042-53. doi: 10.1002/pmic.201600057.
77. Kondo Y, Oubridge C, van Roon AM, Nagai K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. eLife. 2015;4. Epub 2015/01/03. doi: 10.7554/eLife.04986. PubMed PMID: 25555158; PubMed Central PMCID: PMCPMC4383343.
78. Wang YAN, Liu J, Huang BO, Xu Y-M, Li J, Huang L-F, et al. Mechanism of alternative splicing and its regulation. Biomedical Reports. 2015;3(2):152-8. doi: 10.3892/br.2014.407. PubMed PMID: PMC4360811.
79. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, et al. Alternative isoform regulation in human tissue transcriptomes. Nature. 2008;456:470. doi: 10.1038/nature07509

https://www.nature.com/articles/nature07509#supplementary-information.
80. McConnell TS, Lokken RP, Steitz JA. Assembly of the U1 snRNP involves interactions with the backbone of the terminal stem of U1 snRNA. RNA. 2003;9(2):193-201. doi: 10.1261/rna.2136103. PubMed PMID: PMC1370385.
81. Will CL, Lührmann R. Spliceosome Structure and Function. Cold Spring Harbor Perspectives in Biology. 2011;3(7):a003707. doi: 10.1101/cshperspect.a003707. PubMed PMID: PMC3119917.
82. Shepard PJ, Hertel KJ. The SR protein family. Genome Biology. 2009;10(10):242-. doi: 10.1186/gb-2009-10-10-242. PubMed PMID: PMC2784316.
83. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. The Biochemical journal. 2009;417(1):15-27. Epub 2008/12/09. doi: 10.1042/bj20081501. PubMed PMID: 19061484.
84. Cao W, Garcia-Blanco MA. A serine/arginine-rich domain in the human U1 70k protein is necessary and sufficient for ASF/SF2 binding. The Journal of biological chemistry. 1998;273(32):20629-35. Epub 1998/08/01. PubMed PMID: 9685421.
85. Berg Michael G, Singh Larry N, Younis I, Liu Q, Pinto Anna M, Kaida D, et al. U1 snRNP Determines mRNA Length and Regulates Isoform Expression. Cell.150(1):53-64. doi: 10.1016/j.cell.2012.05.029.
86. Cheng Z, Shang Y, Gao S, Zhang T. Overexpression of U1 snRNA induces decrease of U1 spliceosome function associated with Alzheimer's disease. Journal of Neurogenetics. 2017;31(4):337-43. doi: 10.1080/01677063.2017.1395425.
87. Cheng Z, Du Z, Shang Y, Zhang Y, Zhang T. A Preliminary Study: PS1 Increases U1 snRNA Expression Associated with AD. Journal of molecular neuroscience : MN. 2017;62(3-4):269-75. Epub 2017/06/04. doi: 10.1007/s12031-017-0932-y. PubMed PMID: 28577205.
88. Cohen TJ, Hwang AW, Unger T, Trojanowski JQ, Lee VM. Redox signalling directly regulates TDP-43 via cysteine oxidation and disulphide cross-linking. EMBO J. 2012;31(5):1241-52. Epub 2011/12/24. doi: 10.1038/emboj.2011.471. PubMed PMID: 22193716; PubMed Central PMCID: PMCPMC3297986.
89. Neumann M, Rademakers R, Roeber S, Baker M, Kretzschmar HA, Mackenzie IR. A new subtype of frontotemporal lobar degeneration with FUS pathology. Brain. 2009;132(Pt

11):2922-31. Epub 2009/08/14. doi: 10.1093/brain/awp214. PubMed PMID: 19674978; PubMed Central PMCID: PMCPMC2768659.
90. Kwong LK, Neumann M, Sampathu DM, Lee VM-Y, Trojanowski JQ. TDP-43 proteinopathy: the neuropathology underlying major forms of sporadic and familial frontotemporal lobar degeneration and motor neuron disease. Acta Neuropathologica. 2007;114(1):63-70. doi: 10.1007/s00401-007-0226-5.
91. Diner I, Hales CM, Rabenold L, Bishof I, Duong DM, Yi H, et al. Aggregation Properties of the Small Nuclear Ribonucleoprotein U1-70K in Alzheimer Disease. The Journal of biological chemistry. 2014. Epub 2014/10/31. doi: 10.1074/jbc.M114.562959. PubMed PMID: 25355317.
92. Maziuk B, Ballance HI, Wolozin B. Dysregulation of RNA Binding Protein Aggregation in Neurodegenerative Disorders. Frontiers in Molecular Neuroscience. 2017;10:89. doi: 10.3389/fnmol.2017.00089. PubMed PMID: PMC5378767.
