

DISSERTATION

UNDERSTANDING THE ROLE OF PRION-LIKE DOMAINS IN RIBONUCLEOPROTEIN
GRANULE DYNAMICS

Submitted by

Amy Elizabeth Boncella

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2019

Doctoral Committee:

Advisor: Eric Ross

Co-Advisor: Alan Kennan

Olve Peersen

Chris Ackerson

Copyright by Amy Elizabeth Boncella 2019

All Rights Reserved

ABSTRACT

UNDERSTANDING THE ROLE OF PRION-LIKE DOMAINS IN RIBONUCLEOPROTEIN GRANULE DYNAMICS

Ribonucleoprotein (RNP) granules are membraneless organelles, comprised of RNA-binding proteins and RNA, that are integrally related with the cellular stress response. Stress granules and processing bodies (p-bodies) are the two primary types of RNP granules that reversibly assemble upon stress. Interestingly, many of the proteins that localize to stress granules and p-bodies contain aggregation-prone prion-like domains (PrLDs). Furthermore, mutations in the PrLDs of a number of stress granule-associated proteins have been linked to various neurodegenerative diseases, leading to the idea that aggregation-promoting mutations in these PrLDs cause stress granule persistence. Altogether, these findings suggest an important role for these domains in the dynamics of these assemblies.

In order to gain a greater understanding of how PrLDs contribute to RNP granule biology, I have taken two different approaches. The first was to investigate how aggregation-promoting mutations affect stress granule and p-body dynamics. I introduced various aggregation-promoting mutations into the PrLDs of different stress granule and p-body proteins and assessed the ability of these granules to disassemble, hypothesizing that these mutations would cause RNP granule persistence, as is observed in disease. Interestingly, despite successfully increasing the aggregation propensity of these PrLDs, stress granules and p-bodies do not persist and can efficiently disassemble after stress relief.

Given that aggregation-promoting mutations in PrLDs of RNP granule proteins fail to cause granule persistence, I took a second, less targeted approach towards understanding the

roles of these domains in RNP granules. I focused on investigating how PrLDs are recruited to RNP granules by screening a set of PrLDs for ability to assemble into foci upon stress. Interestingly, many PrLDs are sufficient to assemble into foci upon various stresses, with robust recruitment to stress granules upon heat shock. Furthermore, several compositional biases are observed among PrLDs that are and are not sufficient to assemble upon stress. Using these biases, we have developed a reasonably accurate composition-based predictor of PrLD recruitment into heat shock-induced stress granules, which has been further validated using rational mutation strategies. This predictor is reasonably successful at predicting whether a PrLD will assemble into stress granules upon stress. Additionally, scrambling of PrLD sequences does not disrupt recruitment to stress granules. Together, these results suggest that PrLD localization to stress granules is based on composition rather than primary sequence.

ACKNOWLEDGEMENTS

This journey to a PhD has been a long road and I could not have made it alone. I am deeply grateful for those who have helped me through.

First, I would like to thank all of my fantastic lab mates in the Ross Lab over the years, especially Kacy Paul and Dr. Sean Cascarina, who have both been great resources as well as friends for my entire graduate career. I appreciate you two always being available for questions, scientific discussions, and feedback on presentations and projects. I would like to thank my advisor, Dr. Eric Ross, for instilling in me essential skills to become a successful scientist. I appreciate all of the advice and mentoring you have given me, as well as the opportunity to explore many different and challenging scientific questions.

I would also like to thank my family for always supporting me through all of my successes and failures I've have through the past six years. I would especially like to thank my dad, Jim, for all of the advice, scientific conversations, and encouragement through this difficult endeavor. Thank you for always being such an outstanding scientific role model.

Finally, I would like to thank my wonderful husband, Max, who has faithfully stood by my side throughout. Thank you for always listening, giving advice, encouraging my excitement and curiosity, understanding the struggles, and always inspiring me to be the best scientist that I can be.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iv
CHAPTER ONE: INTRODUCTION.....	1
RNP Granules	1
Stress Granules.....	1
Processing Bodies	4
Comparison of Stress Granules and P-Bodies	5
Mechanisms of RNP Granule Assembly	7
Prion-like Domains in RNP Granules.....	7
Liquid-liquid Phase Separation of RNP Granule Proteins.....	9
Mechanisms of RNP Granule Disassembly	11
Chaperones.....	12
Autophagy and the Ubiquitin-Proteasome System.....	12
Sequence Effects on PrLD Behavior	13
Sequence Effects on LLPS of PrLDs.....	14
Sequence Effects on Aggregation and Amyloid Formation of PrLDs	16
Stress Granules in Disease	17
Disease-Associated Mutations in PrLDs	18
Aberrant LLPS as a Mechanism of Disease-Associated Aggregation.....	18
Defects in PQC Machinery as a Mechanism of Disease-Associated Aggregation.....	19
Understanding PrLD Contributions to RNP Granule Dynamics	21
Investigating the Effects of Aggregation-Promoting Mutations on RNP Granule Dynamics	21
Investigating Sequence Determinants of Stress-Induced PrLD Assembly.....	23
REFERENCES	24
CHAPTER TWO: INVESTIGATING THE EFFECTS OF AGGREGATION-PROMOTING MUTATIONS ON THE P-BODY PROTEIN LSM4	29
Introduction.....	29
Materials and Methods.....	31
Strains and Growth Conditions	31
Cloning Methods.....	32
Mutation Design.....	33
Overexpression and Stress Conditions.....	33
Confocal Microscopy	34
Results.....	34
Aggregation-Promoting Mutations Enhance Aggregation of the Lsm4 PrLD	34
Aggregation-Promoting Mutations Enhance Aggregation of Full-Length Lsm4.....	36
Aggregation-Promoting Mutations in Lsm4 Do Not Disrupt P-body Recovery Post-Stress	39
Deleting Portions of the Lsm4 PrLD Does Not Prevent P-body Assembly	41
Lsm4 Joins P-Bodies Later in Assembly	44
Discussion	47
REFERENCES	51

CHAPTER THREE: UNDERSTANDING THE EFFECTS OF PROLINE DELETIONS ON THE AGGREGATION PROPENSITY OF PRION-LIKE DOMAINS FROM STRESS GRANULE PROTEINS	52
Introduction.....	52
Materials and Methods.....	54
Strains and Growth Conditions.....	54
Cloning Methods.....	55
Overexpression and Stress Conditions.....	56
Confocal Microscopy.....	57
Thioflavin T Staining.....	57
SDD-AGE Methods.....	57
Chronic Heat Shock Assay	58
Results.....	58
Deleting Proline Residues from Stress Granule PrLDs is Predicted to Increase PrLD Aggregation.....	58
Deleting Proline Residues from Stress Granule PrLDs Enhances Foci Formation.....	60
Deleting Proline Residues from Stress Granule PrLDs Enhances Aggregation.....	62
Aggregation-Promoting Mutations Do Not Prevent Localization to Stress Granules.....	64
Deleting Proline Residues from the Pab1 and Pbp1 PrLDs Does Not Prevent Stress Granule Disassembly	66
Deleting Chaperone Proteins Delays Pab1 Recovery After Heat Shock.....	67
Deleting Proline Residues from the Pab1 and Pbp1 PrLDs Does Not Perturb Cell Survival Upon Chronic Heat Shock	71
Discussion.....	71
REFERENCES	78
CHAPTER FOUR: STRESS-INDUCED ASSEMBLY OF PRION-LIKE DOMAINS IS STRONGLY INFLUENCED BY COMPOSITION	80
Introduction.....	80
Materials and Methods.....	82
Strains and Growth Conditions.....	82
Cloning PrLDs	82
Mutation Design.....	82
Stress Conditions	83
Confocal Fluorescence Microscopy.....	83
Western Blotting.....	83
PrLD Composition Analyses	84
Algorithm Generation	84
Results.....	84
PrLDs Form Reversible Assemblies Upon Heat Shock	84
PrLDs Form Stress-Induced Assemblies in Response to Different Stresses.....	91
Composition is the Primary Determinant of Stress-Induced Assembly	91
Modulating Stress-Induced Assembly	102
Discussion.....	109
REFERENCES	116
CHAPTER FIVE: CONCLUSION	119
Aggregation-Promoting Mutations in the Lsm4 PrLD Do Not Disrupt P-body Disassembly	119
Lsm4 Future Directions	120

Aggregation-Promoting Mutations in the PrLDs of Core Stress Granule Proteins Are Not Sufficient to Prevent Stress Granule Disassembly.....	122
Stress Granule PrLD Mutation Future Directions	123
Stress-Induced Assembly of PrLDs	124
PrLD Stress Assembly Future Directions	125
REFERENCES	129
APPENDIX ONE: THE EFFECTS OF MUTATIONS ON THE AGGREGATION PROPENSITY OF THE HUMAN PRION-LIKE PROTEIN HNRNPA2B1	
Introduction.....	130
Materials and Methods.....	134
Yeast Strains and Media	134
Prion Formation in Yeast.....	134
Western Blot	135
Fly Stocks and Culture.....	136
Preparation of Adult Fly Muscle for Immunofluorescence	136
Fly Thoraces Fractionation Protocol.....	136
In Vitro Aggregation Assays	137
Results.....	138
Hydrophobic and Aromatic Residues Promote Aggregation.....	138
Additive and Compensatory Mutations	143
Zipper Segments are Neither Necessary nor Sufficient for Prion Aggregation.....	146
Effects of Mutations in Drosophila.....	148
In Vitro Analysis of Mutants	151
Discussion.....	153
REFERENCES	158

CHAPTER ONE: INTRODUCTION

RNP Granules

How a cell organizes its components is critical for proper function. One way that cells can compartmentalize certain components and reactions is with organelles, which are often bound by membranes. However, recent studies have shown that cells do not necessarily require a membrane to sequester elements, and can instead partition components into membraneless inclusions in both the nucleus and the cytoplasm (1). Many of these membraneless organelles are concentrated in mRNA and protein and have thus been termed ribonucleoprotein (RNP) granules (2). Two of the most common conserved cytoplasmic RNP granule types are stress granules and processing bodies (p-bodies). Stress granules are cytoplasmic inclusions that are mostly composed of translation initiation factors and non-translating mRNAs (3, 4). These granules only form in response to stress and then dissipate once the stressful conditions are eliminated. Using this mechanism, the cell can focus on translating only the factors necessary for survival, but without degrading other mRNAs. Another similar type of RNP granule is a processing body (p-body). P-bodies are inclusions comprised of mRNA as well as decapping and decay factors that are thought to serve as sites of mRNA degradation and storage (5). Although present under normal cellular conditions, p-bodies increase in both size and number upon stress (5, 6). Due to the transient nature of these structures, highly reversible assembly of the components involved is required for normal cellular function. However, the details and mechanisms of these reversible aggregation processes remain unclear.

Stress Granules

Stress granules were initially discovered three decades ago in tomato plant cells exposed to heat shock (7). Upon heat shock, many mRNAs encoding heat shock proteins are

preferentially translated, while other housekeeping mRNAs disassemble from polysomes and are sequestered in granules for the duration of the stress (7). Upon recovery, these non-translating mRNAs are released from these granules and are able to resume translation, indicating that they are being stored, rather than degraded (7). Stress granules are also present in mammalian and yeast cells, although they assemble by different mechanisms in each case (8-10). Mammalian stress granule assembly is initiated by the phosphorylation of eIF2-alpha, whereas yeast stress granule assembly is initiated independently of eIF2-alpha phosphorylation (8, 10). In all cases, protein composition of these stress granules is enriched in translation initiation factors, with assembly affected by reagents that inhibit translation (11, 12). Stress granules are inhibited by reagents that destabilize polysome assembly, such as cycloheximide and emetine, but are unaffected by reagents that stabilize polysome assembly, such as puromycin (13), suggesting that stress granules are composed of stalled translation initiation complexes. Together, all of these results have led to a general hypothesis that stress granules form to sequester housekeeping mRNAs, while freeing up the translation machinery to synthesize chaperones and other proteins capable of mitigating stressful conditions (14).

For years, the complete composition of stress granules remained elusive due to their transient nature as well as lack of a delimiting membrane. Isolating stress granules by standard methods is too harsh to recover all components, especially those only transiently associated. However, development of new techniques has allowed the isolation and identification of many new stress granule components. It was recently discovered that the proteins that are concentrated in the interior of stress granules are stably associated and that these so-called stress granule cores are stable in lysates (15). While the components of these cores can be easily identified using mass spectrometry, this method is too harsh to isolate transiently associated factors. Additional studies have employed *in vivo* labeling strategies to first label and then isolate stress granule

components (16, 17). Because the labeling occurs during the period of stress inside the cell, factors that are transiently associated are also effectively labeled and can thus be successfully extracted and identified. Altogether, these studies have identified a much larger group of proteinaceous stress granule constituents than originally thought. Furthermore, the protein composition of stress granules appears to be dependent on the type of stress the cell is subjected to, as well as cell type (16, 18). Interestingly, many of these newly discovered components of stress granules do not even bind mRNA, suggesting that mRNA binding is not a prerequisite for proteins to associate with stress granules (15). Among newly discovered protein components are members of the protein quality control (PQC) system as well as remodeling ATPases (15, 16). These two systems may function to keep granules in a soluble, reversibly assembling state to ensure that persistent aggregates do not form.

In addition to protein constituents, stress granules also contain diverse RNA components. Although mRNAs make up the majority of the RNA types found in stress granules, ncRNAs and lincRNAs can accumulate within these granules, albeit to a much smaller extent than mRNAs (19). mRNA species within stress granules are quite varied. A wide range of mRNAs are recruited to stress granules, with no particular subset of mRNAs being preferentially recruited (19). Furthermore, different types of stress granules appear to recruit similar mRNAs, indicating that there is some specificity with regard to which mRNAs are able to associate (20).

Stress granules are thought to have a layered structure, with more stably associated proteins forming a core and more loosely associated proteins forming a shell around the core (15). Evidence has shown that stress granule cores form prior to the shells and disassemble last, after the shells have dissipated (21). Consistent with these results, many stress granule proteins interact with one another under normal conditions (16, 17). These pre-existing interactions could potentially nucleate stress granules, possibly as the beginnings of the core structures that have

been observed to form first in stress granule assembly. Although stress granules are thought to only be present upon stress, these data suggest that they might already be present under normal conditions, but are too small to observe, similar to p-bodies (5, 6).

Processing Bodies

A second, closely related type of RNP granule is the processing body or p-body. P-bodies were first discovered in yeast as microscopically visible foci enriched in mRNA decay and decapping factors (22). Upon stress, p-bodies can increase in both size and number (23). Although mRNAs can be degraded at these sites, p-body-associated mRNAs are also able to return to translation, indicating that p-bodies may also function to store these mRNAs (24). Because p-bodies are enriched in mRNA decapping and decay factors, they were originally thought to primarily be sites of mRNA decapping and decay; however, it was later discovered that mRNA decay can also occur outside of these entities in the cytoplasm (25). In fact, mRNA decay is evenly distributed throughout the cytoplasm (26), indicating that, while mRNA decay can occur in p-bodies, they are likely not specific sites of mRNA decay. Additionally, inhibition of p-body formation does not prevent mRNA decay or cell viability (27). Altogether, this data suggests that p-bodies may be redundant sites of mRNA decay, or that perhaps only select mRNAs can be degraded within these granules, yet their exact function remains unknown.


Elucidating the components and structure of p-bodies is one possible way to begin to understand their true function. P-bodies form through interactions involving different proteins and RNAs, with RNA being a requirement for p-body formation (23). Interestingly, no single protein component is necessary for p-body formation (28). In fact, recent studies suggest that p-bodies form by redundant mechanisms, indicating that the summation of interactions is more important than which specific interactions are employed (29). P-bodies have been thought to be composed of fewer and more specific protein components than stress granules (5); however,


recent studies have successfully isolated them to more precisely determine the constituent protein components and have discovered a broader range of constituents than originally thought. One study analyzed p-body components in yeast by immunoprecipitating the p-body protein Dhh1 and using mass spectrometry to identify components under both normal conditions and under conditions of stress (30). This study found that proteins involved in protein folding, nucleotide synthesis, and glycolysis as well as the chaperone protein Ydj1 were all associated with p-bodies upon stress. A second study employed fluorescence-activated particle sorting to isolate and analyze p-bodies in mammalian cells (31). This study verified many known components of p-bodies as well as discovering an enrichment of myosins. Together, these studies may help to uncover the function of p-bodies, by investigating the purpose of these newly discovered p-body components.

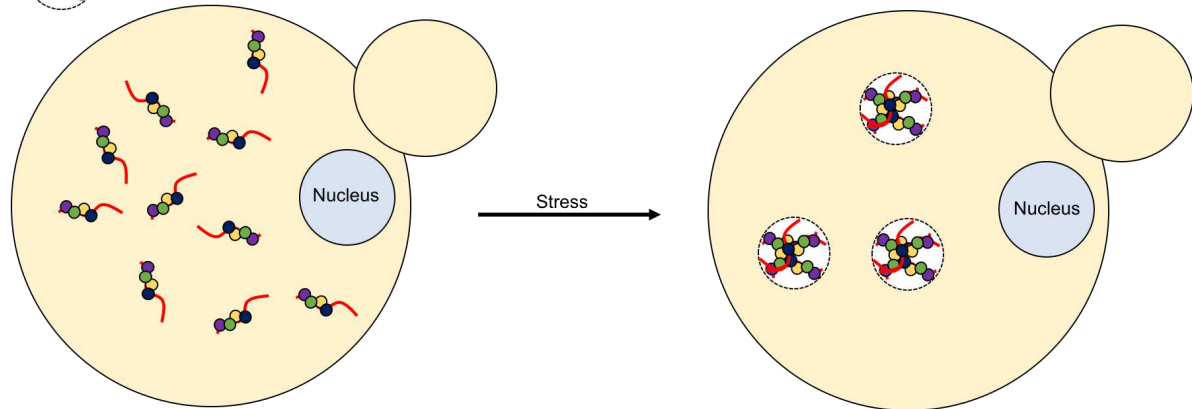
Comparison of Stress Granules and P-Bodies

P-bodies and stress granules are similar cytoplasmic, membraneless organelles, containing non-translating mRNAs, which form upon stress (Figure 1.1). However, p-bodies and stress granules do have some distinct qualities. First, p-bodies are enriched in mRNA decay factors, whereas stress granules are enriched in stalled translation initiation complexes (4, 6), indicating that p-bodies are involved in mRNA degradation, whereas stress granules are likely more suitable for mRNA storage. Additionally, p-bodies can be microscopically observed under normal conditions (22), whereas stress granules are only microscopically visible upon stress (4). However, recent studies have suggested that tiny, incomplete stress granules may be present under normal conditions too, but they cannot be resolved by microscopy (16, 17), a finding that would indicate that p-bodies are at least larger than stress granules under normal conditions. Finally, P-bodies and stress granules can dock, rather than completely merge, with each other during stress, perhaps exchanging factors (9, 32). One possible reason for this behavior is to allow these

A Stress Granules


 = Translation Initiation Complex

 = Stress Granules



B P-bodies

 = mRNA Decay Complex

 = P-body

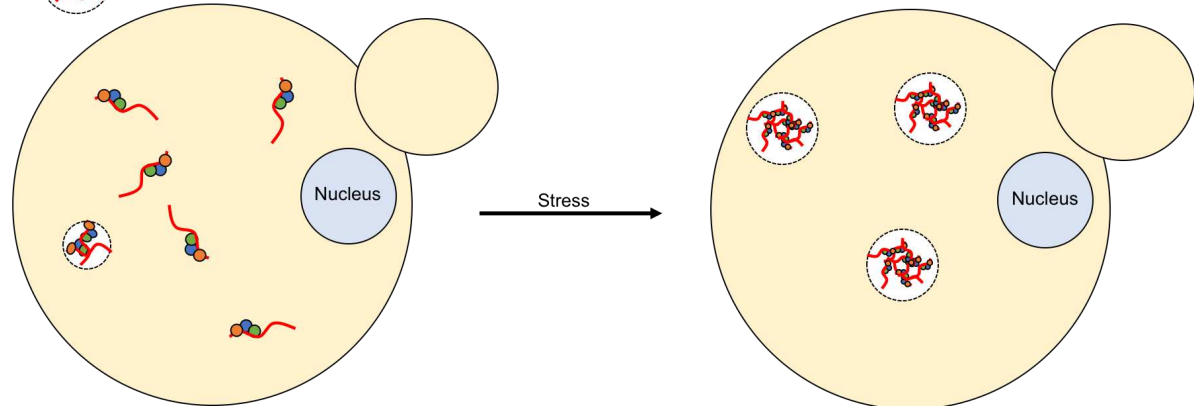


Figure 1.1: Comparing stress granules and p-bodies. (A) Stress granules are composed of stalled translation initiation complexes and are only present upon stress. **(B)** P-bodies are composed of mRNA decay complexes and are present normally, but increase in size and number upon stress.

docked granules to exchange mRNAs for degradation (stress granule to p-body) or exchange mRNAs for storage (p-bodies to stress granules).