93. Ramaswami M, Taylor JP, Parker R. Altered ribostasis: RNA-protein granules in degenerative disorders. Cell. 2013;154(4):727-36. Epub 2013/08/21. doi: 10.1016/j.cell.2013.07.038. PubMed PMID: 23953108; PubMed Central PMCID: PMCPMC3811119.
94. Seyfried NT, Gozal YM, Donovan LE, Herskowitz JH, Dammer EB, Xia Q, et al. Quantitative analysis of the detergent-insoluble brain proteome in frontotemporal lobar degeneration using SILAC internal standards. Journal of proteome research. 2012;11(5):2721-38. Epub 2012/03/16. doi: 10.1021/pr2010814. PubMed PMID: 22416763; PubMed Central PMCID: PMC3357000.
95. Papatriantafyllou M. Protein aggregation: The secret recipe for RNA granules. Nat Rev Mol Cell Biol. 2012;13(7):405-.
96. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell. 2012;149(4):753-67. Epub 2012/05/15. doi: 10.1016/j.cell.2012.04.017. PubMed PMID: 22579281.
97. Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, et al. Phosphorylation-Regulated Binding of RNA Polymerase II to Fibrous Polymers of Low-Complexity Domains. Cell. 2013;155(5):1049-60. doi: http://dx.doi.org/10.1016/j.cell.2013.10.033.
98. Shaw DJ, Morse R, Todd AG, Eggleton P, Lorson CL, Young PJ. Identification of a self-association domain in the Ewing's sarcoma protein: a novel function for arginine-glycine-glycine rich motifs? J Biochem. 2010;147(6):885-93. Epub 2010/03/10. doi: 10.1093/jb/mvq025. PubMed PMID: 20211855; PubMed Central PMCID: PMCPMC2912030.
99. Wang IF, Chang HY, Hou SC, Liou GG, Way TD, James Shen CK. The self-interaction of native TDP-43 C terminus inhibits its degradation and contributes to early proteinopathies. Nat Commun. 2012;3:766. Epub 2012/04/05. doi: 10.1038/ncomms1766. PubMed PMID: 22473010.
100. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science (New York, NY). 2008;319(5870):1668-72. Epub 2008/03/01. doi: 10.1126/science.1154584. PubMed PMID: 18309045.
101. Liu-Yesucevitz L, Lin AY, Ebata A, Boon JY, Reid W, Xu Y-F, et al. ALS-Linked Mutations Enlarge TDP-43-Enriched Neuronal RNA Granules in the Dendritic Arbor. The Journal of Neuroscience. 2014;34(12):4167-74. doi: 10.1523/jneurosci.2350-13.2014.
102. Matsumoto T, Matsukawa K, Watanabe N, Kishino Y, Kunugi H, Ihara R, et al. Self-assembly of FUS through its low-complexity domain contributes to neurodegeneration. Human molecular genetics. 2018;27(8):1353-65. doi: 10.1093/hmg/ddy046.
103. Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj Anderson P, Kim Hong J, et al. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. Cell.163(1):123-33. doi: 10.1016/j.cell.2015.09.015.

104. Liu-Yesucevitz L, Lin AY, Ebata A, Boon JY, Reid W, Xu YF, et al. ALS-linked mutations enlarge TDP-43-enriched neuronal RNA granules in the dendritic arbor. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2014;34(12):4167-74. Epub 2014/03/22. doi: 10.1523/jneurosci.2350-13.2014. PubMed PMID: 24647938; PubMed Central PMCID: PMCPMC3960463.
105. Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science (New York, NY). 2009;323(5918):1208-11. Epub 2009/03/03. doi: 10.1126/science.1165942. PubMed PMID: 19251628.
106. Hyman AA, Weber CA, Jülicher F. Liquid-Liquid Phase Separation in Biology. Annual Review of Cell and Developmental Biology. 2014;30(1):39-58. doi: 10.1146/annurev-cellbio-100913-013325. PubMed PMID: 25288112.
107. Mitrea DM, Kriwacki RW. Phase separation in biology; functional organization of a higher order. Cell Communication and Signaling : CCS. 2016;14:1. doi: 10.1186/s12964-015-0125-7. PubMed PMID: PMC4700675.
108. Gabaldón T, Pittis AA. Origin and evolution of metabolic sub-cellular compartmentalization in eukaryotes. Biochimie. 2015;119:262-8. doi: 10.1016/j.biochi.2015.03.021. PubMed PMID: PMC4678951.
109. Hetzer MW. The Nuclear Envelope. Cold Spring Harbor Perspectives in Biology. 2010;2(3):a000539. doi: 10.1101/cshperspect.a000539. PubMed PMID: PMC2829960.
110. Gellerich FN, Trumbeckaite S, Opalka JR, Seppet E, Rasmussen HN, Neuhoff C, et al. Function of the mitochondrial outer membrane as a diffusion barrier in health and diseases. Biochemical Society transactions. 2000;28(2):164-9. Epub 2000/05/18. PubMed PMID: 10816120.
111. Aguilera-Gomez A, Rabouille C. Membrane-bound organelles versus membrane-less compartments and their control of anabolic pathways in Drosophila. Developmental Biology. 2017;428(2):310-7. doi: https://doi.org/10.1016/j.ydbio.2017.03.029.
112. Anderson P, Kedersha N. RNA granules. The Journal of Cell Biology. 2006;172(6):803-8. doi: 10.1083/jcb.200512082. PubMed PMID: PMC2063724.
113. Brangwynne CP, Mitchison TJ, Hyman AA. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(11):4334-9. Epub 2011/03/04. doi: 10.1073/pnas.1017150108. PubMed PMID: 21368180; PubMed Central PMCID: PMCPMC3060270.
114. Handwerger KE, Cordero JA, Gall JG. Cajal bodies, nucleoli, and speckles in the Xenopus oocyte nucleus have a low-density, sponge-like structure. Molecular biology of the cell. 2005;16(1):202-11. Epub 2004/10/29. doi: 10.1091/mbc.E04-08-0742. PubMed PMID: 15509651; PubMed Central PMCID: PMCPMC539164.
115. Mahen R, Venkitaraman AR. Pattern formation in centrosome assembly. Current opinion in cell biology. 2012;24(1):14-23. Epub 2012/01/17. doi: 10.1016/j.ceb.2011.12.012. PubMed PMID: 22245706.
116. Decker CJ, Parker R. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb Perspect Biol. 2012;4(9):a012286. Epub 2012/07/06. doi: 10.1101/cshperspect.a012286. PubMed PMID: 22763747; PubMed Central PMCID: PMCPMC3428773.
117. Spector DL, Lamond AI. Nuclear Speckles. Cold Spring Harbor Perspectives in Biology. 2011;3(2):a000646. doi: 10.1101/cshperspect.a000646. PubMed PMID: PMC3039535.
118. Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell. 2015;163(1):123-33. doi: 10.1016/j.cell.2015.09.015. PubMed PMID: PMC5149108.

119. Han TW, Kato M, Xie S, Wu LC, Mirzaei H, Pei J, et al. Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell. 2012;149(4):768-79. Epub 2012/05/15. doi: 10.1016/j.cell.2012.04.016. PubMed PMID: 22579282.
120. Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, et al. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. Cell. 2017;171(3):615-27 e16. Epub 2017/09/26. doi: 10.1016/j.cell.2017.08.048. PubMed PMID: 28942918; PubMed Central PMCID: PMCPMC5650524.
121. Li P, Banjade S, Cheng HC, Kim S, Chen B, Guo L, et al. Phase transitions in the assembly of multivalent signalling proteins. Nature. 2012;483(7389):336-40. Epub 2012/03/09. doi: 10.1038/nature10879. PubMed PMID: 22398450; PubMed Central PMCID: PMCPMC3343696.
122. Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, Gharakhani J, et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science (New York, NY). 2009;324(5935):1729-32. Epub 2009/05/23. doi: 10.1126/science.1172046. PubMed PMID: 19460965.
123. Kroschwald S, Maharana S, Mateju D, Malinovska L, Nuske E, Poser I, et al. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. eLife. 2015;4:e06807. Epub 2015/08/05. doi: 10.7554/eLife.06807. PubMed PMID: 26238190; PubMed Central PMCID: PMCPMC4522596.
124. Kato M, Lin Y, McKnight SL. Cross-beta polymerization and hydrogel formation by low-complexity sequence proteins. Methods (San Diego, Calif). 2017;126:3-11. Epub 2017/06/19. doi: 10.1016/j.ymeth.2017.06.011. PubMed PMID: 28624540; PubMed Central PMCID: PMCPMC5583018.
125. Smith J, Seydoux G. Liquid-like P granules require ATP hydrolysis to avoid solidification. bioRxiv. 2018. doi: 10.1101/245878.
126. Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, Parker R. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell. 2016;164(3):487-98. Epub 2016/01/19. doi: 10.1016/j.cell.2015.12.038. PubMed PMID: 26777405; PubMed Central PMCID: PMCPMC4733397.
127. Lin Y, Currie SL, Rosen MK. Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. Journal of Biological Chemistry. 2017;292(46):19110-20. doi: 10.1074/jbc.M117.800466.
128. Su X, Ditlev JA, Hui E, Xing W, Banjade S, Okrut J, et al. Phase separation of signaling molecules promotes T cell receptor signal transduction. Science (New York, NY). 2016;352(6285):595-9. Epub 2016/04/09. doi: 10.1126/science.aad9964. PubMed PMID: 27056844; PubMed Central PMCID: PMCPMC4892427.