Both of these granules appear to form via a functional reversible aggregation mechanism, yet this mechanism is poorly understood. Thus, there is a need to obtain a greater understanding of RNP granule dynamics, assembly, and disassembly mechanisms in order to gain a greater understanding of how cells compartmentalize components.

Mechanisms of RNP Granule Assembly

Although many components of RNP granules have been identified, how they interact and contribute to granule assembly is still poorly understood. The transient nature of these structures necessitates that the interactions involved be able to form and break quickly. Membraneless compartments provide an elegant way for the cell to quickly compartmentalize different components in an energy efficient manner; yet, lack of a delimiting membrane raises the question of how the cell retains specificity within these granules. In order to understand more about the assembly characteristics and structuring of these RNP granules, we explored common sequence features as well as physical characteristics of the protein components involved.

Prion-like Domains in RNP Granules

Analyzing common sequence features and motifs among proteins that localize to RNP granules can provide insight into their assembly and structural characteristics. Many proteins that localize to RNP granules contain one or more intrinsically disordered regions (IDRs) (2). IDRs are regions within a protein that are not predicted to form a stable secondary structure; however, they are still able to contribute nonspecific interactions with other proteins (33). These regions tend to be enriched in charged and polar amino acids and depleted in hydrophobic amino acids (33). One class of IDR that is common among RNP granule proteins is the prion-like domain (PrLD). PrLDs are regions of protein that are predicted to be intrinsically disordered and

additionally contain amino acid compositions similar to those of known yeast prion forming domains (34). They tend to be enriched in glutamine and asparagine residues and depleted in charged and hydrophobic residues (35, 36). While these domains are classified as being similar to prion domains, the two are very distinct. Prions are infectious, whereas prion-like proteins are not; however, both prion forming domains and PrLDs share similar sequence characteristics and can be aggregation-prone. The presence of aggregation-prone PrLDs in RNP granule proteins has led to the idea that these domains may provide important assembly or stabilizing interactions to these granules (37). Indeed, some PrLDs are necessary for targeting certain proteins to different RNP granules (27, 38), suggesting that PrLDs play a role in both assembly as well as specificity of granule components.

In addition to possible roles in protein targeting in granule assembly, PrLDs have the potential to provide many different types of stabilizing interactions within RNP granules, once localized. Many PrLDs are inherently aggregation-prone and multiple studies have shown that PrLDs contribute more promiscuous interactions within these granules, rather than specific ones (39, 40). The Lsm4 and TIA1 PrLDs, which are necessary for p-body and stress granule formation respectively, can be swapped with other PrLDs without hindering RNP granule formation, indicating that composition of these domains, rather than sequence is responsible for localization (27, 38). These results further suggest that the interactions contributed by PrLDs are more promiscuous, rather than specific, in nature.

Within RNP granules, PrLDs may interact with one another homotypically, through self-aggregation, or heterotypically, forming interactions with other PrLDs or IDRs (2). The PrLD of the stress granule-associated protein TDP-43 contains a helical motif within its PrLD that that is required for self-association of the TDP-43 PrLD (41). Another example is the hnRNPA2 PrLD, which can self-interact, but can also induce the assembly of TDP-43, indicating that the PrLD

can interact with other domains heterotypically (42). Additionally, PrLDs can interact with other protein species within RNP granules. One way that PrLDs can form these interactions with other protein species is through short linear motifs (SLiMs) located on the PrLD. These are short regions that provide interaction sites for more structured proteins, and are common among IDRs in p-body proteins (43). One example of this phenomenon is the helical repeats on the IDR of the p-body protein Dcp2, which can interact with a structured domain within the p-body protein Edc3 to promote p-body formation (44). Given that PrLDs are intrinsically disordered, the segments of the domain containing SLiMs are often exposed and thus available for interactions. These motifs also increase the valency of the PrLD-containing protein, thus increasing the number of interactions it has the potential to form. Additionally, IDRs and PrLDs can sometimes interact with RNA molecules (45). Given that RNA is a major component of RNP granule structure, this promiscuity could provide a wealth of possible interaction sites for PrLDs and IDRs to promote strengthening and stabilization of these granules.

Altogether, these interactions between PrLDs and other PrLDs, structured protein domains, or RNAs could be easily summed to help form and structure these RNP granules. The lack of structuring of these PrLDs could enable interactions with other RNP granule components in more nonspecific ways, which could in turn result in interactions that are more easily disrupted, as might be favored for transient structures that need to be able to form and dissipate quickly. More specific and structured interactions might require more energy to disrupt, whereas promiscuous, unstructured interactions are likely less energy intensive to make and break, thus enabling the formation and disassembly of these structures to be rapid and less energy-intensive.

Liquid-liquid Phase Separation of RNP Granule Proteins

The transient nature of RNP granules indicates that these entities may not be classical solid-like aggregates. In agreement with this idea, some RNP granule protein components

display liquid-like properties both *in vitro* and *in vivo* (46, 47). When isolated *in vitro*, many RNP granule proteins, as well as their isolated PrLDs, are capable forming droplets in response to different stimuli, a process called liquid-liquid phase separation (LLPS) (47-52). These protein droplets exhibit properties of liquids including fusion, wetting, and a dynamic interior (52, 53). The concentration at which proteins phase separate is dependent on salt concentration, temperature, pH, as well as crowding agents and RNA (46-51). Together, these observations have led to the theory that RNP granule formation proceeds via a LLPS mechanism of the constituent proteins and RNAs, wherein these species demix from the cytoplasm, and are thus effectively sequestered upon the appropriate environmental conditions.

LLPS is thought to be primarily driven by multivalent interactions among species within membraneless organelles such as RNP granules, and several studies using engineered proteins with modular domain repeats support this theory (1, 54, 55). Phase separation of proteins is a concentration-dependent process (53), and one way to increase the effective protein concentration is to increase the valency of the protein. By providing more interaction sites, and thus increasing the valency, the concentration of protein required for phase separation to occur is effectively lowered (1, 52). PrLDs and IDRs often contain repetitive sequence elements that can serve as interaction sites (44, 47, 56, 57), which increases their valency. Many PrLD- or IDR-containing protein constituents of RNP granules can also bind to RNA through RNA recognition motifs (RRMs). These RNA interaction sites, in addition to multivalent PrLDs or IDRs, result in proteins that are uniquely suited to phase separate in response to environmental changes, such as those induced by certain stress conditions.

The phase separation model of RNP granule assembly is attractive because it potentially explains how proteins and RNA molecules can undergo the reversible aggregation observed in RNP granule biology. If these granules are more liquid-like in nature, then that would indicate

that they could more easily be solubilized, rather than more stable aggregates that require more energy to dissolve (40, 58), thus providing an elegant way for the cell to compartmentalize proteins and RNAs in an energy-efficient, reversible manner. However, most phase separation research has been performed using *in vitro* systems, raising the question of whether or not this phenomenon is relevant to *in vivo* cellular systems. Some experiments, utilizing photobleaching, have demonstrated that these RNP granule assemblies are dynamic (46, 47, 59), but evidence that they are actually behaving as dynamic liquids has remained elusive. The fact that stress granules are structured as a stable core surrounded by a loosely associated shell is only partially consistent with a LLPS mechanism. The LLPS model is compatible with the model of the stress granule shell region in which constituent proteins are loosely associated with the granules. However, the behavior of the core region is less likely to be a LLPS mechanism, given that stress granule cores are stable in lysates and phase separated protein droplets are not (15). Additionally, the fact that certain stress granule proteins already interact normally, prior to the onset of stress conditions (16, 17), argues that the stress granule nucleation step is not an LLPS mechanism. Although the LLPS model is promising, further investigation is required to determine the exact role of this mechanism in RNP granule dynamics and structuring.

Mechanisms of RNP Granule Disassembly

Although plenty of studies have investigated how RNP granules assemble, comparatively little research has focused on understanding the highly important mechanisms of granule disassembly. Without efficient disassembly strategies, RNP granules run the risk of persisting as unwanted aggregates in the cell. RNP granules are enriched in factors associated with protein quality control (60), an example being stress granules, which, along with translation initiation factors, are also enriched in different components of the protein quality control (PQC) machinery (60). These PQC factors are presumably in place to help maintain these granules in a fluid-like

state to prevent formation of persistent, irreversible aggregates. Among the various forms of PQC machinery in the cell, chaperones, autophagy factors, and the ubiquitin-proteasome system (UPS) have all been linked to RNP granule disassembly (60).

Chaperones

When cells are exposed to stress, protein misfolding becomes widespread in the cell (61). One way cells have evolved to cope with this problem is by utilizing chaperone proteins, which aid in the refolding of misfolded proteins (62). Chaperones have been heavily linked to stress granule disassembly, often fulfilling different roles. The Hsp40 chaperones Sis1 and Ydj1 are both implicated in stress granule disassembly after oxidative stress in yeast; however, the two chaperones target stress granules to different fates (63). Sis1 appears to help target stress granules for autophagy, whereas Ydj1 appears to target stress granules for reentry to translation. Another chaperone, Hsp104, is required for stress granule disassembly in response to acute heat shock in yeast (40, 58), possibly functioning to allow translation initiation complexes to exit stress granules and resume translation (58). In mammalian cells, no homolog of Hsp104 exists, but in some cases these cells do appear to rely instead on Hsp70 proteins for stress granule clearance (58). Altogether, chaperone networks appear to work jointly to provide the appropriate fates to stress granule components upon granule disassembly.

Autophagy and the Ubiquitin-Proteasome System

Besides chaperones, autophagy and the ubiquitin-proteasome system (UPS) also contribute to RNP granule clearance. The ubiquitin-binding protein VCP appears to promote stress granule disassembly through targeting to autophagy. Additionally, removal or inhibition of VCP causes stress granule persistence in cells (64). However, given that autophagy is a disposal mechanism, it may not be the preferred pathway of stress granule clearance if the components are to be reused. One possibility is that autophagy may be specifically employed for eliminating

aberrantly aggregated stress granules, whereas chaperones are capable of disassembling normal stress granules without destroying the components (60). Another clearance system, the UPS, has also been implicated in RNP granule disassembly. Inhibition of the UPS in mammalian cells induces stress granule formation, indicating that this system is normally required to prevent accumulation of stress granules and stress granule components (65). Additionally, UBQLN2, a proteasome shuttling protein, was found to be a member of stress granules, possibly functioning by shuttling aberrant proteins out of granules for destruction by the UPS (66). Altogether, it is likely that the combination of chaperones, autophagy, and the UPS function together to disassemble stress granules, as well as normally maintaining these granules in a fluid-like state.

Sequence Effects on PrLD Behavior

To understand the purpose of PrLDs in RNP granule dynamics and assembly, we can examine sequence features. Because amino acid composition and sequence determine the physical nature of protein domains, and because PrLDs are defined partly by their amino acid composition (37), these domains likely have conserved compositional features that in turn drive behavior of these domains. Understanding how the specific sequence features of these PrLDs drive their behavior in the context of RNP granules will help us to better understand the role that these domains play in granule assembly and dynamics. As discussed previously, amino acid composition of PrLDs can contribute to liquid-liquid phase separation, yet amino acid composition also influences other aggregate behavior, such as that underlying amyloid formation (67). We have a robust understanding of the sequence features that underlie the formation of very specific aggregates, such as amyloid (67); however, fewer efforts have been done examining the effects of sequence on LLPS of reversibly aggregating proteins, as may be more applicable to RNP granule formation. Efforts have begun to elucidate the sequence and compositional features

that drive these types of reversible aggregation which are likely more ubiquitous in the cell under normal, physiological conditions.

Sequence Effects on LLPS of PrLDs

Several different sequence and compositional trends have emerged among phase separating IDRs and PrLDs (1, 52). Among these, electrostatic interactions from charged residues, cation- π interactions resulting from patterning of aromatic and arginine residues, dipolar interactions between glutamine, asparagine, and serine residues, as well as hydrophobic interactions all appear to contribute to phase separation (50, 52).

Charged residues are enriched among PrLDs and IDRs that can phase separate, which is indicative of electrostatic interactions modulating LLPS (52). These opposing charges could provide electrostatic interactions that might either facilitate homotypic interactions within a phase separating protein or heterotypic interactions between a protein and another charged molecule (1). Heterotypic interactions provide an elegant way to increase the valency of phase separating proteins. Additionally, patterning of these charges appears to be particularly important (56, 57, 68). Patterning simply refers to blocks of charge (mostly positive or mostly negative) that alternate throughout the protein sequence. For example, the IDR of the germ granule protein DDX4 is capable of phase separating, with LLPS largely modulated by electrostatic interactions (56). However, even when the net charge remains the same, the patterning is essential for LLPS to occur in DDX4 (56). This charge patterning feature has similarly been shown to be a determinant of LLPS in other phase separating proteins (57, 68).

The important contribution of aromatic residues to phase separation was discovered in early phase separation studies, when tyrosine residues were found to be necessary for phase separation and subsequent hydrogel formation of the PrLD of the disease-associated RNA-binding protein FUS (49). It was later discovered that these tyrosine residues in the FUS PrLD

can interact with arginine residues within the RNA-binding region of FUS as a multivalent interaction promoting phase separation (69, 70). These interactions have been proposed to be cation- π interactions between the arginine and tyrosine residues which function to enhance phase separation of FUS (69, 70). Although tyrosine and arginine appear to provide the strongest cation- π pair (69), cation π interactions between phenylalanine and arginine residues have been proposed to promote phase separation of the DDX4 IDR (56), indicating that tyrosine and arginine are not the only cation- π pair capable of promoting phase separation. The importance of aromatic residues contributing to phase separation appears to be quite ubiquitous amongst phase separating IDRs and PrLDs (51, 68, 71, 72), and it is likely that more examples of this interaction influencing LLPS of different proteins will appear.

Another type of interaction that has recently been suggested to underlie phase separation is interactions between low-complexity aromatic-rich kinked segments, or LARKS (73). These motifs consist of aromatic residues, glycine, as well as the polar residues serine, glutamine, and asparagine, which are all residues that are prevalent amongst phase separating PrLDs and IDRs (73). The aromatic and glycine residue patterning provides a kinked beta sheet structure that facilitates interactions between different sheets, yet these interactions are not strong enough to promote the formation of steric zippers, as is observed for amyloid species. In these segments, serine, glutamine, and asparagine residues likely interact via dipole-dipole interactions, with the aromatic residues stacking to form π - π interactions within or between sheets. This result was further validated in a study examining the phase separation determinants of FUS, which demonstrated that glycine promotes fluidity, likely through flexibility, whereas serine and glutamine promote hardening, possibly through dipole-dipole interactions (69).

Hydrophobic residues are another feature that have been shown to be important for phase separation of certain proteins. For example, Pab1 is a yeast stress granule protein that has a large

number of hydrophobic residues in its IDR (50). These residues are thought to promote phase separation by promoting collapse of the protein due to the fact that hydrophobic residues are more likely to be shielded from the hydrophilic cytosol rather than exposed to it. Mutating these residues reduces the ability of this protein to undergo collapse and phase separation and also reduces the ability of yeast to survive chronic stress (50), thus demonstrating how phase separation can contribute to cellular survival.

Altogether, these different sequence features have emerged as being important for phase separation, mostly in some combination with each other. Importantly, none of these interactions are expected to form contacts that are particularly rigid in nature, which is likely necessary to ensure protein solubility and ability to reversibly associate.

Sequence Effects on Aggregation and Amyloid Formation of PrLDs

In contrast to the amorphous, dynamic aggregates that constitute membraneless organelles through phase separation, amyloid aggregates are very structured and thus have more strict requirements dictating which amino acid residues and sequence features are permitted (67). Yeast prion domains provide good examples of how amino acid composition influences protein domain physical behavior and aggregation ability. Yeast prions form amyloids (74), with structural studies indicating that the Sup35 yeast prion forms in-register parallel beta sheets (75). These prion domains are composed of a very standard set of amino acids, which function to dictate their physical properties and behaviors. Yeast prion domains are overrepresented in glutamine and asparagine residues (35) and are depleted in charged and hydrophobic residues (36). Additionally, yeast prion-forming domains are depleted in proline residues. These residues can introduce turns and kinks into protein structures and are thus considered detrimental for proteins that adopt specific structures, such as amyloid. Together, these rules defining prion

domain sequences ensure that these protein domains can adopt an amyloid structure and are able to stably propagate.

Even though PrLDs are compositionally similar to yeast prion domains, the two are very distinct in function. Aggregates formed by PrLDs are not actually transmissible, which distinguishes them from prions. However, because PrLDs are compositionally similar to prion domains, certain sequence features must be different to account for differences in structure and function of the two. It would appear that charged residues in particular have opposing effects on prion formation and LLPS of PrLDs (1, 36). LLPS of PrLD-containing proteins must be very exquisitely tuned to allow for their reversible aggregation and perhaps the prevalence of charged residues helps to prevent any aberrant aggregation of these granules that form by these LLPS mechanisms. Consolidating proteins and RNAs together in granules is inherently risky for the cell. If these structures cannot dissipate once the cellular conditions permit, then aggregates of protein can build up which may either be toxic to the cell or could also be detrimental because of loss of function of proteins that are tied up in these aberrant aggregates.