129. Dao TP, Kolaitis R-M, Kim HJ, O'Donovan K, Martyniak B, Colicino E, et al. Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. Molecular Cell.69(6):965-78.e6. doi: 10.1016/j.molcel.2018.02.004.
130. Mitrea DM, Cika JA, Stanley CB, Nourse A, Onuchic PL, Banerjee PR, et al. Self-interaction of NPM1 modulates multiple mechanisms of liquid–liquid phase separation. Nature Communications. 2018;9(1):842. doi: 10.1038/s41467-018-03255-3.
131. Banani SF, Lee HO, Hyman AA, Rosen MK. Biomolecular condensates: organizers of cellular biochemistry. Nature Reviews Molecular Cell Biology. 2017;18:285. doi: 10.1038/nrm.2017.7

https://www.nature.com/articles/nrm.2017.7#supplementary-information.
132. Ambadipudi S, Biernat J, Riedel D, Mandelkow E, Zweckstetter M. Liquid–liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. Nature communications. 2017;8:275. doi: 10.1038/s41467-017-00480-0. PubMed PMID: PMC5561136.

133. Wegmann S, Eftekharzadeh B, Tepper K, Zoltowska KM, Bennett RE, Dujardin S, et al. Tau protein liquid–liquid phase separation can initiate tau aggregation. The EMBO Journal. 2018. doi: 10.15252/embj.201798049.
134. Zhang X, Lin Y, Eschmann NA, Zhou H, Rauch JN, Hernandez I, et al. RNA stores tau reversibly in complex coacervates. PLoS biology. 2017;15(7):e2002183. Epub 2017/07/07. doi: 10.1371/journal.pbio.2002183. PubMed PMID: 28683104; PubMed Central PMCID: PMCPMC5500003.
135. Vanderweyde T, Yu H, Varnum M, Liu-Yesucevitz L, Citro A, Ikezu T, et al. Contrasting pathology of the stress granule proteins TIA-1 and G3BP in tauopathies. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2012;32(24):8270-83. Epub 2012/06/16. doi: 10.1523/jneurosci.1592-12.2012. PubMed PMID: 22699908; PubMed Central PMCID: PMCPMC3402380.
136. Vanderweyde T, Apicco DJ, Youmans-Kidder K, Ash PEA, Cook C, da Rocha EL, et al. Interaction of tau with the RNA-Binding Protein TIA1 Regulates tau Pathophysiology and Toxicity. Cell reports. 2016;15(7):1455-66. doi: 10.1016/j.celrep.2016.04.045. PubMed PMID: PMC5325702.
137. Apicco DJ, Ash PEA, Maziuk B, LeBlang C, Medalla M, Al Abdullatif A, et al. Reducing the RNA binding protein TIA1 protects against tau-mediated neurodegeneration in vivo. Nature Neuroscience. 2017. doi: 10.1038/s41593-017-0022-z.
138. Mota B, Herculano-Houzel S. All brains are made of this: a fundamental building block of brain matter with matching neuronal and glial masses. Frontiers in Neuroanatomy. 2014;8:127. doi: 10.3389/fnana.2014.00127. PubMed PMID: PMC4228857.
139. Herculano-Houzel S. The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. Frontiers in Human Neuroscience. 2009;3:31. doi: 10.3389/neuro.09.031.2009. PubMed PMID: PMC2776484.
140. King OD, Gitler AD, Shorter J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Research. 2012;1462:61-80. doi: 10.1016/j.brainres.2012.01.016. PubMed PMID: PMC3372647.
141. Jacob J, Duclohier H, Cafiso DS. The role of proline and glycine in determining the backbone flexibility of a channel-forming peptide. Biophysical journal. 1999;76(3):1367-76. Epub 1999/02/27. doi: 10.1016/s0006-3495(99)77298-x. PubMed PMID: 10049319; PubMed Central PMCID: PMCPMC1300115.
142. Wolozin B, Apicco D. RNA binding proteins and the genesis of neurodegenerative diseases. Advances in experimental medicine and biology. 2015;822:11-5. doi: 10.1007/978-3-319-08927-0_3. PubMed PMID: PMC4694570.
143. Vanderweyde T, Yu H, Varnum M, Liu-Yesucevitz L, Citro A, Ikezu T, et al. Contrasting Pathology of the Stress Granule Proteins TIA-1 and G3BP in Tauopathies. The Journal of Neuroscience. 2012;32(24):8270-83. doi: 10.1523/jneurosci.1592-12.2012.
144. Raj T, Li Y, Wong G, Ramdhani S, Wang Y-c, Ng B, et al. Integrative analyses of splicing in the aging brain: role in susceptibility to Alzheimer's Disease. bioRxiv. 2017. doi: 10.1101/174565.
145. Banani SF, Rice AM, Peeples WB, Lin Y, Jain S, Parker R, et al. Compositional Control of Phase-Separated Cellular Bodies. Cell. 2016;166(3):651-63. doi: http://dx.doi.org/10.1016/j.cell.2016.06.010.
146. Nott TJ, Petsalaki E, Farber P, Jervis D, Fussner E, Plochowietz A, et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Molecular cell. 2015;57(5):936-47. Epub 2015/03/10. doi: 10.1016/j.molcel.2015.01.013. PubMed PMID: 25747659; PubMed Central PMCID: PMCPMC4352761.
147. Falkenberg CV, Carson JH, Blinov ML. Multivalent Molecules as Modulators of RNA Granule Size and Composition. Biophysical journal. 2017;113(2):235-45. Epub 2017/03/01. doi:

10.1016/j.bpj.2017.01.031. PubMed PMID: 28242011; PubMed Central PMCID: PMCPMC5529162.
148. Coletta A, Pinney JW, Solis DY, Marsh J, Pettifer SR, Attwood TK. Low-complexity regions within protein sequences have position-dependent roles. BMC Syst Biol. 2010;4:43. Epub 2010/04/14. doi: 10.1186/1752-0509-4-43. PubMed PMID: 20385029; PubMed Central PMCID: PMCPMC2873317.
149. Perutz M. Polar zippers: their role in human disease. Protein Science : A Publication of the Protein Society. 1994;3(10):1629-37. PubMed PMID: PMC2142615.
150. Wittig I, Braun H-P, Schagger H. Blue native PAGE. Nat Protocols. 2006;1(1):418-28.
151. Romac JM, Keene JD. Overexpression of the arginine-rich carboxy-terminal region of U1 snRNP 70K inhibits both splicing and nucleocytoplasmic transport of mRNA. Genes Dev. 1995;9(11):1400-10. Epub 1995/06/01. PubMed PMID: 7797079.
152. Stejskalová E, Staněk D. The splicing factor U1-70K interacts with the SMN complex and is required for nuclear gem integrity. Journal of Cell Science. 2014;127(18):3909-15. doi: 10.1242/jcs.155838.
153. Howell VM, Jones JM, Bergren SK, Li L, Billi AC, Avenarius MR, et al. Evidence for a direct role of the disease modifier SCNM1 in splicing. Human molecular genetics. 2007;16(20):2506-16. Epub 2007/07/28. doi: 10.1093/hmg/ddm206. PubMed PMID: 17656373.
154. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. Molecular & Cellular Proteomics. 2014;13(9):2513-26. doi: 10.1074/mcp.M113.031591.
155. Karpievitch YV, Dabney AR, Smith RD. Normalization and missing value imputation for label-free LC-MS analysis. BMC Bioinformatics. 2012;13(Suppl 16):S5-S. doi: 10.1186/1471-2105-13-S16-S5. PubMed PMID: PMC3489534.
156. Raj T, Shulman JM, Keenan BT, Chibnik LB, Evans DA, Bennett DA, et al. Alzheimer disease susceptibility loci: evidence for a protein network under natural selection. American journal of human genetics. 2012;90(4):720-6. Epub 2012/04/10. doi: 10.1016/j.ajhg.2012.02.022. PubMed PMID: 22482808; PubMed Central PMCID: PMC3322230.
157. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics. 2008;24(5):719-20. Epub 2007/11/21. doi: 10.1093/bioinformatics/btm563. PubMed PMID: 18024473.
158. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559. doi: 10.1186/1471-2105-9-559.
159. Zambon AC, Gaj S, Ho I, Hanspers K, Vranizan K, Evelo CT, et al. GO-Elite: a flexible solution for pathway and ontology over-representation. Bioinformatics. 2012;28(16):2209-10. Epub 2012/06/30. doi: 10.1093/bioinformatics/bts366. PubMed PMID: 22743224; PubMed Central PMCID: PMCPMC3413395.
160. Egloff S, Vitali P, Tellier M, Raffel R, Murphy S, Kiss T. The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression. EMBO J. 2017;36(7):934-48. Epub 2017/03/04. doi: 10.15252/embj.201695740. PubMed PMID: 28254838; PubMed Central PMCID: PMCPMC5376971.
161. Morris GE. The Cajal body. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. 2008;1783(11):2108-15. doi: http://dx.doi.org/10.1016/j.bbamcr.2008.07.016.
162. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics. 2000;25(1):25-9. Epub 2000/05/10. doi: 10.1038/75556. PubMed PMID: 10802651; PubMed Central PMCID: PMCPMC3037419.