Stress Granules in Disease

Stress granules have garnered recent interest due to their association with the neurodegenerative diseases amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) (76, 77). A common pathology of these diseases is the accumulation of cytoplasmic inclusions in certain neurons and/or muscle tissues of patients afflicted with these disorders (78). These inclusions stain with known stress granule markers, a finding that suggests that these inclusions are actually persistent stress granules unable to be cleared by the cellular machinery (79, 80). Mutations in several different RNA-binding proteins, specifically TDP-43, FUS, TAF-15, EWSR1, hnRNPA1, hnRNPA2B1, and TIA1, have each been linked to different forms of ALS and/or FTLD through various mutations (81-88). Many of these disease-associated

RNA-binding proteins localize to stress granules *in vivo* normally (38, 81, 89); however, mutations are thought to enhance the aggregation propensities of these proteins, thus enabling them to cause stress granule persistence. Altogether, these data suggest a central role for stress granules and their constituent proteins in these diseases.

Disease-Associated Mutations in PrLDs

Interestingly, many disease-associated mutations occur in the PrLDs of these RNA-binding proteins (81, 90). Given that PrLDs are naturally aggregation-prone, mutations in these domains might be uniquely poised convert these otherwise soluble proteins to insoluble aggregates. PrLDs of some stress granule proteins may lie on the edge of aggregation such that they can still reversibly associate with stress granules under a normal stress response, but, upon mutation, these proteins are pushed over the edge of aggregation, leading to the accumulation of irreversibly associated cytoplasmic inclusions. This situation appears to be true for the stress granule RNA-binding protein hnRNPA2. In its normal form hnRNPA2 is soluble; however, a single disease-associated mutation of an aspartic acid residue to a valine residue within the PrLD promotes hnRNPA2 to an insoluble aggregate that accumulates in the cytoplasm (81). Despite these observations, whether these mutations are actually the sole cause of pathological stress granule formation remains to be determined. Other factors may contribute to aggregate formation and persistence, as detailed in the following sections.

Aberrant LLPS as a Mechanism of Disease-Associated Aggregation

Many of these disease-associated RNA-binding proteins, namely FUS, TDP-43, hnRNPA1, hnRNPA2, and TIA1, are capable of undergoing LLPS *in vitro* under various conditions (41, 42, 46, 47, 88), with some demonstrating liquid-like qualities *in vivo* (46, 47). Interestingly, when disease-associated mutations are introduced to the PrLDs of these proteins, these normally soluble proteins tend to lose their liquid-like properties and become more solid-

like as assessed by various biochemical techniques (41, 42, 46-48, 91). These findings are consistent with the idea that these disease-associated mutations transform these proteins from soluble species to insoluble aggregates, thus causing stress granules to transition from highly dynamic, liquid-like states to stably associated aggregates or fibers (Figure 1.2). Additionally, *in vitro* aging experiments suggest that the liquid-like protein droplets of these proteins can mature over time and lose their liquid-like properties, with disease-associated mutations accelerating this process (46, 48). These results indicate that aging plays a role in aggregation of these proteins, which is consistent with the late onset of the associated neurodegenerative diseases. Together, these findings identify LLPS as a mechanism that may promote aberrant stress granule formation by sequestering aggregation-prone proteins together in the cell.

Defects in PQC Machinery as a Mechanism of Disease-Associated Aggregation

As mentioned above, many of these stress granule-associated neurodegenerative diseases are late onset, occurring in the elderly population (78). Another factor that could play a role in disease progression is the degradation of the protein quality control (PQC) machinery, which is a process that occurs slowly over time and causes cells to lose their ability to combat protein misfolding and aggregation (92). One explanation for why stress granules persist in disease is that disease-associated stress granules can be efficiently cleared by either autophagy or the proteasome normally, however, upon degradation of this defense machinery, the PQC can no longer keep up with disposal of the aberrant cytoplasmic inclusions that appear (93, 94).

Additionally, mutations in some proteins related to the PQC machinery are linked to neurodegenerative diseases. VCP is an ATPase which is linked to ALS when mutated (95). Normally, VCP acts to target stress granules for autophagy; however, when disease-associated mutations are present, stress granules persist (64). Additionally, recent evidence suggests that the proteasome shuttle protein UBQLN2 also plays a role in stress granule dynamics, possibly

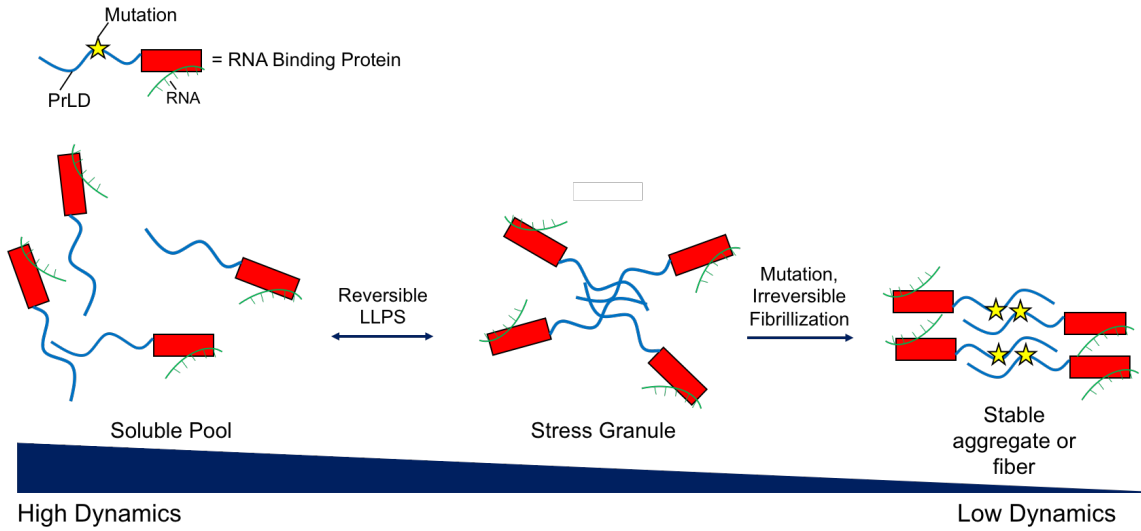


Figure 1.2: Disease-associated mutations promote transition of stress granules to stable aggregates. Normally, stress granule-associated RNA-binding proteins exist in a highly dynamic, soluble pool and can undergo LLPS to reversibly assemble into stress granules. Upon mutation in PrLDs of stress granule proteins, stress granules can undergo an irreversible transition to a stable aggregate or fiber.

shuttling components out of stress granules to the proteasome for degradation (66, 96). Similar to VCP, UBQLN2 is also linked to ALS upon mutation (97), and these mutations hinder stress granule dynamics (96). Interestingly, UBQLN2 is prone to phase separation, with mutations promoting conversion to aberrant aggregation (66, 96, 98), suggesting that in addition to losing its normal function in keeping stress granules clear of aberrant components, it can itself be a cause of stress granule persistence. Ultimately, poorly functioning PQC machinery appears to prevent cells from effectively eliminating persistent aggregates, which could lead to disease.

All of these mechanisms likely play roles in the onset and progression of these neurodegenerative diseases, either in combination, or separately. Further investigations into RNP granule dynamics as well as phase separation mechanisms will help us understand how these diseases manifest. This knowledge can then be used in the development of therapeutics towards combating disease.

Understanding PrLD Contributions to RNP Granule Dynamics

Although plenty of work has been done to advance our understanding of RNP granule biology, there remain gaps in our knowledge with regard to the function and purpose of these entities, as well as their role in disease. The goal of this work is to gain insight into the role of PrLDs in RNP granule biology and, more specifically, to determine how the composition of these domains affects granule dynamics. Here, I have made progress towards understanding the compositional features that govern association of PrLDs to RNP granules in yeast using different systems studying both p-bodies and stress granules.

Investigating the Effects of Aggregation-Promoting Mutations on RNP Granule Dynamics

Initially I wanted to examine how aggregation-promoting mutations made in the PrLDs of RNP granule proteins would affect granule dynamics. I made mutations in the PrLD of the yeast p-body protein Lsm4 that were designed to increase its aggregation propensity. Because the

Lsm4 PrLD is necessary for p-body formation in some cases (27), I reasoned that aggregation-promoting mutations in its PrLDs might cause hyper-aggregation of p-bodies, thus preventing proper disassembly. These mutations were designed by inserting amino acids that are known to promote aggregation and prion formation in yeast (36). Initial observations of the mutant PrLDs in isolation showed that the mutations introduced successfully increased aggregation of the domain; however, when these mutations were introduced into full-length Lsm4 in an endogenous system, p-body disassembly was unaffected. Additionally, I discovered that the Lsm4 PrLD is not absolutely necessary for p-body formation, contrary to what was previously reported (27), possibly explaining why my aggregation-promoting mutations failed to inhibit p-body disassembly.

Following these experiments, I sought out a different system in which to examine the effects of aggregation-promoting mutations in RNP granule protein PrLDs. Stress granules are more solid-like in yeast than p-bodies (40), so I reasoned that introducing aggregation-promoting mutations into stress granule PrLDs might be more effective at causing stress granule components to hyper-aggregate than introducing these mutations into p-body components. I observed that several stress granule PrLDs and IDRs are enriched in proline residues. Given that proline is detrimental to amyloid formation (99), I hypothesized that these residues might be conserved to help prevent aberrant aggregation of stress granule proteins. To test this hypothesis, I deleted the proline residues from the PrLDs of the core stress granule proteins Pab1 and Pbp1, which both normally contain an overrepresentation of proline residues in their PrLDs. Similar to the Lsm4 results, these proline deletions successfully enhanced aggregation of these PrLDs in isolation; however, upon introduction of these mutations into an endogenous system, stress granule dynamics remained unaffected.

Investigating Sequence Determinants of Stress-Induced PrLD Assembly

These two unexpected results suggest that PrLD association to RNP granules may be governed by different interactions than prion-like aggregation. To determine the compositional requirements for PrLD association to stress granules, I screened a selection of 56 yeast PrLDs for their ability to reversibly localize to stress granules upon different stresses, and then analyzed the compositions of the domains that were and were not sufficient to localize. We found several compositional biases among domains that localize to stress granules and interestingly, many of the compositional features that promote assembly into stress granules are features that are underrepresented among prion domains, such as hydrophobic and charged residues (36). Using these biases, we created a predictor which allowed us to predict with reasonable accuracy which PrLDs would localize to stress granules and which would not. Additionally, these biases were sufficient to guide the design of rational mutations of different PrLDs to either promote or prevent assembly into foci. Finally, scrambling the sequences of different PrLDs did not affect their ability to localize to stress granules, suggesting that composition, rather than primary sequence, is responsible for PrLD recruitment to stress granules.

REFERENCES

1. Banani SF, Lee HO, Hyman AA, & Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol*.
2. Mittag T & Parker R (2018) Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly. *J Mol Biol*.
3. Buchan JR & Parker R (2009) Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* 36(6):932-941.
4. Protter DS & Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26(9):668-679.
5. Parker R & Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Mol Cell* 25(5):635-646.
6. Luo Y, Na Z, & Slavoff SA (2018) P-Bodies: Composition, Properties, and Functions. *Biochemistry*.
7. Nover L, Scharf KD, & Neumann D (1989) Cytoplasmic heat shock granules are formed from precursor particles and are associated with a specific set of mRNAs. *Mol Cell Biol* 9(3):1298-1308.
8. Kedersha NL, Gupta M, Li W, Miller I, & Anderson P (1999) RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *J Cell Biol* 147(7):1431-1442.
9. Buchan JR, Muhrad D, & Parker R (2008) P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 183(3):441-455.
10. Hoyle NP, Castelli LM, Campbell SG, Holmes LE, & Ashe MP (2007) Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. *J Cell Biol* 179(1):65-74.
11. Kedersha N, *et al.* (2002) Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol Cell* 13(1):195-210.
12. Kimball SR, Horetsky RL, Ron D, Jefferson LS, & Harding HP (2003) Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. *American journal of physiology. Cell physiology* 284(2):C273-284.
13. Kedersha N, *et al.* (2000) Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J Cell Biol* 151(6):1257-1268.
14. Anderson P & Kedersha N (2006) RNA granules. *J Cell Biol* 172(6):803-808.
15. Jain S, *et al.* (2016) ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164(3):487-498.
16. Markmiller S, *et al.* (2018) Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell* 172(3):590-604 e513.
17. Youn JY, *et al.* (2018) High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol Cell* 69(3):517-532 e511.
18. Buchan JR, Yoon JH, & Parker R (2011) Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* 124(Pt 2):228-239.
19. Khong A, *et al.* (2017) The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Mol Cell* 68(4):808-820 e805.
20. Namkoong S, Ho A, Woo YM, Kwak H, & Lee JH (2018) Systematic Characterization of Stress-Induced RNA Granulation. *Mol Cell* 70(1):175-187 e178.

21. Wheeler JR, Matheny T, Jain S, Abrisch R, & Parker R (2016) Distinct stages in stress granule assembly and disassembly. *eLife* 5.
22. Sheth U & Parker R (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300(5620):805-808.
23. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, & Parker R (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 11(4):371-382.
24. Brengues M, Teixeira D, & Parker R (2005) Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* 310(5747):486-489.
25. Eulalio A, Behm-Ansmant I, Schweizer D, & Izaurralde E (2007) P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* 27(11):3970-3981.
26. Tutucci E, *et al.* (2017) An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat Methods*.
27. Decker CJ, Teixeira D, & Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179(3):437-449.
28. Teixeira D & Parker R (2007) Analysis of P-body assembly in *Saccharomyces cerevisiae*. *Mol Biol Cell* 18(6):2274-2287.
29. Rao BS & Parker R (2017) Numerous interactions act redundantly to assemble a tunable size of P bodies in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*.
30. Cary GA, Vinh DB, May P, Kuestner R, & Dudley AM (2015) Proteomic Analysis of Dhh1 Complexes Reveals a Role for Hsp40 Chaperone Ydj1 in Yeast P-Body Assembly. *G3 (Bethesda)* 5(11):2497-2511.
31. Hubstenberger A, *et al.* (2017) P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Mol Cell* 68(1):144-157 e145.
32. Kedersha N, *et al.* (2005) Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* 169(6):871-884.
33. van der Lee R, *et al.* (2014) Classification of intrinsically disordered regions and proteins. *Chem Rev* 114(13):6589-6631.
34. Li L, McGinnis JP, & Si K (2018) Translational Control by Prion-like Proteins. *Trends Cell Biol*.
35. DePace AH, Santoso A, Hillner P, & Weissman JS (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93(7):1241-1252.
36. Toombs JA, McCarty BR, & Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-332.
37. March ZM, King OD, & Shorter J (2016) Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res*.
38. Gilks N, *et al.* (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15(12):5383-5398.
39. Protter DSW, *et al.* (2018) Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* 22(6):1401-1412.
40. Kroschwald S, *et al.* (2015) Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife* 4:e06807.

41. Conicella AE, Zerze GH, Mittal J, & Fawzi NL (2016) ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure*.
42. Ryan VH, *et al.* (2018) Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell* 69(3):465-479 e467.
43. Jonas S & Izaurralde E (2013) The role of disordered protein regions in the assembly of decapping complexes and RNP granules. *Genes Dev* 27(24):2628-2641.
44. Fromm SA, *et al.* (2014) In vitro reconstitution of a cellular phase-transition process that involves the mRNA decapping machinery. *Angew Chem Int Ed Engl* 53(28):7354-7359.
45. Hentze MW, Castello A, Schwarzl T, & Preiss T (2018) A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol* 19(5):327-341.
46. Patel A, *et al.* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162(5):1066-1077.
47. Molliex A, *et al.* (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* 163(1):123-133.
48. Lin Y, Protter DS, Rosen MK, & Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60(2):208-219.
49. Kato M, *et al.* (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149(4):753-767.
50. Riback JA, *et al.* (2017) Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* 168(6):1028-1040 e1019.
51. Xiang S, *et al.* (2015) The LC Domain of hnRNPA2 Adopts Similar Conformations in Hydrogel Polymers, Liquid-like Droplets, and Nuclei. *Cell* 163(4):829-839.
52. Brangwynne Clifford P, Tompa P, & Pappu Rohit V (2015) Polymer physics of intracellular phase transitions. *Nature Physics* 11(11):899-904.
53. Hyman AA, Weber CA, & Julicher F (2014) Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* 30:39-58.
54. Li P, *et al.* (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483(7389):336-340.
55. Banani SF, *et al.* (2016) Compositional Control of Phase-Separated Cellular Bodies. *Cell* 166(3):651-663.
56. Nott TJ, *et al.* (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57(5):936-947.
57. Elbaum-Garfinkle S, *et al.* (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* 112(23):7189-7194.
58. Cherkasov V, *et al.* (2013) Coordination of translational control and protein homeostasis during severe heat stress. *Curr Biol* 23(24):2452-2462.
59. Bolognesi B, *et al.* (2016) A Concentration-Dependent Liquid Phase Separation Can Cause Toxicity upon Increased Protein Expression. *Cell Rep* 16(1):222-231.
60. Alberti S, Mateju D, Mediani L, & Carra S (2017) Granulostasis: Protein Quality Control of RNP Granules. *Front Mol Neurosci* 10:84.
61. Morimoto RI (2011) The heat shock response: systems biology of proteotoxic stress in aging and disease. *Cold Spring Harb Symp Quant Biol* 76:91-99.
62. Richter K, Haslbeck M, & Buchner J (2010) The heat shock response: life on the verge of death. *Mol Cell* 40(2):253-266.

63. Walters RW, Muhlrad D, Garcia J, & Parker R (2015) Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in *Saccharomyces cerevisiae*. *RNA* 21(9):1660-1671.
64. Buchan JR, Kolaitis RM, Taylor JP, & Parker R (2013) Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* 153(7):1461-1474.
65. Mazroui R, Di Marco S, Kaufman RJ, & Gallouzi IE (2007) Inhibition of the ubiquitin-proteasome system induces stress granule formation. *Mol Biol Cell* 18(7):2603-2618.
66. Dao TP, *et al.* (2018) Ubiquitin Modulates Liquid-Liquid Phase Separation of UBQLN2 via Disruption of Multivalent Interactions. *Mol Cell*.
67. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry* 75:333-366.
68. Pak CW, *et al.* (2016) Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* 63(1):72-85.
69. Wang J, *et al.* (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*.
70. Qamar S, *et al.* (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- π Interactions. *Cell* 173(3):720-734 e715.
71. Lin Y, Currie SL, & Rosen MK (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem*.
72. Jiang H, *et al.* (2015) Phase transition of spindle-associated protein regulate spindle apparatus assembly. *Cell* 163(1):108-122.
73. Hughes MP, *et al.* (2018) Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. *Science* 359(6376):698-701.
74. Liebman SW & Chernoff YO (2012) Prions in yeast. *Genetics* 191(4):1041-1072.
75. Shewmaker F, Wickner RB, & Tycko R (2006) Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure. *Proc Natl Acad Sci U S A* 103(52):19754-19759.
76. Ramaswami M, Taylor JP, & Parker R (2013) Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154(4):727-736.
77. Li YR, King OD, Shorter J, & Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201(3):361-372.
78. Taylor JP, Brown RH, & Cleveland DW (2016) Decoding ALS: from genes to mechanism. *Nature* 539(7628):197-206.
79. Liu-Yesucevitz L, *et al.* (2010) Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* 5(10):e13250.
80. Dewey CM, *et al.* (2012) TDP-43 aggregation in neurodegeneration: are stress granules the key? *Brain Res* 1462:16-25.
81. Kim HJ, *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495(7442):467-473.
82. Neumann M, *et al.* (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314(5796):130-133.
83. Arai T, *et al.* (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun* 351(3):602-611.
84. Vance C, *et al.* (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323(5918):1208-1211.