163. Seraphin B, Rosbash M. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. Cell. 1989;59(2):349-58. doi: http://dx.doi.org/10.1016/0092-8674(89)90296-1.

164. Kwon I, Kato M, Xiang S, Wu L, Theodoropoulos P, Mirzaei H, et al. Phosphorylation-regulated Binding of RNA Polymerase II to Fibrous Polymers of Low Complexity Domains. Cell. 2013;155(5):1049-60. doi: 10.1016/j.cell.2013.10.033. PubMed PMID: PMC4010232.
165. Zhou A, Ou AC, Cho A, Benz EJ, Jr., Huang SC. Novel splicing factor RBM25 modulates Bcl-x pre-mRNA 5' splice site selection. Molecular and cellular biology. 2008;28(19):5924-36. Epub 2008/07/30. doi: 10.1128/mcb.00560-08. PubMed PMID: 18663000; PubMed Central PMCID: PMCPMC2546994.
166. Li Y, Ito M, Sun S, Chida T, Nakashima K, Suzuki T. LUC7L3/CROP inhibits replication of hepatitis B virus via suppressing enhancer II/basal core promoter activity. Sci Rep. 2016;6:36741. Epub 2016/11/20. doi: 10.1038/srep36741. PubMed PMID: 27857158; PubMed Central PMCID: PMCPMC5114668.
167. Cherry JD, Zeineddin A, Dammer EB, Webster JA, Duong D, Seyfried NT, et al. Characterization of Detergent Insoluble Proteome in Chronic Traumatic Encephalopathy. Journal of neuropathology and experimental neurology. 2018;77(1):40-9. Epub 2017/11/18. doi: 10.1093/jnen/nlx100. PubMed PMID: 29145658.
168. Hales CM, Dammer EB, Diner I, Yi H, Seyfried NT, Gearing M, et al. Aggregates of Small Nuclear Ribonucleic Acids (snRNAs) in Alzheimer's Disease. Brain Pathology. 2014;24(4):344-51. doi: 10.1111/bpa.12133.
169. Dammer EB, Fallini C, Gozal YM, Duong DM, Rossoll W, Xu P, et al. Coaggregation of RNA-binding proteins in a model of TDP-43 proteinopathy with selective RGG motif methylation and a role for RRM1 ubiquitination. PLoS One. 2012;7(6):e38658. Epub 2012/07/05. doi: 10.1371/journal.pone.0038658. PubMed PMID: 22761693; PubMed Central PMCID: PMCPMC3380899.
170. Seyfried NT, Dammer EB, Swarup V, Nandakumar D, Duong DM, Yin L, et al. A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer’s Disease. Cell Systems. 2017;4(1):60-72.e4. doi: 10.1016/j.cels.2016.11.006.
171. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology. 2008;26:1367. doi: 10.1038/nbt.1511

https://www.nature.com/articles/nbt.1511#supplementary-information.
172. Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. Journal of proteome research. 2011;10(4):1794-805. doi: 10.1021/pr101065j.
173. Luber CA, Cox J, Lauterbach H, Fancke B, Selbach M, Tschopp J, et al. Quantitative proteomics reveals subset-specific viral recognition in dendritic cells. Immunity. 2010;32(2):279-89. doi: 10.1016/j.immuni.2010.01.013. PubMed PMID: 20171123.
174. Maaten Lvd, Hinton G. Visualizing data using t-SNE. Journal of Machine Learning Research. 2008;9(Nov):2579-605.
175. Webb CH, Hertel KJ. Preparation of splicing competent nuclear extracts. Methods in molecular biology (Clifton, NJ). 2014;1126:117-21. Epub 2014/02/20. doi: 10.1007/978-1-62703-980-2_8. PubMed PMID: 24549659; PubMed Central PMCID: PMCPMC4501576.
176. Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology. 1991;41(4):479-86. Epub 1991/04/01. PubMed PMID: 2011243.
177. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol. 1991;82(4):239-59. Epub 1991/01/01. PubMed PMID: 1759558.
178. Umoh ME, Dammer EB, Dai J, Duong DM, Lah JJ, Levey AI, et al. A proteomic network approach across the ALS-FTD disease spectrum resolves clinical phenotypes and genetic

vulnerability in human brain. EMBO Molecular Medicine. 2018;10(1):48-62. doi: doi:10.15252/emmm.201708202.
179. Patil A, Kinoshita K, Nakamura H. Domain distribution and intrinsic disorder in hubs in the human protein-protein interaction network. Protein Sci. 2010;19(8):1461-8. Epub 2010/05/29. doi: 10.1002/pro.425. PubMed PMID: 20509167; PubMed Central PMCID: PMCPMC2923499.
180. Elkon R, Ugalde AP, Agami R. Alternative cleavage and polyadenylation: extent, regulation and function. Nature reviews Genetics. 2013;14(7):496-506. Epub 2013/06/19. doi: 10.1038/nrg3482. PubMed PMID: 23774734.
181. Herai RH, Negraes PD, Muotri AR. Evidence of nuclei-encoded spliceosome mediating splicing of mitochondrial RNA. Human molecular genetics. 2017;26(13):2472-9. doi: 10.1093/hmg/ddx142.
182. Glanzer J, Miyashiro KY, Sul JY, Barrett L, Belt B, Haydon P, et al. RNA splicing capability of live neuronal dendrites. Proceedings of the National Academy of Sciences of the United States of America. 2005;102(46):16859-64. Epub 2005/11/09. doi: 10.1073/pnas.0503783102. PubMed PMID: 16275927; PubMed Central PMCID: PMCPMC1277967.
183. Buckley PT, Khaladkar M, Kim J, Eberwine J. Cytoplasmic intron retention, function, splicing, and the sentinel RNA hypothesis. Wiley interdisciplinary reviews RNA. 2014;5(2):223-30. doi: 10.1002/wrna.1203. PubMed PMID: PMC4449146.
184. Onyango IG, Dennis J, Khan SM. Mitochondrial Dysfunction in Alzheimer's Disease and the Rationale for Bioenergetics Based Therapies. Aging and Disease. 2016;7(2):201-14. doi: 10.14336/AD.2015.1007. PubMed PMID: PMC4809610.
185. Hir HL, Saulière J, Wang Z. The exon junction complex as a node of post-transcriptional networks. Nature Reviews Molecular Cell Biology. 2015;17:41. doi: 10.1038/nrm.2015.7.
186. Scotti MM, Swanson MS. RNA mis-splicing in disease. Nature Reviews Genetics. 2015;17:19. doi: 10.1038/nrg.2015.3.
187. Vance C, Scotter EL, Nishimura AL, Troakes C, Mitchell JC, Kathe C, et al. ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules. Human molecular genetics. 2013;22(13):2676-88. Epub 2013/03/12. doi: 10.1093/hmg/ddt117. PubMed PMID: 23474818; PubMed Central PMCID: PMCPMC3674807.
188. Romac JM, Graff DH, Keene JD. The U1 small nuclear ribonucleoprotein (snRNP) 70K protein is transported independently of U1 snRNP particles via a nuclear localization signal in the RNA-binding domain. Molecular and cellular biology. 1994;14(7):4662-70. Epub 1994/07/01. PubMed PMID: 7516470; PubMed Central PMCID: PMCPMC358839.
189. Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, et al. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature. 2013;495(7442):467-73. Epub 2013/03/05. doi: 10.1038/nature11922. PubMed PMID: 23455423; PubMed Central PMCID: PMCPMC3756911.
190. Harrison AF, Shorter J. RNA-binding proteins with prion-like domains in health and disease. The Biochemical journal. 2017;474(8):1417-38. doi: 10.1042/BCJ20160499. PubMed PMID: PMC5639257.
191. Hales CM, Dammer EB, Deng Q, Duong DM, Gearing M, Troncoso JC, et al. Changes in the detergent-insoluble brain proteome linked to amyloid and tau in Alzheimer's Disease progression. Proteomics. 2016;16(23):3042-53. doi: 10.1002/pmic.201600057. PubMed PMID: PMC5462625.
192. Oakley H, Cole SL, Logan S, Maus E, Shao P, Craft J, et al. Intraneuronal β-Amyloid Aggregates, Neurodegeneration, and Neuron Loss in Transgenic Mice with Five Familial Alzheimer's Disease Mutations: Potential Factors in Amyloid Plaque Formation. The Journal of Neuroscience. 2006;26(40):10129-40. doi: 10.1523/jneurosci.1202-06.2006.

193. Gouras GK, Tampellini D, Takahashi RH, Capetillo-Zarate E. Intraneuronal β-amyloid accumulation and synapse pathology in Alzheimer's disease. Acta neuropathologica. 2010;119(5):523-41. doi: 10.1007/s00401-010-0679-9. PubMed PMID: PMC3183823.
194. Hernández-Vega A, Braun M, Scharrel L, Jahnel M, Wegmann S, Hyman BT, et al. Local Nucleation of Microtubule Bundles through Tubulin Concentration into a Condensed Tau Phase. Cell Reports.20(10):2304-12. doi: 10.1016/j.celrep.2017.08.042.
195. Ambadipudi S, Biernat J, Riedel D, Mandelkow E, Zweckstetter M. Liquid–liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. Nature Communications. 2017;8(1):275. doi: 10.1038/s41467-017-00480-0.