85. Kwiatkowski TJ, Jr., *et al.* (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323(5918):1205-1208.
86. Couthouis J, *et al.* (2011) A yeast functional screen predicts new candidate ALS disease genes. *Proc Natl Acad Sci U S A* 108(52):20881-20890.
87. Neumann M, *et al.* (2011) FET proteins TAF15 and EWS are selective markers that distinguish FTLN with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. *Brain* 134(Pt 9):2595-2609.
88. Mackenzie IR, *et al.* (2017) TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* 95(4):808-816 e809.
89. Colombrita C, *et al.* (2009) TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem* 111(4):1051-1061.
90. King OD, Gitler AD, & Shorter J (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* 1462:61-80.
91. Murakami T, *et al.* (2015) ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* 88(4):678-690.
92. Klaips CL, Jayaraj GG, & Hartl FU (2018) Pathways of cellular proteostasis in aging and disease. *J Cell Biol* 217(1):51-63.
93. Alberti S & Carra S (2018) Quality Control of Membraneless Organelles. *J Mol Biol.*
94. Alberti S & Hyman AA (2016) Are aberrant phase transitions a driver of cellular aging? *Bioessays.*
95. Johnson JO, *et al.* (2010) Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68(5):857-864.
96. Alexander EJ, *et al.* (2018) Ubiquilin 2 modulates ALS/FTD-linked FUS-RNA complex dynamics and stress granule formation. *Proc Natl Acad Sci U S A* 115(49):E11485-E11494.
97. Deng HX, *et al.* (2011) Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477(7363):211-215.
98. Sharkey LM, *et al.* (2018) Mutant UBQLN2 promotes toxicity by modulating intrinsic self-assembly. *Proc Natl Acad Sci U S A* 115(44):E10495-E10504.
99. Theillet FX, *et al.* (2013) The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord Proteins* 1(1):e24360.

CHAPTER TWO: INVESTIGATING THE EFFECTS OF AGGREGATION-PROMOTING MUTATIONS ON THE P-BODY PROTEIN LSM4

Introduction

Mutations in a number of RNA-binding proteins have been linked to various neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS), frontotemporal lobar degeneration (FTLD), and multisystem proteinopathy (MSP) (1, 2). These disease-associated RNA-binding proteins are also associated with ribonucleoprotein (RNP) granule formation, of which there are two main types: stress granules and processing bodies (3). Stress granules are assemblies of translation initiation factors bound to non-translating mRNAs that form when the cell is subjected to various forms of stress, thus allowing the cell to focus on translating only the factors necessary for survival (4). P-bodies are inclusions comprised of mRNA and degradation factors that are thought to serve as sites of decapping and decay of mRNA (5, 6). Although observed in the cell normally, p-body size and number increase upon stress (7). Given that these RNP granules must be able to form and dissipate quickly, dynamic and reversible association of the constituent proteins is essential to functionality. Current research efforts suggest that when this process becomes irreversible, protein aggregates accumulate and can cause disease (1, 2). Although so far all of the disease-associated mutations that are linked to RNP granules are linked to stress granule proteins, it remains possible that p-body proteins may also be prone to this type of aberrant aggregation given that they assemble similarly.

Aggregation of these RNA-binding proteins appears to be a requirement for the formation of cytoplasmic mRNP granules and an emerging area of research is directed towards identifying those regions of a given stress granule or p-body protein that are required for aggregation. These regions must be able to coalesce upon certain conditions, but also must be able to disassemble once the stress has been removed. Several stress granule and p-body proteins contain

aggregation-prone regions termed prion-like domains (PrLDs) (8, 9). These domains are defined as having amino acid compositions similar to those of known yeast prions, with high glutamine and asparagine content (2). In some cases, p-body and stress granule formation is dependent on the PrLD of certain proteins (10, 11). Interestingly, many disease-linked mutations lie in the PrLDs of stress granule proteins (12-14), leading to the idea that these mutations promote the transition of these stress granules from dynamic, reversible aggregates, to toxic, irreversible inclusions. These observations have sparked interest in learning more about these aggregation-prone PrLDs and the features that promote them to a hyper-aggregated state.

Given that many disease-relevant RNA-binding proteins contain PrLDs that resemble yeast prions, it is reasonable to predict that the sequence features that enhance prion formation and aggregation might also enhance aggregation of PrLDs in p-body or stress granule proteins. Our lab previously developed an algorithm, called the Prion Aggregation Prediction Algorithm (PAPA), to predict the prion propensity of glutamine/asparagine (Q/N)-rich sequences in yeast (15). Noticeably, aromatic and hydrophobic amino acids tend to increase prion propensity, whereas proline and charged residues tend to inhibit prion formation (15). This algorithm was used to successfully predict the behavior of three disease-associated mutations in the PrLDs of the stress granule proteins hnRNPA1 and hnRNPA2B1 (13), all of which cause the protein to transition from a soluble state to the disease-associated insoluble state. This result suggests that the features promoting prion formation might similarly be able to promote transitions of PrLD-containing proteins from soluble to insoluble states, possibly affecting RNP granule dynamics.

Here, I have investigated how mutations made in the PrLD of the yeast decapping activator Lsm4 alter the dynamics of p-body assembly and disassembly. Lsm4 was chosen because it localizes to p-bodies and contains a C-terminal Q/N-rich PrLD (9, 11). Additionally, the Lsm4 PrLD is required for P-body formation when the scaffolding p-body protein Edc3 is

knocked out (11). Together, these attributes provide an ideal system to assess how mutations in a PrLD of an RNP granule protein might affect RNP granule dynamics. Using the prion prediction algorithm, PAPA, mutations were inserted to either enhance or inhibit the aggregation propensity of the Lsm4 PrLD, and the effects on aggregation and p-body dynamics were observed. I initially designed two aggregation-promoting mutants and one aggregation-inhibiting PrLD mutant and fused them to GFP. Upon overexpression, the aggregation-promoting mutants formed foci faster than the wt Lsm4 PrLD, indicating that aggregation propensity was successfully increased. However, when these mutations were introduced into the full-length protein under endogenous expression, p-body dynamics were not disrupted. Upon further analysis of the Lsm4 PrLD I discovered that the Lsm4 PrLD does not appear to prevent p-body formation entirely, thus possibly explaining why the aggregation-promoting mutations introduced to the PrLD did not prevent p-body disassembly.

Materials and Methods

Strains and Growth Conditions

All experiments were performed in *Saccharomyces cerevisiae* using standard yeast handling and growth conditions (16). All plasmids were transformed into different yeast strains using standard yeast transformation protocols (16). Overexpressed Lsm4 and Lsm4 mutant PrLDs were transformed into YER826 (*α kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 sup35::KanMx [pER589(URA3)]*), overexpressed full-length Lsm4 and Lsm4 mutants were transformed into YER632 (*α kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 sup35::KanMx [pJ533(URA3)]*), and all of the untagged or mCherry-tagged, endogenously expressed Lsm4 and Lsm4 mutants were transformed into YER988 (*MATa lsm4::NEO edc3::NEO DCP2GFP::NEO leu2 lys2 trp1 ura3 [Lsm4pRS416-pRP1548]*), which is a strain shared with us from the Roy Parker Lab. YER988 contains a chromosomal knockout of Lsm4, but since this gene is essential,

a copy of *Lsm4* is maintained on a *URA3* maintenance plasmid. Each plasmid containing a *Lsm4* mutant was shuffled into this strain with the maintenance plasmid shuffled out using FOA counterselection.

Cloning Methods

The overexpressed PrLDs were cloned into pER760, which is a derivative of YePlac112 (17) containing the GAL1 promoter followed by GFP. Each construct was built using PCR with the mutated region synthetically built using overlapping primers containing the mutations. A start codon was inserted before each PrLD to ensure expression. These constructs were cloned in between the GAL1 promoter and GFP in between the BamHI and XhoI restriction sites in pER760 using standard restriction enzyme cloning procedures.

Full-length overexpressed *Lsm4* and mutant versions were constructed from *Lsm4* which was first amplified from the genome. Mutations were introduced using overlapping primers to synthetically build the mutated region of each protein. These constructs were cloned into pER760 in between the BamHI and XhoI restriction sites, using the same cloning methods described above for the overexpressed PrLDs

Full-length, endogenously expressed *Lsm4* +FY, +3F3Y, and Δ 35 mutants were constructed using two-piece PCR with an overlapping region between the two pieces. The first piece consisted of the *Lsm4* promoter and the N-terminal portion of *Lsm4* preceding the PrLD. The second piece contained the *Lsm4* terminator and the C-terminal *Lsm4* PrLD, which was synthetically built with overlapping primers to introduce the mutations. After PCR amplification, these *Lsm4* constructs were then cloned into pER857, a derivative of YcPlac111 (17) between the PstI and EcoRI restriction sites using standard restriction enzyme cloning procedures.

Full-length, endogenously expressed, untagged Δ 1st 50, Δ Mid 50, Δ Last 50, and Δ C were PCR amplified from genomic yeast DNA to include the *Lsm4* promoter, but to also delete the

designated regions, using overlapping primers. Once amplified, these products were inserted into pER1318 (derivative of pER857 containing the Lsm4 terminator between the BamHI and EcoRI restriction sites) between the PstI and BamHI restriction sites using standard restriction enzyme cloning procedures.

Full-length, endogenously expressed, mCherry-tagged Lsm4 and Lsm4 deletion mutant constructs were made by PCR amplifying the promoter region and Lsm4 from yeast genomic DNA, with primers designed to delete the appropriate regions for each construct. These PCR products were cloned into pER1201 (derivative of pER857 containing the mCherry- T_{Lsm4} cassette between the XhoI and EcoRI restriction sites) between the BamHI and XhoI restriction sites using standard restriction enzyme cloning procedures.

Mutation Design

To design the +FY and +3F3Y aggregation-enhancing mutants, F and Y residues were randomly positioned in the PrLD using the Excel Random Number generator to determine the amino acid after which each F or Y residue would be inserted.

Overexpression and Stress Conditions

For overexpression of PrLDs and full-length Lsm4 constructs, cells were grown in SGal/Raff -Trp +Ade media (to select for the plasmids) for 2, 4, or 21 hours to induce the *GALI* promoter to overexpress each construct.

To perform p-body assembly experiments, cells were grown to $OD_{600} \approx 0.3-0.7$ at 30°C in YPA + 2% glucose prior to stress induction. For glucose deprivation, 1mL of cells was harvested and washed once in SC-glucose media before being resuspended in 1mL SC-glucose and incubated at 30°C with shaking for 10-40 minutes prior to imaging. To recover from glucose deprivation, Cells were harvested by centrifugation and SC-glucose media was exchanged for SC -Leu media before being returned to 30°C with shaking for 20 minutes before imaging.

Confocal Microscopy

Cells were imaged using an Olympus IX83 confocal spinning disk microscope. Images were collected as single planes.

Results

Aggregation-Promoting Mutations Enhance Aggregation of the Lsm4 PrLD

Mutations were initially designed to modulate the aggregation propensity of the Lsm4 PrLD using PAPA to guide mutation selection. Given that the Lsm4 PrLD is necessary for p-body formation (11) we reasoned that this assembly process might proceed via prion-like aggregation. If this is the case, the sequence features promoting prion formation should also enhance p-body assembly, possibly reducing dynamics. The aggregation-enhancing mutations were introduced as insertions rather than substitutions to ensure that no effect of any particular amino acid in the PrLD was lost. To increase aggregation propensity, the amino acid residues phenylalanine (F) and tyrosine (Y) were chosen as additions because they both increase prion propensity (15). Two mutants were constructed that were designed to enhance aggregation: one a subtle mutation inserting an F and Y residue, denoted +FY and the other a more aggressive mutation inserting 3 F and 3 Y residues, denoted +3F3Y (Table 2.1). A third mutant was designed to decrease Lsm4 aggregation propensity and consisted of a deletion of the region of the Lsm4 PrLD containing the 35 most prion-promoting amino acids, denoted as $\Delta 35$ (Table 2.1). The Lsm4 PrLD is already predicted to be prion-like by PAPA; however, addition of aromatic residues was predicted to further increase the prion-like character of the Lsm4 PrLD (Table 2.1). Conversely, removing the most prion-like region of the Lsm4 PrLD ($\Delta 35$ mutant) was predicted to decrease prion-like character of the domain (Table 2.1).

All of the mutant PrLDs, as well as the wild-type (wt) Lsm4 PrLD, were expressed from plasmids under the strong, inducible *GALI* promoter and were fused to the N-terminus of GFP

Table 2.1: Mutations introduced to the Lsm4 PrLD.

Mutation	PAPA Score	Sequence
Wild-type Lsm4 PrLD	0.12	QQINSNNNSNSNGPGHKRYYNRDSNNNRGNYNRRNNNNNGNSNRRPY SQNRQYNNSSSNINNSINSINSNNQNMNGLGGSVQHHFNSSSPQKVE F
Insertion of two aromatic residues (+FY)	0.15	QQINSNNNSNSNGPGHKRYYNRDSNNNRGNYNRRNNNNNGNSNRRPY SQNRQYNNSSSNINNN F Y INSINSNNQNMNGLGGSVQHHFNSSSPQK VEF
Insertion of six aromatic residues (+3F3Y)	0.21	QQINSNNNSNSNGPGHKRYYNRDSNNNRGNYNRRNNNNNGNSNRRPY SQNRQYNNSSSN Y NNNSINS Y SN F NQ Y M F N N G F LGGSVQHHFNSS SPQKVEF
Deletion of 35 prion-promoting residues (Δ35)	-0.01	QQINSNNNSNSNGPGHKRYYNRDSNNNRGNYNRRNNNNNGNSNRRPY SQNRQNSSSPQKVEF

PAPA score listed is for the entire Lsm4 protein sequence with the indicated mutations introduced into the PrLD. Scores above 0.05 are predicted to be prion-like, and scores below 0.05 are not predicted to be prion-like. Sequences of the Lsm4 PrLD are shown with red amino acid residues indicating where aromatic residues were inserted.

for visualization. To determine whether or not the mutations increased the kinetics of aggregation of the Lsm4 PrLD, each PrLD was overexpressed for 4 and 21 hours before imaging. After 4 hours of overexpression, the +3F3Y mutant formed distinct foci, indicative of aggregation (Figure 2.1). While some foci formed for the +FY mutant, most of the GFP signal remained diffuse after 4 hours (Figure 2.1). In contrast, the wt Lsm4 PrLD and the $\Delta 35$ mutant remained diffuse at the 4-hour timepoint (Figure 2.1). After 21 hours of overexpression, the +3F3Y mutant formed similar, but slightly larger foci and the +FY mutant appeared more concentrated in foci than at the 4-hour timepoint (Figure 2.1). In contrast, the Lsm4 PrLD and the $\Delta 35$ PrLD still remained diffuse after 21 hours (Figure 2.1), indicating that overexpression was not inducing aggregation. GFP remained diffuse throughout the timecourse, indicating that the fluorescent tag was not the cause of the aggregation (Figure 2.1). These results suggest that addition of aromatic residues effectively increases the aggregation propensity of the Lsm4 PrLD. Although the $\Delta 35$ mutant never formed foci, whether or not the aggregation propensity of the Lsm4 PrLD was altered cannot be determined due to the fact that the wt Lsm4 PrLD never formed foci throughout this timecourse.

Aggregation-Promoting Mutations Enhance Aggregation of Full-Length Lsm4

Given that the aggregation propensity of the Lsm4 PrLD could be rationally modulated, I asked whether or not the same PrLD mutations could similarly alter the aggregation propensity of the full-length protein. The same PrLD mutations were introduced in the context of the full-length protein and, similar to the previous experiment, each Lsm4 mutant was fused to the N-terminus of GFP and overexpressed under the *GALI* promoter for different timepoints. After 2 hours of overexpression, both the +FY and +3F3Y mutant constructs formed robust foci in most cells (Figure 2.2). However, neither wt Lsm4 nor the $\Delta 35$ mutant formed foci (Figure 2.2), indicating that the aggregation-enhancing mutations in the PrLD were also capable of

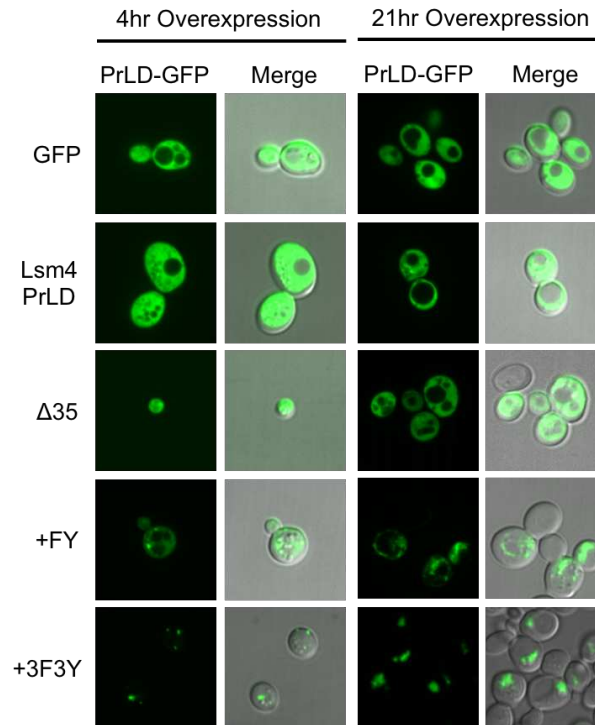


Figure 2.1: Aggregation-promoting mutations increase the rate of foci formation of the Lsm4 PrLD. Each Lsm4 PrLD variant was fused to the N-terminus of GFP and overexpressed for 4 and 21 hours prior to imaging.