196. Fitzpatrick AWP, Falcon B, He S, Murzin AG, Murshudov G, Garringer HJ, et al. Cryo-EM structures of tau filaments from Alzheimer's disease. Nature. 2017;547(7662):185-90. doi: 10.1038/nature23002.
197. Mayer ML, Hieter P. Protein networks—built by association. Nature Biotechnology. 2000;18:1242. doi: 10.1038/82342.
198. Michels AA, Nguyen VT, Fraldi A, Labas V, Edwards M, Bonnet F, et al. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. Molecular and cellular biology. 2003;23(14):4859-69. Epub 2003/07/02. PubMed PMID: 12832472; PubMed Central PMCID: PMCPMC162212.
199. Prelich G. Gene Overexpression: Uses, Mechanisms, and Interpretation. Genetics. 2012;190(3):841-54. doi: 10.1534/genetics.111.136911. PubMed PMID: PMC3296252.
200. Fabrini R, De Luca A, Stella L, Mei G, Orioni B, Ciccone S, et al. Monomer-dimer equilibrium in glutathione transferases: a critical re-examination. Biochemistry. 2009;48(43):10473-82. Epub 2009/10/03. doi: 10.1021/bi901238t. PubMed PMID: 19795889.
201. Yang AJ, Knauer M, Burdick DA, Glabe C. Intracellular A beta 1-42 aggregates stimulate the accumulation of stable, insoluble amyloidogenic fragments of the amyloid precursor protein in transfected cells. The Journal of biological chemistry. 1995;270(24):14786-92. Epub 1995/06/16. PubMed PMID: 7782344.
202. Gouras GK, Tampellini D, Takahashi RH, Capetillo-Zarate E. Intraneuronal beta-amyloid accumulation and synapse pathology in Alzheimer's disease. Acta Neuropathol. 2010;119(5):523-41. Epub 2010/04/01. doi: 10.1007/s00401-010-0679-9. PubMed PMID: 20354705; PubMed Central PMCID: PMCPMC3183823.
203. Gunawardana CG, Mehrabian M, Wang X, Mueller I, Lubambo IB, Jonkman JEN, et al. 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. Molecular & Cellular Proteomics. 2015;14(11):3000-14. doi: 10.1074/mcp.M115.050724.
204. Meier S, Bell M, Lyons DN, Rodriguez-Rivera J, Ingram A, Fontaine SN, et al. Pathological Tau Promotes Neuronal Damage by Impairing Ribosomal Function and Decreasing Protein Synthesis. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2016;36(3):1001-7. Epub 2016/01/23. doi: 10.1523/jneurosci.3029-15.2016. PubMed PMID: 26791227; PubMed Central PMCID: PMCPMC4719006.
205. Bou Samra E, Buhagiar-Labarchède G, Machon C, Guitton J, Onclercq-Delic R, Green MR, et al. A role for Tau protein in maintaining ribosomal DNA stability and cytidine deaminase-deficient cell survival. Nature Communications. 2017;8(1):693. doi: 10.1038/s41467-017-00633-1.
206. Ding Q, Markesbery WR, Chen Q, Li F, Keller JN. Ribosome Dysfunction Is an Early Event in Alzheimer's Disease. The Journal of Neuroscience. 2005;25(40):9171-5. doi: 10.1523/jneurosci.3040-05.2005.
207. Van Treeck B, Protter DSW, Matheny T, Khong A, Link CD, Parker R. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. Proceedings of the National Academy of Sciences. 2018. doi: 10.1073/pnas.1800038115.

208. Ganassi M, Mateju D, Bigi I, Mediani L, Poser I, Lee HO, et al. A Surveillance Function of the HSPB8-BAG3-HSP70 Chaperone Complex Ensures Stress Granule Integrity and Dynamism. Molecular Cell. 2016;63(5):796-810. doi: https://doi.org/10.1016/j.molcel.2016.07.021.
209. Ryan VH, Dignon GL, Zerze GH, Chabata CV, Silva R, Conicella AE, et al. Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. Mol Cell. 2018;69(3):465-79 e7. Epub 2018/01/24. doi: 10.1016/j.molcel.2017.12.022. PubMed PMID: 29358076; PubMed Central PMCID: PMCPMC5801700.
210. Aumiller WM, Jr., Keating CD. Phosphorylation-mediated RNA/peptide complex coacervation as a model for intracellular liquid organelles. Nature chemistry. 2016;8(2):129-37. Epub 2016/01/23. doi: 10.1038/nchem.2414. PubMed PMID: 26791895.
211. Wippich F, Bodenmiller B, Trajkovska Maria G, Wanka S, Aebersold R, Pelkmans L. Dual Specificity Kinase DYRK3 Couples Stress Granule Condensation/Dissolution to mTORC1 Signaling. Cell.152(4):791-805. doi: 10.1016/j.cell.2013.01.033.