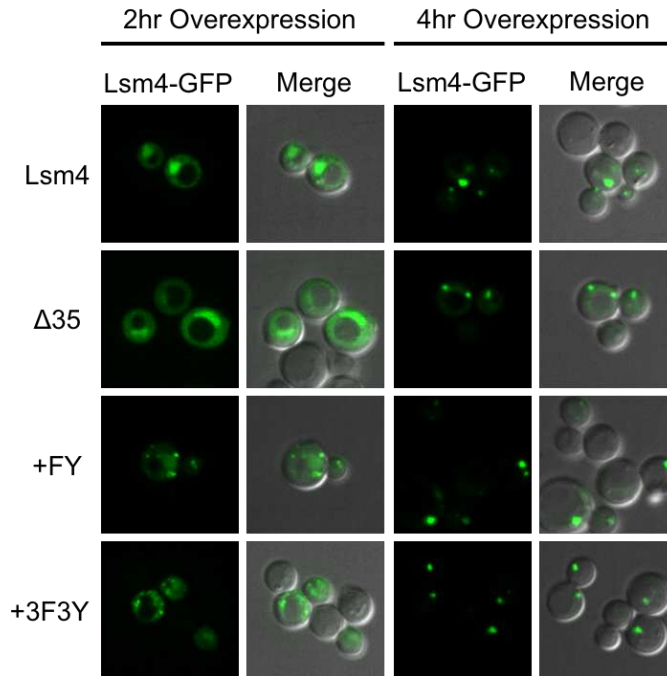


Figure 2.2: Aggregation-promoting mutations in the Lsm4 PrLD increase the rate of foci formation of full-length Lsm4. Lsm4 (with the indicated mutations in the PrLD) was fused to the N-terminus of GFP and overexpressed for 2 and 4 hours prior to imaging.

increasing the aggregation propensity of the full-length protein. Interestingly, after 4 hours of overexpression, all four constructs aggregated to form distinct foci (Figure 2.2). This result indicates that the full-length protein is much more inherently aggregation-prone than the PrLD. One possible explanation for this observation is that there are more contact sites in the full-length protein, thus enhancing the overall aggregation propensity.

Aggregation-Promoting Mutations in Lsm4 Do Not Disrupt P-body Recovery Post-Stress

Given that the aggregation-promoting mutations appeared to enhance aggregation of the Lsm4 PrLD both in isolation and in the context of the full-length protein, I asked whether these mutations could reduce p-body disassembly. A previous study demonstrated that the Lsm4 PrLD is necessary for p-body formation in a strain in which the scaffolding p-body protein Edc3 is knocked out (11). If the PrLD is necessary for p-body formation in this particular strain, I reasoned that the aggregation-promoting mutations within the PrLD might disrupt p-body dynamics, resulting in p-body persistence upon recovery from stress. To test this, each mutated full-length version of Lsm4 was cloned into a plasmid under the *LSM4* promoter. These constructs were each transformed into a yeast strain in which both Edc3 and Lsm4 are knocked out so that each mutant provided the only copy of Lsm4 in the cell. Additionally, the p-body protein Dcp2 was tagged with GFP, to mark p-bodies. Because the Lsm4 PrLD is necessary for p-body formation when Edc3 is missing, any effect these Lsm4 mutants might have on p-body assembly and disassembly could be monitored using the Dcp2-GFP p-body marker, so the Lsm4 constructs were untagged. Cells were depleted of glucose for 10 minutes to induce robust p-body formation, as shown in (7, 11). After stress, glucose was returned to cells for 20 minutes to allow them to recover before reassessing cells for p-bodies. After glucose deprivation, p-bodies were able to form in cells expressing wt Lsm4 as well as in cells expressing each mutant; however, p-bodies appeared to be decreased in the $\Delta 35$ mutant (Figure 2.3A). P-body formation did appear

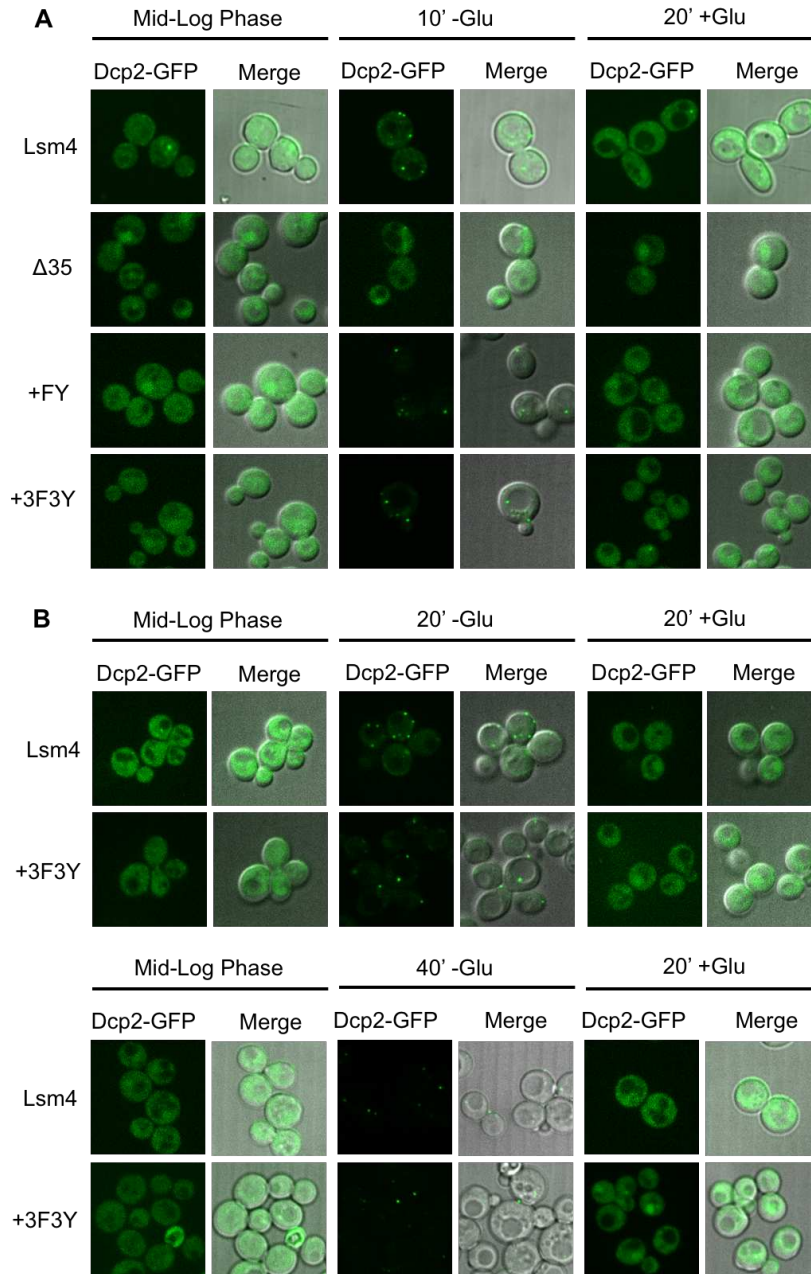


Figure 2.3: Mutations in the Lsm4 PrLD do not disrupt p-body disassembly post-stress.

Lsm4 variants containing the indicated mutations in the PrLD were untagged and expressed from the *LSM4* promoter on a plasmid in a *edc3Δlsm4Δ* strain, which also expresses Dcp2-GFP as a p-body marker. Cells were imaged during mid-log phase growth, after (A) 10 minutes of glucose deprivation and after (B) 20 or 40 minutes of glucose deprivation, and once more after 20 minutes of recovery in glucose replete media.

to be more robust in the cells in which the +FY and +3F3Y mutants were expressed; however, upon recovery from glucose deprivation, p-body levels were reduced, despite the increase in aggregation propensity of Lsm4 (Figure 2.3A).

This result was contrary to what I expected, so I tried increasing the length of exposure to stress to increase the time for the aggregation-promoting PrLDs to make more rigid contacts with other proteins in p-bodies. I only analyzed the most aggregation-promoting mutant, +3F3Y for this experiment reasoning that this one would be the most likely to promote irreversible association of proteins within p-bodies. After both 20 minutes and 40 minutes of glucose deprivation, p-bodies formed (Figure 2.3B); however, after 20 minutes of recovery in glucose replete media, p-body formation disappeared and Dcp2 was able to revert to a diffuse state (Figure 2.3B). Altogether, these results indicate that although the mutations introduced did effectively increase aggregation propensity, they were not sufficient to disrupt p-body recovery after stress.

Deleting Portions of the Lsm4 PrLD Does Not Prevent P-body Assembly

Because the aggregation-promoting mutations do not appear to disrupt p-body disassembly, I thought that perhaps I was not effectively targeting the appropriate region of the PrLD for mutation. Three different Lsm4 PrLD deletion mutants were constructed to attempt to more accurately determine which part of the PrLD was necessary for p-body assembly. Deletions of the first 50 ($\Delta 1^{st}$ 50), middle 50 (Δ Mid 50), and last 50 (Δ Last 50) amino acids of the Lsm4 PrLD were made as well as a deletion of the entire PrLD (Lsm4 Δ C) with sequences shown in Table 2.2. P-body formation and disassembly was assessed in the same *edc3 Δ lsm4 Δ* knockout strain as used previously. Surprisingly, after glucose deprivation all of the cells harboring the Lsm4 PrLD deletion mutants were still able to form p-bodies (Figure 2.4), indicating that none of the deletions could effectively prevent p-body assembly. Even cells expressing the Lsm4 Δ C

Table 2.2: Deletions made within the Lsm4 PrLD.

Mutation	Sequence
Lsm4 PrLD	QQINSNNNSNSNGPGHKRYNNRDSNNNRGNYNRRNNNNGNSNRRPYSQ NRQYNNSNSSNINNSINSINSNNQNMNNGLGGSVQHFFNSSSPQKVEF
Lsm4 Δ1st 50	RQYNNSNSSNINNSINSINSNNQNMNNGLGGSVQHFFNSSSPQKVEF
Lsm4 ΔMid 50	QQINSNNNSNSNGPGHKRYNNRDSNNGLGGSVQHFFNSSSPQKVEF
Lsm4 ΔLast 50	QQINSNNNSNSNGPGHKRYNNRDSNNNRGNYNRRNNNNGNSNRRPY

Sequences of the Lsm4 PrLD with the indicated amino acid deletions are shown.

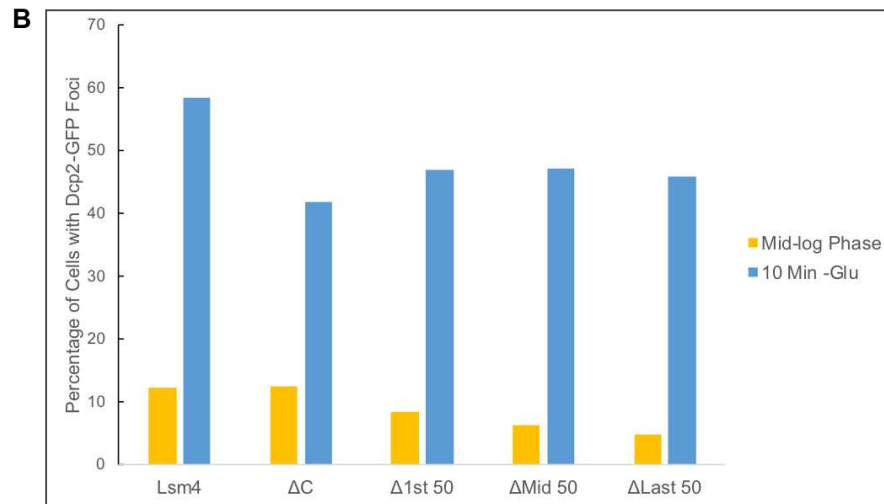
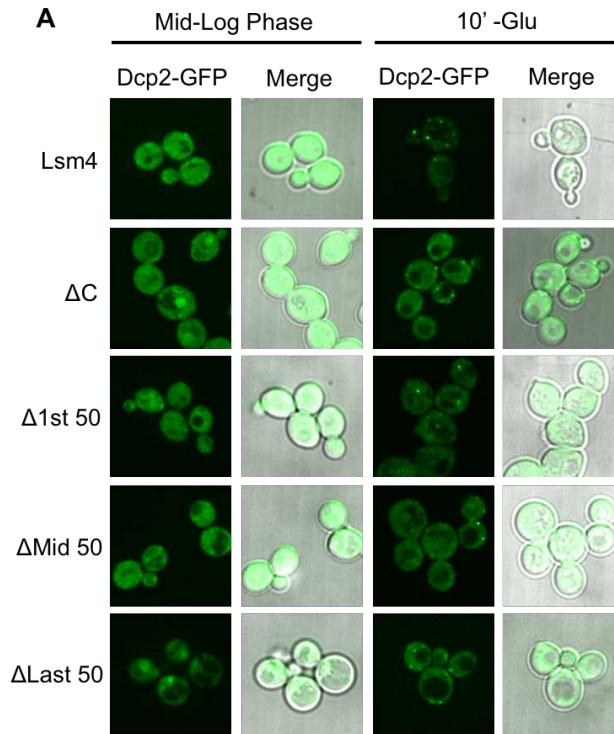


Figure 2.4: Deletions in the Lsm4 PrLD do not prevent p-body assembly upon stress. (A) Lsm4 variants containing the indicated mutations in the PrLD were untagged and expressed from the *LSM4* promoter on a plasmid in a *edc3 Δ lsm4 Δ* strain, which also expresses Dcp2-GFP as a p-body marker. Cells were imaged during mid-log phase growth and after 10 minutes of glucose deprivation. **(B)** Quantification of the percentage of cells containing at least one Dcp2-GFP focus for each Lsm4 variant during mid-log phase growth (yellow bars) and after 10 minutes of glucose deprivation (blue bars).

mutant were still able to form p-bodies, although not as robustly as when wt Lsm4 was expressed (Figure 2.4). Even though these mutations could partially disrupt p-body formation, none were able to completely prevent p-body formation.

In all of the previous experiments, Dcp2 was tagged with GFP as a p-body marker, but none of the Lsm4 mutants were tagged, making it impossible to determine whether or not the Lsm4 mutations were affecting aggregation of Lsm4 specifically. To analyze the localization of the Lsm4 mutants upon stress, all of the Lsm4 deletion mutants were tagged with mCherry and co-expressed in the same $\Delta edc3 \Delta lsm4$ Dcp2-GFP strain. Surprisingly, although p-body formation was not prevented, the Lsm4 $\Delta 1^{st} 50$ and Lsm4 ΔC deletion mutants were themselves precluded from localizing to p-bodies after glucose deprivation (Figure 2.5). However, Lsm4 itself does not appear to localize to p-bodies very robustly, only forming foci in about 11% of cells (Figure 2.5). Together, these results indicate that Lsm4 is not absolutely necessary for p-body formation in this strain as previously thought (11).

Lsm4 Joins P-Bodies Later in Assembly

Given that Lsm4 does not appear to localize to p-bodies after ten minutes of glucose deprivation, I hypothesized that it might localize to p-bodies at a later timepoint after stress induction, once p-bodies are already formed. To test this theory, p-body assembly was monitored over time, comparing both wt Lsm4 and the Lsm4 ΔC mutant that has more limited p-body formation. Cells expressing either wt Lsm4 or the Lsm4 ΔC mutant were imaged after 10', 15', 20', 25', and 30' of glucose deprivation. While Dcp2-GFP labeled p-bodies were able to form in the majority of cells as early as 10 minutes into glucose deprivation in the wt Lsm4 strain, p-bodies did not accumulate to wild-type levels until 20 minutes into glucose deprivation in the Lsm4 ΔC strain (Figure 2.6B). The 10-minute glucose deprivation timepoint may not provide enough time for p-bodies to efficiently form to wild-type levels when the Lsm4 PrLD is absent;

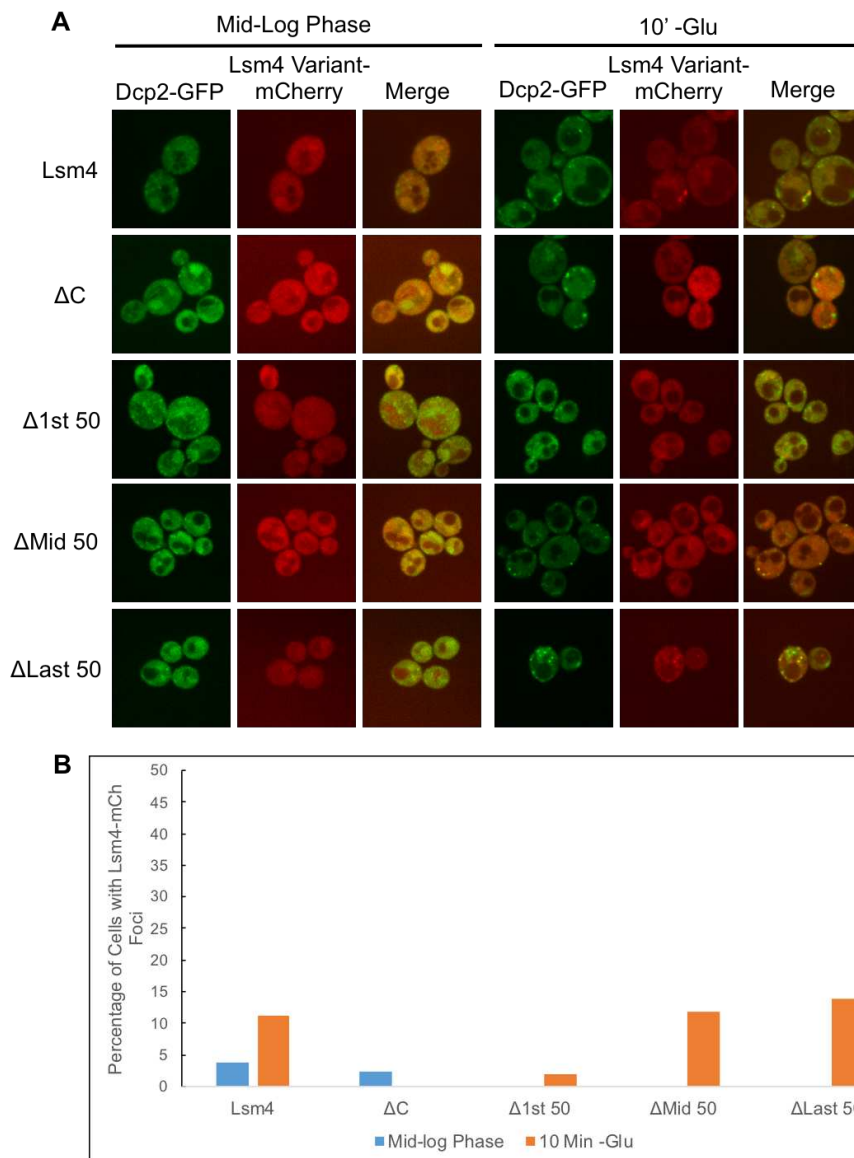


Figure 2.5: Lsm4 is not robustly recruited to p-bodies upon stress. (A) Lsm4 variants containing the indicated mutations in the PrLD were tagged with mCherry and expressed from the *LSM4* promoter on a plasmid in a *edc3 Δ lsm4 Δ* strain, which also expresses Dcp2-GFP as a p-body marker. Cells were imaged during mid-log phase growth and after 10 minutes of glucose deprivation. **(B)** Quantification of the percentage of cells containing at least one Lsm4-mCherry focus for each Lsm4 variant during mid-log phase growth (blue bars) and after 10 minutes of glucose deprivation (orange bars).

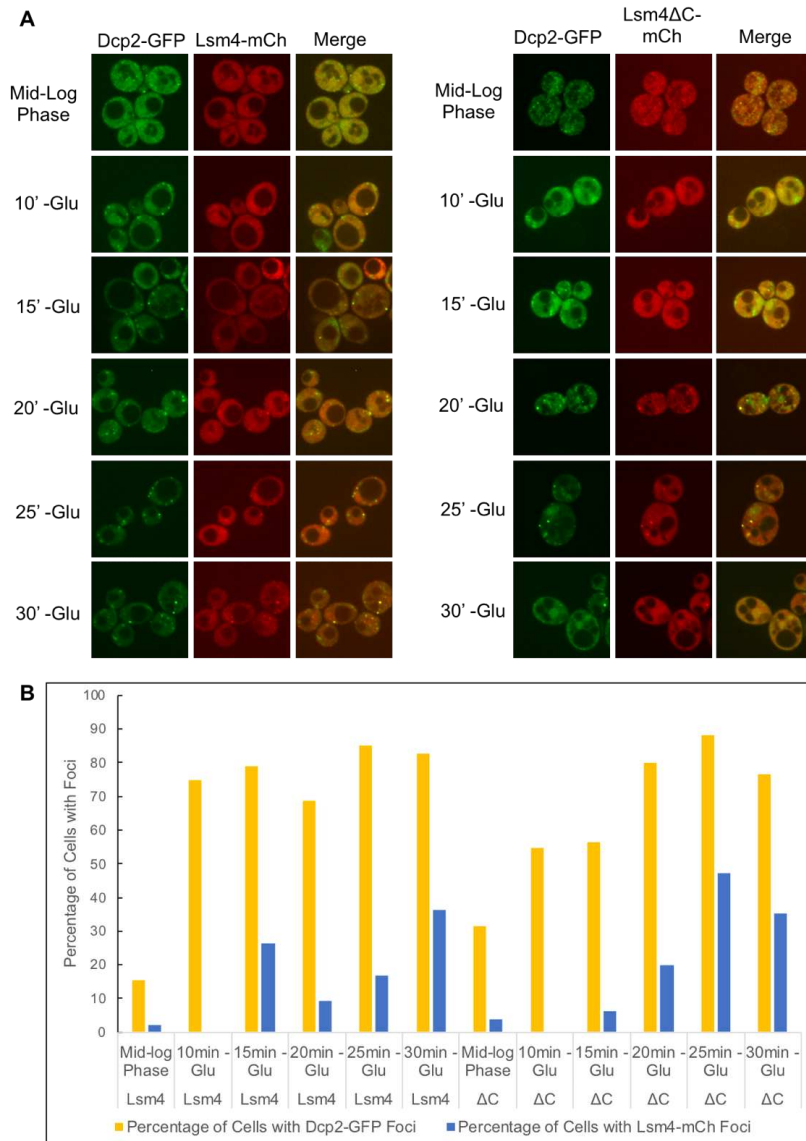


Figure 2.6: Lsm4 is mildly recruited to p-bodies after extended stress. (A) Lsm4 (left panel) and Lsm4 Δ C (right panel) were tagged with mCherry and expressed from the *LSM4* promoter on a plasmid in a *edc3 Δ lsm4 Δ* strain, which also expresses Dcp2-GFP as a p-body marker. Cells were imaged during mid-log phase growth and after 10, 15, 20, 25, and 30 minutes of glucose deprivation. **(B)** Quantification of the percentage of cells containing at least one Dcp2-GFP focus (yellow bars) or at least one (blue bars) for Lsm4 and Lsm4 Δ C during mid-log phase growth and each of the indicated timepoints after the onset of glucose deprivation.

however, p-bodies are still capable of eventually increasing to normal levels once the stress has been applied for a longer period of time. When analyzing p-body formation of mCherry-tagged Lsm4, neither wt Lsm4 nor Lsm4 Δ C accumulate to very high levels after 30 minutes of glucose deprivation (Figure 2.6B). These data suggest that Lsm4 joins p-bodies in later stages of assembly, implying that its PrLD is not necessary for p-body assembly, unless only a very small amount of Lsm4 is required, in which case localization to p-bodies may not be able to be resolved by microscopy.

Discussion

While we understand the amino acid determinants that increase aggregation and amyloid formation in the context of prions, these determinants do not appear to be extended to PrLDs that are involved in the reversible aggregation process of p-body formation. Although I successfully increased the inherent aggregation propensity of the Lsm4 PrLD (Figure 2.1), these mutations were not sufficient to detectably disrupt p-body dynamics when the PrLD is an integral component determining p-body assembly (Figure 2.3A). These data indicate that either the mutations introduced do not make Lsm4 as aggregation-prone as thought, or that the cell simply has robust systems working to remove aberrantly aggregated p-bodies.

In the context of the Lsm4 PrLD alone, inserting aggregation-promoting aromatic residues increased the rate at which the overexpressed PrLD could form foci in cells (Figure 2.1); however, introducing these same mutations to the PrLD in the context of full-length Lsm4 did not result in persistently aggregated p-bodies as predicted (Figure 2.3A). A few explanations could account for this discrepancy between the results obtained when investigating the PrLD overexpression system and the endogenous p-body system. One reason is that the PrLD is only a piece of the entire protein and the full-length protein could behave differently than the PrLD upon introduction of these aggregation-promoting mutations. However, the full-length protein

did appear to aggregate more rapidly when the +FY mutations were introduced compared to the wild-type protein (Figure 2.2), indicating that these mutations in the PrLD were capable of modulating the aggregation propensity of the full-length protein. Another caveat to consider is that the PrLDs were overexpressed in this system, thus artificially inflating the concentration of the aggregation-prone species compared to normal levels in the cell. One possible consequence of this overexpression system is that at these higher concentrations, the protein quality control machinery (PQC), consisting of chaperones, the autophagy pathway, and the ubiquitin-proteasome system, is overwhelmed and cannot destroy these aggregates. Perhaps when these aggregation-prone species are produced at normal expression levels, the PQC machinery is capable of disassembling hyper-aggregated p-bodies. It is worth noting that the p-body formation and disassembly assay is low-resolution, thus it remains possible that the mutations did effect p-body dynamics, but not to a detectable level. One additional method to determine whether or not the mutations introduced disrupted p-body dynamics would be to perform FRAP (Fluorescence Recovery After Photobleaching) experiments to determine the dynamic state of these p-bodies. This technique simply requires photobleaching a spot within a p-body and then analyzing recovery of fluorescence. If these aggregation-promoting mutations did make these p-bodies more aggregate-like, I would expect them to behave less dynamically, indicated by slower recovery times after photobleaching. However, one caveat is that p-bodies are very tiny, so FRAP may not be a viable method to examine these entities.

Analysis of the Lsm4 PrLD deletion mutants indicated that the Lsm4 PrLD may not be as essential to p-body formation as previously thought. Indeed, deleting the Lsm4 PrLD does result in formation of fewer p-bodies upon stress (Figure 2.4); however, deletion of the Lsm4 PrLD does not result in complete loss of p-body formation, as was previously reported (11). Given that this PrLD does not appear to be absolutely necessary for p-body formation, the interactions it is

making in the p-body might not be as important for p-body structural integrity as originally thought, possibly explaining why aggregation-promoting mutations in this region were not sufficient to disrupt p-body dynamics (Figure 2.3). Although deleting the Lsm4 PrLD and the first 50 amino acids of the Lsm4 PrLD prevented Lsm4 from forming foci, wt Lsm4, as well as the other PrLD deletion mutants, only localized to p-bodies in a fraction of the cells (Figure 2.5), suggesting that Lsm4 itself may not even be a necessary p-body component. Overall, deleting the Lsm4 PrLD does partially inhibit p-body formation, but not to the extent required to effectively perform a mutational study such as the one described here.

Finally, when fused to a fluorescent protein tag, Lsm4 appears to join p-bodies in about 35% of cells after 30 minutes of glucose deprivation (Figure 2.6), further suggesting that it is not an essential structural component of p-bodies. The low level of Lsm4 p-body localization could be caused by interference from the mCherry tag, or insufficient glucose deprivation treatment; however, the latter explanation is unlikely given that Dcp2 showed a substantial increase in p-body formation upon glucose deprivation (Figure 2.6). If Lsm4 is not an integral structural p-body component, it may not form contacts that are sufficient to cause p-body persistence when these interactions with other components (or with itself) are strengthened by aggregation-promoting mutations.

Taking into account other observations about p-bodies, they may not be an ideal system to study the persistent aggregation of RNP granules. P-bodies are present both normally and under stress (7), possibly indicating that p-body disassembly is well regulated enough to avoid persistent aggregation of its components. Additionally, p-bodies have not been linked to disease or any sort of persistent aggregation, even though they appear to behave similarly to stress granules, which are heavily tied to aggregation and disease. Thus, stress granules may provide a better model with which to investigate this problem because they are only present under stress

and have also been linked to persistent aggregation events (4). Additionally, recent observations suggest that different interactions from different proteins can substitute for one another in p-body formation (18), possibly indicating that other interactions are compensating in place of Lsm4's interactions when it's PrLD is mutated. This same study also showed that when Edc3, Dhh1, and the Lsm4 PrLD are all deleted, p-bodies can only form in about 15% of cells (18). This might be a more appropriate strain to conduct these mutational studies in, since it provides a more complete knockdown of p-bodies than the $\Delta edc3\Delta lsm4C$ strain used here, in which p-bodies still form in about 40% of cells (Figure 2.4).

Altogether, the results from this study suggest that, despite increasing the aggregation propensity of the Lsm4 PrLD, p-body dynamics remain unperturbed. Even considering the fact that the Lsm4 PrLD may not be as important for p-body formation as previously thought, it is still remarkable that presence of a highly aggregation-prone region appears to have no effect on p-body disassembly. Although it remains possible that the Lsm4 PrLD, and p-bodies in particular, are not well-suited for studying persistent aggregation of RNP granules, other systems, such as the PQC machinery, could be also be contributing to aggregate dissolution, making the results difficult to interpret. Future work will be aimed at understanding why RNP granules can persistently aggregate in some situations, but not others.

REFERENCES

1. Li YR, King OD, Shorter J, & Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201(3):361-372.
2. Ramaswami M, Taylor JP, & Parker R (2013) Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154(4):727-736.
3. Buchan JR (2014) mRNP granules. *RNA Biology*:e29034.
4. Protter DS & Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26(9):668-679.
5. Parker R & Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Mol Cell* 25(5):635-646.
6. Luo Y, Na Z, & Slavoff SA (2018) P-Bodies: Composition, Properties, and Functions. *Biochemistry*.
7. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, & Parker R (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 11(4):371-382.
8. March ZM, King OD, & Shorter J (2016) Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res.*
9. Reijns MA, Alexander RD, Spiller MP, & Beggs JD (2008) A role for Q/N-rich aggregation-prone regions in P-body localization. *J Cell Sci* 121(Pt 15):2463-2472.
10. Gilks N, *et al.* (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15(12):5383-5398.
11. Decker CJ, Teixeira D, & Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179(3):437-449.
12. King OD, Gitler AD, & Shorter J (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* 1462:61-80.
13. Kim HJ, *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495(7442):467-473.
14. Mackenzie IR, *et al.* (2017) TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* 95(4):808-816 e809.
15. Toombs JA, McCarty BR, & Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-332.
16. Sherman F (1991) Getting started with yeast. *Methods in enzymology* 194:3-21.
17. Gietz RD & Sugino A (1988) New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* 74(2):527-534.
18. Rao BS & Parker R (2017) Numerous interactions act redundantly to assemble a tunable size of P bodies in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*.

CHAPTER THREE: UNDERSTANDING THE EFFECTS OF PROLINE DELETIONS ON THE AGGREGATION PROPENSITY OF PRION-LIKE DOMAINS FROM STRESS GRANULE PROTEINS

Introduction

Upon stress, cells have the ability to sequester different RNA and protein components into membraneless compartments called stress granules. Stress granules are primarily composed of stalled translation initiation complexes and are thought to function to free up the translation machinery for synthesis of only the factors essential for survival (1, 2). Although this process appears to be quite important for cell survival upon stress, the mechanism of assembly is poorly understood. Interestingly, many stress granule proteins contain aggregation-prone prion-like domains (PrLDs). PrLDs are glutamine/asparagine (Q/N) rich protein domains that compositionally resemble yeast prion forming domains (3). In some cases, these domains are required for stress granule formation (4), indicating that they may aid in the assembly of these granules. Because stress granule formation is a reversible aggregation process, it is possible that PrLDs may normally promote loose aggregation of stress granule components, but not to the extent that stable aggregates are formed, which could prevent stress granule disassembly.

Interestingly, mutations in many different stress granule proteins have been linked to neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) (5, 6). A common pathology of these diseases is the formation of cytoplasmic inclusions containing one or more of these mutated proteins in certain neurons and/or muscle tissues of patients with these diseases (5, 7). Inclusions found in disease-associated tissue stain with known stress granule markers, leading to the idea that these inclusions are actually persistent stress granules unable to be cleared by the protein quality control (PQC) machinery (8, 9). These mutations are often located in the PrLDs of these stress

granule proteins and are sometimes associated with an increase in the aggregation propensity of the PrLD (10, 11). These observations suggest that mutations in stress granule PrLDs might sufficiently enhance the aggregation of these domains to promote irreversible aggregation of stress granules, resulting in persistent aggregates in the cytoplasm of the cell. Understanding the mechanisms of stress granule assembly and disassembly, and the role of PrLDs in this process, will be essential to gain insight into these disease pathologies.

Research into the mechanisms of stress granule assembly has revealed the ability of some of these proteins, including the disease-associated proteins FUS, hnRNPA1, hnRNPA2B1, TDP43, and TIA1, to undergo liquid-liquid phase separation (LLPS) *in vitro* (12-17). LLPS is a process by which changes in solution conditions cause the solute (protein in this case) to drop out of solution in a concentration-dependent manner to form dynamic droplets that behave as liquids (18). LLPS of these proteins can be promoted by the addition of RNA and crowding agents, which better mimic the cellular environment (12, 13, 19). Some experiments performed on stress granules *in vivo* also suggest that stress granules normally behave as dynamic, liquid-like entities (12, 13). Together, these observations suggest that stress granule assembly is triggered by demixing of proteins and RNAs in the cytoplasm to create a membraneless liquid-like compartment, which can easily and rapidly disassemble after the stress has dissipated. Interestingly, when disease-associated mutations are introduced into the PrLDs of these stress granule proteins, LLPS droplets lose their liquid-like qualities and convert to more stable aggregates (12, 13, 15, 16, 20), further bolstering the argument that disease-associated inclusions are aberrantly aggregated stress granules.

Altogether, stress granule formation appears to play an important role in different neurodegenerative disease pathologies, with PrLDs appearing to contribute as well. However, further research into characteristics that promote aberrant aggregation of these proteins is

necessary to gain a more complete understanding of the disease process. Towards this goal I developed a project designed to assess the effects of aggregation-promoting mutations in the PrLDs of different stress granule proteins on stress granule clearance. Using the rationale that PrLDs might undergo prion-like aggregation within stress granules to help them assemble, I tested whether addition of aggregation-promoting mutations in these domains might enhance stress granule assembly and prevent disassembly, resulting in persistent aggregates, as is thought to occur in disease. I noticed that the core stress granule proteins, Pab1 and Pbp1, contain high numbers of proline residues within their PrLDs. Because proline residues are disfavored for amyloid formation (21), I reasoned that removing these residues might promote conversion of these PrLDs to stable amyloids, which would in turn disrupt stress granule dynamics. Deletion of the proline residues from the Pab1 and Pbp1 PrLDs successfully increased the aggregation propensity of these domains in isolation; however, upon introduction of these mutations into the full-length proteins, stress granule disassembly was unaffected. Additionally, deletion of chaperone proteins involved in stress granule disassembly was found to slow recovery of the Pab1 and Pbp1 foci when the proline residues are deleted, but does not completely inhibit disassembly, suggesting that other components of the PQC machinery may act redundantly to prevent accumulation of aberrant aggregates.

Materials and Methods

Strains and Growth Conditions

All experiments were performed in *Saccharomyces cerevisiae* using standard yeast handling and growth conditions (22). Yeast were grown at 30°C unless otherwise specified. Plasmids were transformed into appropriate yeast strains using standard yeast transformation protocols (22). Plasmids carrying overexpressed PrLDs and full-length proteins for microscopy, ThioflavinT staining, and SDD-AGE were transformed into BY4741 (*MATa his3D1 leu2D0*

met15D0 ura3D0) (23). Plasmids carrying the Pbp1 PrLD-mCh and Pbp1 PrLD Δ Pro-mCherry constructs for overexpression in addition to heat shock were transformed into YER1220 (*MATa his3D1 leu2D0 met15D0 ura3D0 Tif4631::Tif4631-GFP-HIS3*), a strain from the yeast GFP collection (24). Plasmids carrying endogenously expressed Pab1 and Pbp1, as well as the versions with deletions of proline residues, were transformed into YER2136 (*MATa his3D1 leu2D0 met15D0 ura3D0 Pub1::Pub1-GFP-HIS3*), YER1503 (*MATa his3D1 leu2D0 met15D0 ura3D0 ydj1::KanMx*) (25), YER1643 (*MATa kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 sup35::KanMx ssa1::TRP1 pJ533(URA3)*), YER2151 (*MATa his3D1 leu2D0 met15D0 ura3D0 ydj1::KanMx Pub1::Pub1-GFP-HIS3*), and YER2168 (*MATa his3D1 leu2D0 met15D0 ura3D0 ssa1::KanMx Pub1::Pub1-GFP-HIS3*) for acute and chronic heat shock assays.

Cloning Methods

The overexpressed GFP fusions were cloned into pER2052, which is a derivative of Yeplac181 (26) containing the *GALI* promoter followed by GFP. Overexpressed PrLD-mCherry-HA fusions used in the ThioflavinT assay were cloned into pER2053, which is a derivative of Yeplac181 containing the *GALI* promoter followed by mCherry-HA.

Overexpressed Pbp1 PrLD and Pbp1 PrLD Δ Pro mCherry fusions used for the overexpression + heat shock assay were cloned into pER1131, a derivative of Yeplac181 containing the *GALI* promoter followed by mCherry. Overexpressed PrLD-HA fusions used for the SDD-AGE assay were cloned into pER1121, which is a derivative of Yeplac181 containing the *GALI* promoter. PrLDs and full-length versions were first amplified from yeast genomic DNA and then re-amplified to synthetically build the mutated regions using overlapping primers deleting the proline residues. A start codon was inserted before each PrLD to ensure expression. PrLDs used in the SDD-AGE assay also had a HA tag synthetically built onto the C-terminus of each PrLD. Each construct was cloned in between the BamHI and XhoI restriction sites, following the *GALI*

promoter, but preceding any fluorescent protein tag (if applicable) in each vector using In-Fusion cloning.

Versions of Pab1 and Pbp1 expressed from their native promoters, both wild-type and with proline deletions, were cloned into pER2040, a derivative of YcPlac111 (26) containing mCherry-HAx3. Each protein was first amplified with its endogenous promoter from yeast genomic DNA and then re-amplified to synthetically build the mutated regions using overlapping primers designed to delete the proline residues in the PrLDs. Each construct was cloned preceding the mCherry-HAx3 cassette, in between the BamHI and XhoI restriction sites, using In-Fusion cloning.

Overexpression and Stress Conditions

For experiments investigating overexpression of all PrLDs and full-length Pab1 and Pbp1 constructs, cells were grown in SGal/Raff -Leu media (to select for the plasmids) for 2, 4, 6, or 24 hours to induce the *GALI* promoter to overexpress each construct.

For overexpression in addition to heat shock, cells were diluted into SRaff -Leu media and grown to mid-log phase before PrLD induction. 2% galactose was added to induce PrLD expression and cells were incubated to 2, 4, or 12 hours. After overexpression, cells were imaged and then exposed to heat shock at 46°C for 30 minutes before imaging again.

To perform stress granule assembly and recovery experiments, cells were grown to $OD_{600} \approx 0.3-0.6$ in SC-Leu prior to stress induction. To induce heat shock, cells were transferred to a 46°C shaking water bath for 30 minutes, and then returned to a 30°C shaker for 2 hours for recovery. 1mL of cells were concentrated to 50 μ L during mid-log phase, after heat shock, and after each recovery timepoint for imaging.

Confocal Microscopy

Cells were imaged using an Olympus IX83 confocal spinning disk microscope. Images were collected as single planes.

Thioflavin T Staining

Cells were grown in SGal/Raff -Leu media for 24 hours to induce overexpression of each PrLD. Cells were then harvested by centrifugation for 5 minutes at 1,500 x1000 rpm. Pellets were resuspended in 3mL of 30 μ M thioflavin T in 10mM Tris/EDTA in 1X TE, pH 7 buffer and shaken at 30°C for 30 minutes. Cells were washed 3 times in 10mM Tris/EDTA in 1X TE, pH 7 buffer prior to imaging. Samples were imaged using a 405ex/525em filter set.

SDD-AGE Methods

Cells were grown in 15mL of SGal/Raff -Leu media for 24 hours to induce the *GALI* promoter to overexpress each PrLD-HA fusion. Cells were harvested by centrifugation for 5 minutes at 3 x 1,000 rpm, 4°C and then washed 1 time in water. Pellets were resuspended in 1mL of spheroplasting solution (1.2M D-Sorbitol, 0.5mM MgCl₂, 50mM β -mercaptoethanol, 0.5mg/mL zymolyase in 20mM Tris-HCl, pH 7.4) and incubated at 30°C for 1 hour. Spheroplasts were harvested by gentle centrifugation for 5 minutes at 800xg and then resuspended in 200 μ L of lysis buffer (10mM β -mercaptoethanol, 10mM AEBSF in 20mM Tris-HCl, pH 7.4). Spheroplasts were lysed by vortexing for 1.5 minutes followed by incubation on ice for 10 minutes. Cell debris was pelleted by centrifugation for 2 minutes at 4 x 1,000 g at 4°C. Protein concentrations were measured using a Bradford Assay. For each sample, 60 μ g of protein was incubated in 2% SDS for 7 minutes before loaded onto a 1.5% agaose + 0.1% SDS gel. The gel was run for ~3 hours at 60V. Protein was transferred onto a PVDF membrane and detected using an anti-HA antibody for primary detection and a secondary antibody (AlexaFluor IR800) as described previously (27).

Chronic Heat Shock Assay

Cells were grown to $OD_{600} \approx 0.3-0.6$ in SC-Leu media and then plated as 10-fold serial dilutions onto SC-Leu plates. Two identical plates were prepared, one incubated at 30°C as a control and the other incubated at 40°C. Cells were allowed to grow for 4 days prior to imaging.

Results

Deleting Proline Residues from Stress Granule PrLDs is Predicted to Increase PrLD

Aggregation

Many different RNP granule proteins contain PrLDs that are rich in proline residues, yet prolines are considered poor for prion formation, due to their low beta sheet propensities (21). Two central yeast stress granule proteins, Pab1 and Pbp1, both contain proline-rich PrLDs (Table 3.1). Neither PrLD is predicted to be prone to prion-like aggregation using the Prion Aggregation Prediction Algorithm (PAPA) (28); however, upon deletion of all proline residues within the PrLD, these domains score as aggregation-prone (Table 3.1). I reasoned that these proline residues might be favored for stress granule assembly by promoting LLPS and can also effectively prevent unwanted persistent aggregation of stress granules by inhibiting amyloid formation of domains that are otherwise predicted to be quite aggregation-prone. This work focuses on the proteins Pab1 and Pbp1 because they are both core stress granule constituents that localize to stress granules in response to many different types of stresses (29-31). These proteins have also both been isolated in stress granule cores (32), suggesting that they might already form more stable contacts within granules. I reasoned that by increasing the aggregation of the PrLDs of these two proteins, stress granules might persist, similar to the stress granule aggregates observed in disease tissue (5, 10).

Table 3.1: Deletions of proline residues in the Pab1 and Pbp1 PrLDs.

Mutation	PAPA Score	Sequence
Wild-type Pab1 PrLD	-0.05	YQQATAAAAAAAAAAGM P GQFM P PMFYGV M P PRGV P FN G P NP Q QMN P M GGM P K N GM P P Q FR N G P VYGV P P Q GG F P RNANDNNQFYQQKQRQALG EQLYKKVSAKTSNEEAAGKITGMILDL P P Q EV F P LLESD E LF E QHYKEAS AAYESFKKEQEQQTEQA
Pab1 PrLD Δ Pro	0.15	YQQATAAAAAAAAAAGMGQFM M FYGV M RGVFN G N Q QMN M GGM K NGM QFR N GVYGV Q GGFR N ANDNNQFYQQKQRQALGEQLYKKVSAKTSNEE AAGKITGMILDLQEV F LLESD E LF E QHYKEASAAAYESFKKEQEQQTEQA
Wild-type Pbp1 PrLD	-0.02	QTRFQ Q RQLNSM G NAV P GM N P AMGM N MGG M MG F P MGG P SAS P NP M M N GFAAGSMGM Y M P F Q P MFY H P SM P Q M M P VMGS N GAE E GG G NI S P H V P AGF M AA G P G A P MGAF G Y P GG I P FQ G MM G SG P SG M P ANGSAM H SHGHSRNYHQTSHHGHHNSSTSGHK
Pbp1 PrLD Δ Pro	0.06	QTRFQ Q RQLNSM G NAVGM N AMGM N MGG M MG F MGGSAS N MM N GFA AGSMGM Y M F Q Q MFY H SM Q MM V MG S N GAE E GG G NI S HVAGF M AA G AMGAF G Y GG I F Q G MM G SG S GMANGSAM H SHGHSRNYHQTSHHGHHN SSTSGHK

Each sequence of the Pab1 and Pbp1 PrLDs is shown with proline residues bolded in red. Sequences of the PrLDs with deletions of the proline residues are also shown as well as PAPA scores for wild-type and mutant versions of each full-length protein. Scores above 0.05 are predicted to be prion-like, and scores below 0.05 are not predicted to be prion-like.

Deleting Proline Residues from Stress Granule PrLDs Enhances Foci Formation

To determine whether deleting proline residues from the Pab1 and Pbp1 PrLDs would be sufficient to enhance aggregation, all of the proline residues were first deleted from the Pab1 and Pbp1 PrLDs. These mutated PrLDs, as well as the wild-type versions of both PrLDs, were fused to GFP and placed under control of the *GALI* promoter to overexpress each construct. Both wild-type and mutant PrLDs were overexpressed for 2, 4, 6, and 24 hours before assessing assembly into foci using fluorescence microscopy. The Pab1 PrLD showed minimal foci formation throughout the timecourse; however, deletion of the proline residues did effectively increase assembly into foci, with foci prevalent after only 4 hours of overexpression (Figure 3.1A). In contrast, both the Pbp1 PrLD and Pbp1 PrLD Δ Pro constructs formed robust foci after only 4 hours of overexpression (Figure 3.1A), indicating that the proline deletions had no detectable effect on aggregation of the Pbp1 PrLD.

Given that foci formation of the Pab1 PrLD was increased by deletion of proline residues, I investigated whether these deletions could similarly affect assembly in the context of the full-length protein. Pbp1 was also tested to see if differences in assembly propensity could be resolved when PrLD mutations were made in the context of the full-length protein. Full-length Pab1 and Pbp1, with and without deletions of all proline residues in each PrLD, were fused to GFP and overexpressed from the *GALI* promoter for 2, 4, 6, and 24 hours before imaging. Surprisingly, the proline deletions had no effect on the foci formation of Pab1 and Pbp1 (Figure 3.1B). This result is not surprising for Pbp1 because the proline deletions did not detectably affect the assembly propensity of the PrLD (Figure 3.1A); however, this result was unexpected for the Pab1 constructs, which appeared to have increased aggregation in response to overexpression of the Pab1 PrLD Δ Pro construct alone. Although the full-length Pab1 Δ Pro

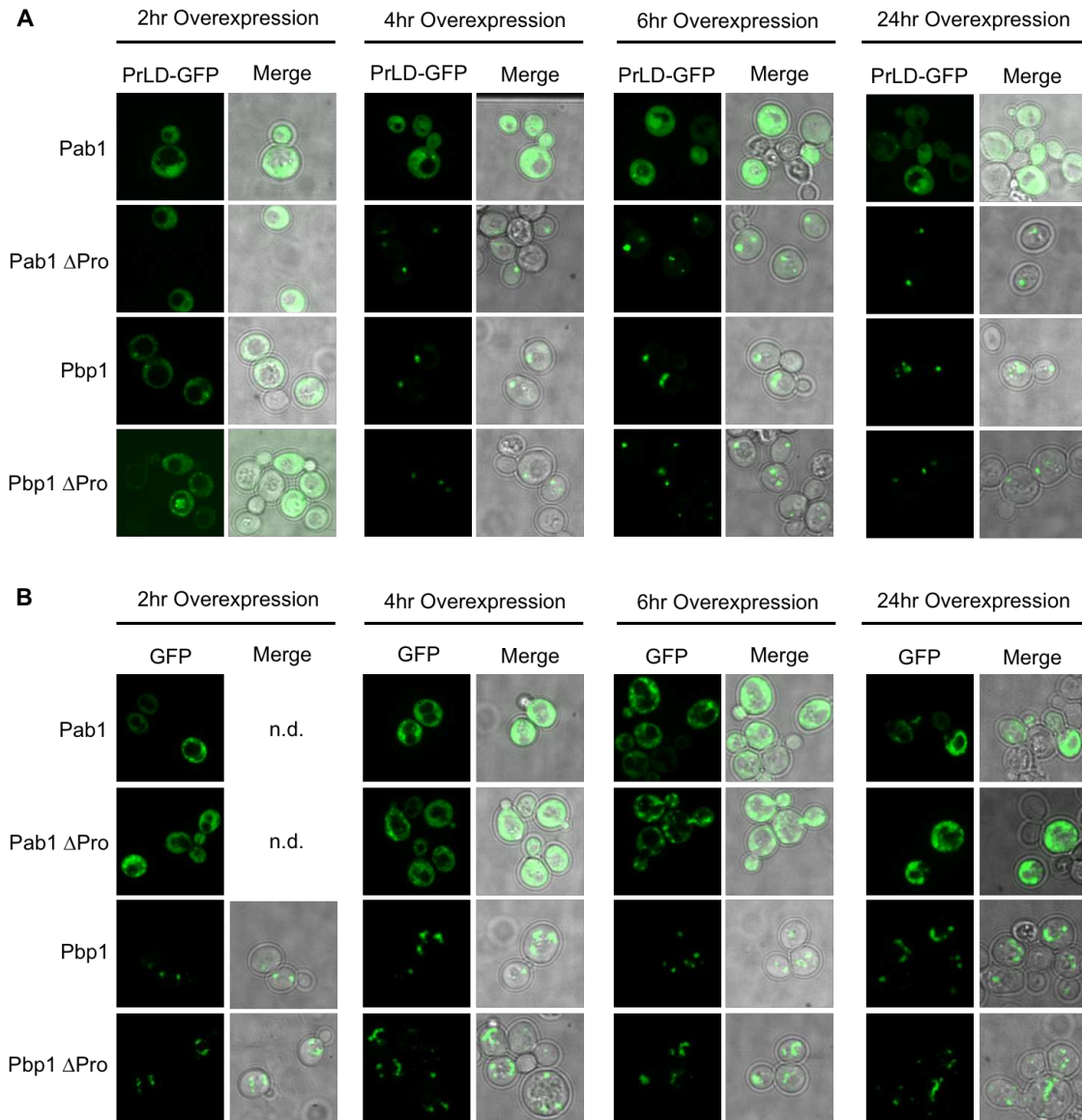


Figure 3.1: Effects of proline deletions on foci formation. PrLDs (A) or full-length proteins with the listed PrLD mutations (B) were fused to the N-terminus of GFP and overexpressed from the *GALI* promoter for 2, 4, 6, and 24 hours before imaging.

construct did form foci (Figure 3.1B), this foci formation was less extensive than was observed upon overexpression of the Pab1 PrLD Δ Pro in isolation (Figure 3.1A), indicating that the full-length protein might help to solubilize the aggregation-prone PrLD. In contrast, full-length Pbp1 appears to be very aggregation-prone, even without mutation.

Deleting Proline Residues from Stress Granule PrLDs Enhances Aggregation

Foci formation in cells indicates that these PrLDs are assembling, but does not yield any information about the types of assemblies being formed. These assemblies could be amorphous and unstructured, or more stable and rigidly structured. Often, more stable and structured aggregates are SDS-resistant, which is a property that can be determined by performing semi-denaturing agarose gel electrophoresis (SDD-AGE) (33). In this technique, cells are lysed and exposed to low concentrations of SDS, which will not denature stably aggregated species, such as amyloids, but will denature amorphous aggregates (33). These lysates are then run on an agarose gel and then transferred to a membrane for protein detection (33). SDS-resistant aggregates will run as a high molecular weight smear, whereas species that are SDS-soluble will appear as a low molecular weight band (33). Pab1 and Pbp1 PrLDs, with and without deletions of proline residues, were overexpressed for 24 hours and subjected to SDD-AGE analysis. Sup35 NM, a yeast prion domain that forms high molecular weight oligomers (27), was also run as a control. Interestingly, both the Pab1 Δ Pro and Pbp1 Δ Pro aggregates are SDS-resistant, observable as smears of higher molecular weight species on the blot, whereas the wild-type PrLD aggregates do not run as smears and thus appear to be soluble in SDS (Figure 3.2A). This data further suggests that deleting the proline residues increases aggregation propensity of these PrLDs and also promotes the formation of more stable aggregates.

Because proline residues can inhibit amyloid formation, I originally reasoned that deleting them might promote amyloid formation of the PrLDs. The stable aggregates formed by the mutated

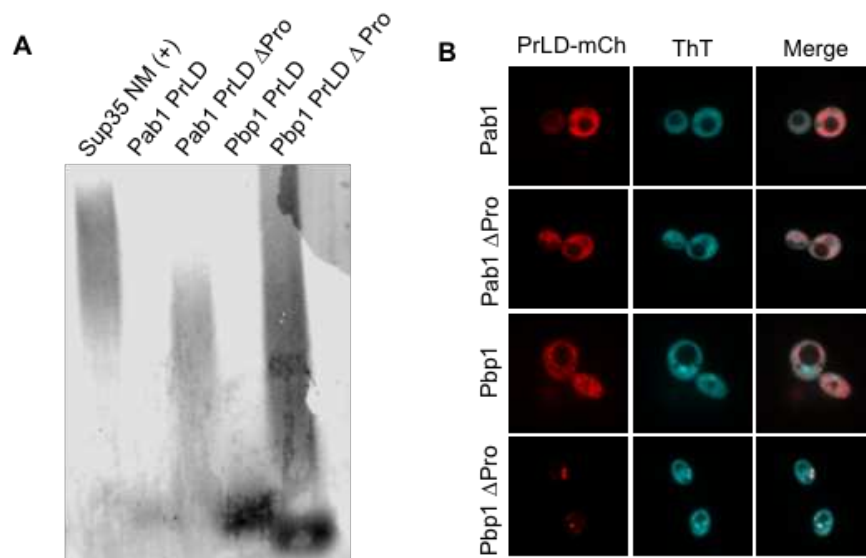


Figure 3.2: Deleting proline residues increases PrLD aggregation. (A) SDD-AGE analysis to assess oligomer formation of PrLDs. PrLDs were fused to HA and overexpressed for 24 hours before lysis. Oligomers appear as high molecular weight smears and the monomers appear as lower molecular weight bands towards the bottom of the blot. (B) The indicated PrLDs were fused to mCherry-HA and overexpressed for 24 hours before staining with thioflavin T to detect amyloid formation.

PrLDs and observed by SDD-AGE could be amyloid fibers, because amyloidogenic species are SDS-insoluble (33). To test this hypothesis and determine whether the Pab1 PrLD Δ Pro or Pbp1 PrLD Δ Pro were forming amyloid upon overexpression, each PrLD construct was fused to mCherry and overexpressed from the *GALI* promoter for 24 hours. After the period of overexpression, thioflavin T (ThT), a dye used to detect amyloid (34), was added to each culture to stain for amyloid. Strikingly, the Pbp1 PrLD Δ Pro forms foci that stain with ThT, indicating that these aggregates are forming amyloid fibrils (Figure 3.2B). However, the wild-type Pab1 and Pbp1 PrLDs, as well as the Pab1 PrLD Δ Pro, all remained diffuse (Figure 3.2B), in contrast to previous results showing foci formation when these PrLDs were fused to GFP (Figure 3.1A). These results suggest that the mCherry tag may be solubilizing these PrLDs, or that the GFP tag may promote aggregation.

Aggregation-Promoting Mutations Do Not Prevent Localization to Stress Granules

A previous study demonstrated that overexpression of the Nrp1 stress granule PrLD can nucleate amyloidogenic stress granules (35), and I reasoned that I might similarly be promoting formation of amyloidogenic stress granules with overexpression of the highly aggregation-prone mutated PrLDs. Because the Pbp1 PrLD, both with and without proline residues, is more aggregation-prone than the Pab1 PrLD (Figures 3.1 & 3.2), I chose this pair of PrLDs to investigate nucleation of stress granules. Both the Pbp1 PrLD and the Pbp1 PrLD Δ Pro were fused to mCherry and expressed in a strain also co-expressing Tif4631-GFP. The Pbp1 PrLD constructs were overexpressed using the inducible *GALI* promoter for 2, 4, and 12 hours prior to imaging. Although the Pbp1 PrLD Δ Pro formed foci during every overexpression timepoint, it never successfully nucleated stress granule formation, as determined by localization of Tif4631-GFP (Figure 3.3A). The wild-type PrLD formed some foci after 4 and 12 hours of overexpression, but similarly was unable to nucleate stress granule formation (Figure 3.3B).

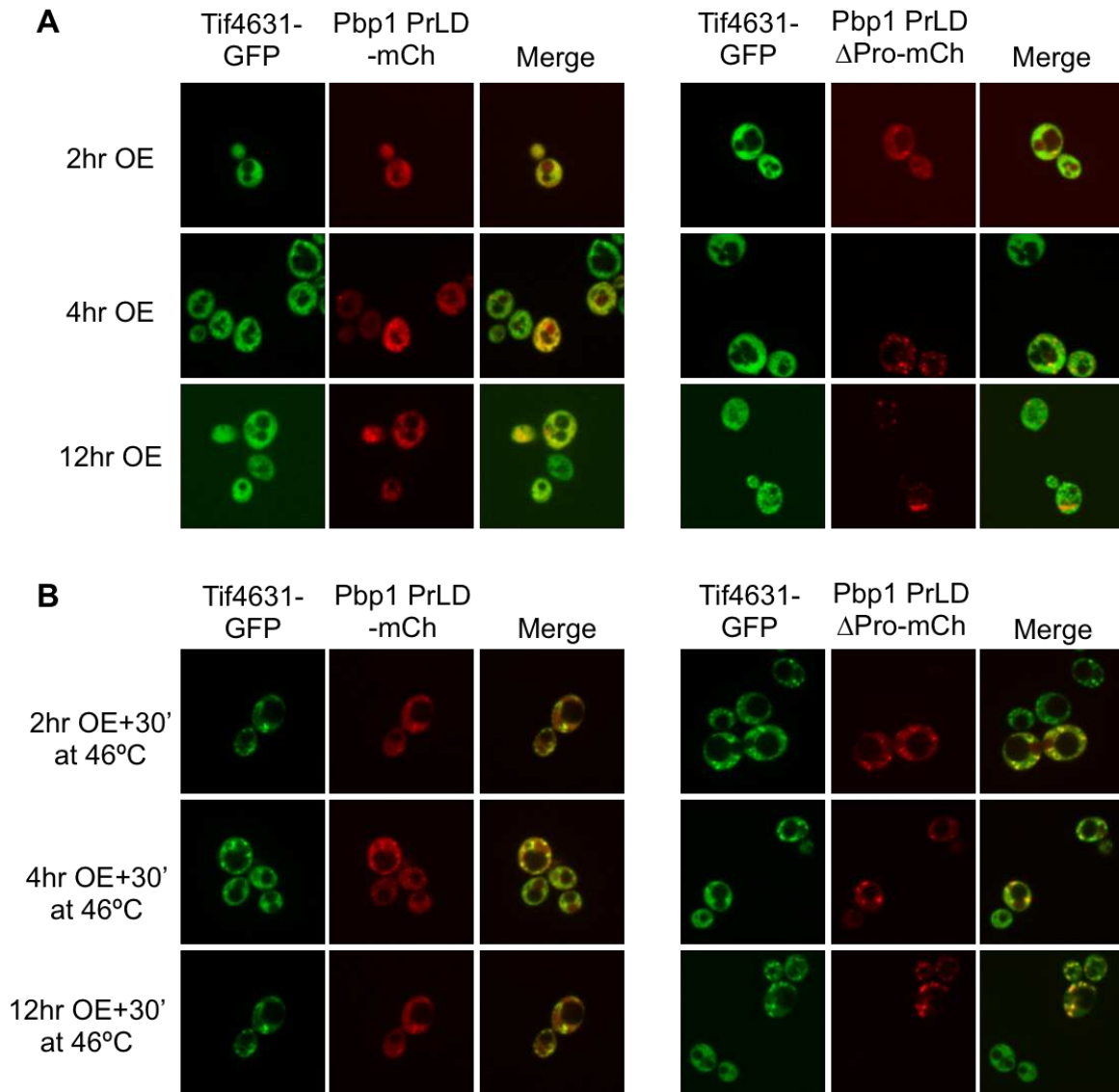


Figure 3.3: Pbp1 PrLD aggregates localize to stress granules upon stress. (A) Each indicated Pbp1 construct was fused to mCherry and overexpressed for 2, 4, and 12 hours prior to imaging. Tif4631-GFP was also co-expressed as a stress granule marker. (B) The Pbp1 PrLD constructs were overexpressed as in (A) and then incubated at 46°C for 30 minutes prior to imaging.

These results are contrary to what was expected based on the results from the previous study with Nrp1 overexpression (35). One potential explanation for this discrepancy is that the specific amyloid species formed by the Nrp1 PrLD are capable of nucleating stress granule formation, but that this ability is unique to the Nrp1 PrLD amyloids and is not generalizable to all amyloid-forming PrLDs. The Pbp1 PrLD and Pbp1 PrLD Δ Pro aggregates might not even be capable of associating with stress granules, let alone nucleating them. To determine if the Pbp1 PrLD and Pbp1 PrLD Δ Pro aggregates are capable of localizing to stress granules, cells were exposed to heat shock at 46°C for 30 minutes after overexpression for 2, 4, or 12 hours. Interestingly, both PrLD constructs colocalize with stress granules after heat shock (Figure 3.3B), suggesting that the initial foci formed by overexpression of the Pbp1 PrLD variants are not very stable assemblies. Additionally, despite extensive mutation, the Pbp1 PrLD Δ Pro construct is still capable of localizing to stress granules, indicating that localization to, and interactions within, stress granules may not be very specific.

Deleting Proline Residues from the Pab1 and Pbp1 PrLDs Does Not Prevent Stress Granule Disassembly

Because proline deletions appear to promote aggregation of these stress granule protein PrLDs (Figures 3.1 & 3.2), perhaps these mutations could similarly increase the aggregation propensity of the endogenously expressed versions of Pab1 and Pbp1 and cause a reduction in stress granule dynamics. Specifically, I wanted to determine whether these mutations were sufficient to cause stress granule persistence, similar to what is observed in disease. Full-length versions of Pab1 and Pbp1 were fused to mCherry and inserted into plasmids. It should be noted that the endogenous version of each protein was still present and untagged in the strains used for these experiments. The stress granule protein Pub1 was also tagged with GFP to serve as a stress granule marker. Cells were exposed to heat shock at 46°C for 30 minutes and then allowed to

recover at 30°C for 2 hours. Cells were imaged under normal conditions, after heat shock, and at 30-minute time intervals during recovery. Surprisingly, deleting the proline residues in the PrLD of Pab1 did not result in stress granule persistence (Figure 3.4A-B). For Pbp1, stress granules appeared to persist slightly more during the 30-minute recovery timepoint in the proline deletion mutant (Figure 3.4C-D), but ultimately stress granules had largely disappeared after 1 hour of recovery. These results indicate that although deleting proline residues effectively increases aggregation of the Pab1 and Pbp1 PrLDs alone, these mutations are not sufficient to reduce stress granule recovery when introduced into the endogenous versions of these proteins.

Deleting Chaperone Proteins Delays Pab1 Recovery After Heat Shock

One possible explanation for why stress granule disassembly was unaffected upon introduction of aggregation-promoting mutations to Pab1 or Pbp1 could be that the PQC machinery is actively disposing of aberrant aggregates formed by these aggregation-prone species. The cell utilizes chaperones, the autophagy pathway, and the ubiquitin-proteasome system to dispose of unwanted aggregates as well as to disassemble stress granules in some cases (36). Initially, I chose to investigate stress granule disassembly upon deletion of the chaperone proteins Ydj1 and Ssa1 because they are both linked stress granule disassembly activity (37). Specifically, I examined whether deleting these chaperones would cause the foci formed by the aggregation-prone Pab1 Δ Pro mutant to persist after heat shock and recovery. I chose to focus on Pab1 for microscopy experiments, since Pbp1 does not express at very high levels and is thus difficult to image (Figure 3.4C). Full-length Pab1-mCherry and Pab1 Δ Pro-mCherry fusions were each transformed into *ydj1* Δ or *ssa1* Δ yeast strains. Cells were subjected to heat shock and then allowed to recover at 30°C for two hours, with images taken every 30 minutes throughout the recovery timecourse. In both the *ydj1* Δ and *ssa1* Δ strains, Pab1 foci disassembly was slower for the Δ Pro mutant than for wild-type Pab1 (Figure 3.5), indicating that these chaperones may

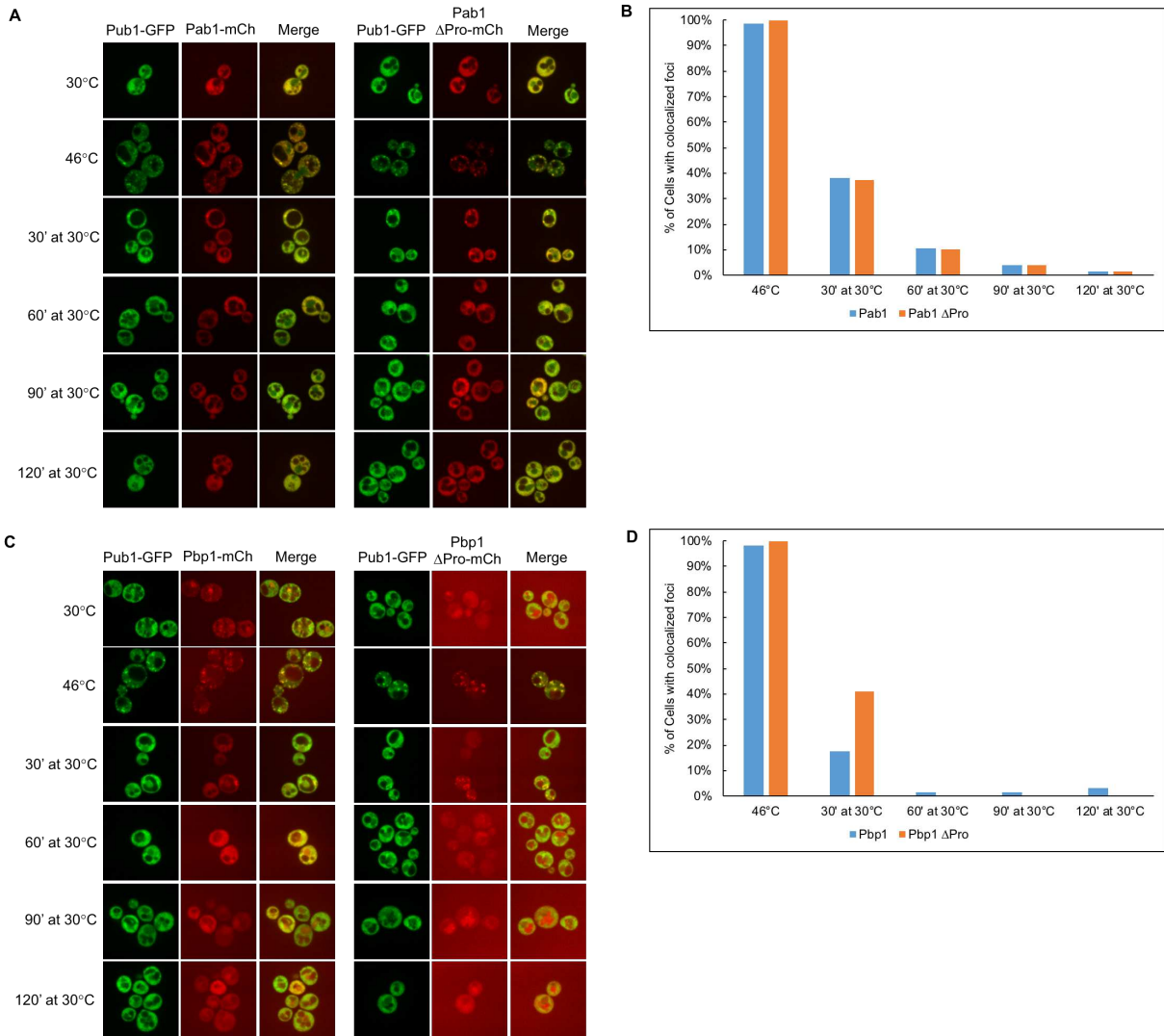


Figure 3.4: Proline deletions do not disrupt stress granule disassembly. Pab1 (A) or Pbp1 (C) with and without proline residues in their PrLDs were fused to mCherry and expressed from their endogenous promoters with Pub1-GFP additionally expressed as a stress granule marker. Cells were subjected to heat shock at 46°C for 30 minutes before being returned to 30°C to recover. Cells were imaged before and after heat shock as well as after 30, 60, 90, and 120 minutes of recovery at 30°C. (B & D) Quantification of percentage of cells containing at least one colocalized focus for each indicated timepoint.

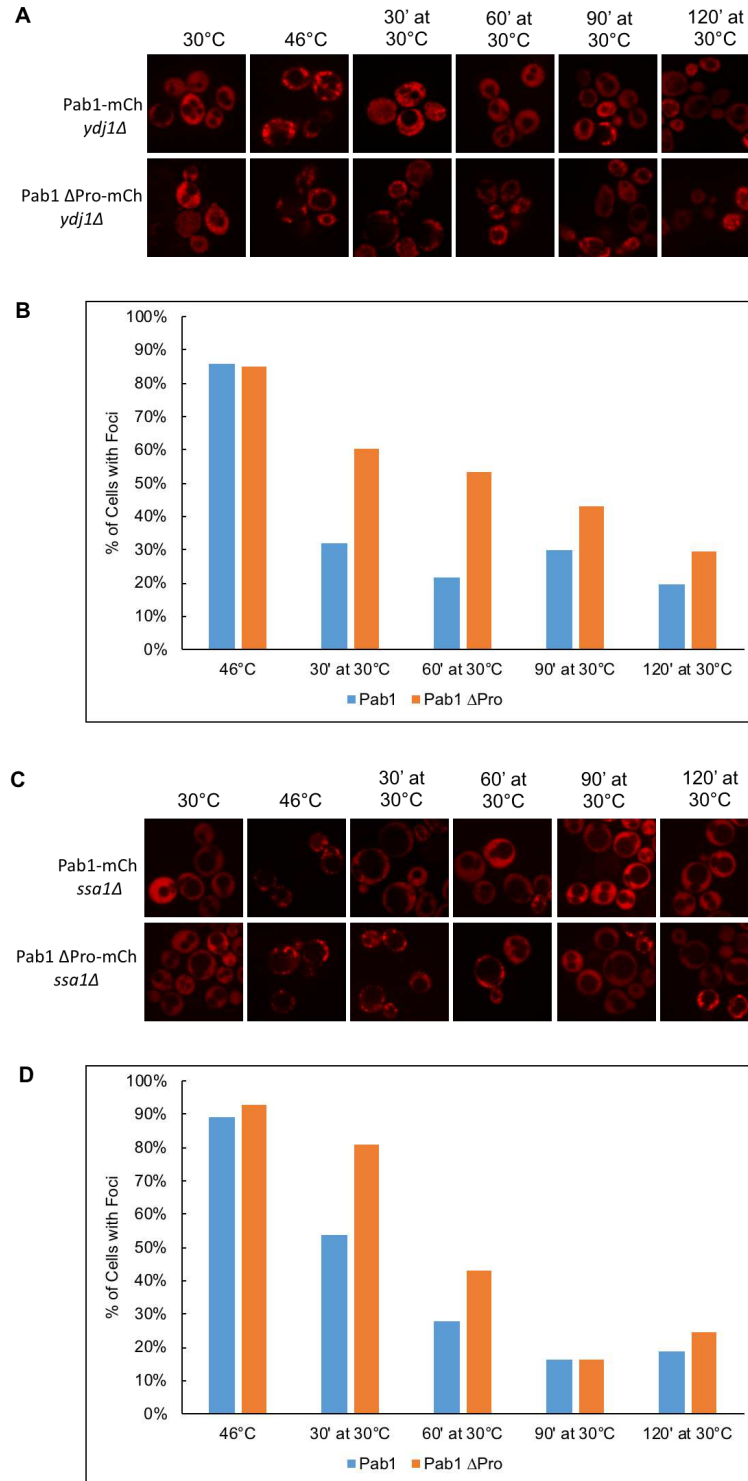


Figure 3.5: Deletion of chaperones slows Pab1 foci disassembly when proline residues are deleted. Each indicated Pab1 construct was fused to mCherry and expressed from its endogenous promoter in a strain containing a knockout of either Ydj1 (A) or Ssa1 (C). Cells were subjected to heat shock at 46°C for 30 minutes before being returned to 30°C to recover. Cells were imaged before and after heat shock as well as after 30, 60, 90, and 120 minutes of recovery at

30°C. **(B & D)** Quantification of percentage of cells containing at least one Pab1-mCherry focus for each indicated timepoint.

be involved in resolving Pab1 Δ Pro aggregates. However, even though the Pab1 Δ Pro foci require more time to disassemble than wild-type Pab1 foci, they do eventually disassemble (Figure 3.5), indicating that other chaperones or other components of the PQC machinery may also contribute to disassembly.

Deleting Proline Residues from the Pab1 and Pbp1 PrLDs Does Not Perturb Cell Survival Upon Chronic Heat Shock

All previous stress experiments in this study were performed using acute heat shock, yet evidence has shown that mutations in the Pab1 PrLD can decrease cell survival upon chronic heat shock (14). These data imply that mutations in the PrLD of Pab1 might affect cell survival in response to this type of stress. To determine whether cells expressing aggregation-prone Pab1 or Pbp1 are able to survive chronic heat shock, cells were plated as serial dilutions and grown for 4 days at both 30°C and 40°C (chronic heat shock condition) as described in (14). Deleting the proline residues from Pab1 and Pbp1 did not appear to affect the ability of cells to survive chronic heat shock (Figure 3.6A), further indicating that these mutations do not adversely affect the cell. Because there was a slight delay in Pab1 Δ Pro foci recovery upon deletion of the Ydj1 and Ssa1 chaperone proteins in response to acute heat shock (Figure 3.5), I chose to examine whether cell survival was inhibited in response to chronic heat shock in the *ydj1* Δ and *ssa1* Δ strains. Although there was no difference in cell survival between the wild-type Pab1 and Pbp1 and the versions of these proteins containing proline deletions in the chaperone deletion strains, cells without Ydj1 were unable to grow at 40°C, regardless of mutation (Figure 3.6B), suggesting an essential role for this chaperone upon chronic stress.

Discussion

Mutations in the PrLDs of different stress granule proteins have been implicated in various neurodegenerative diseases (10, 11, 17), which has spurred interest in understanding how

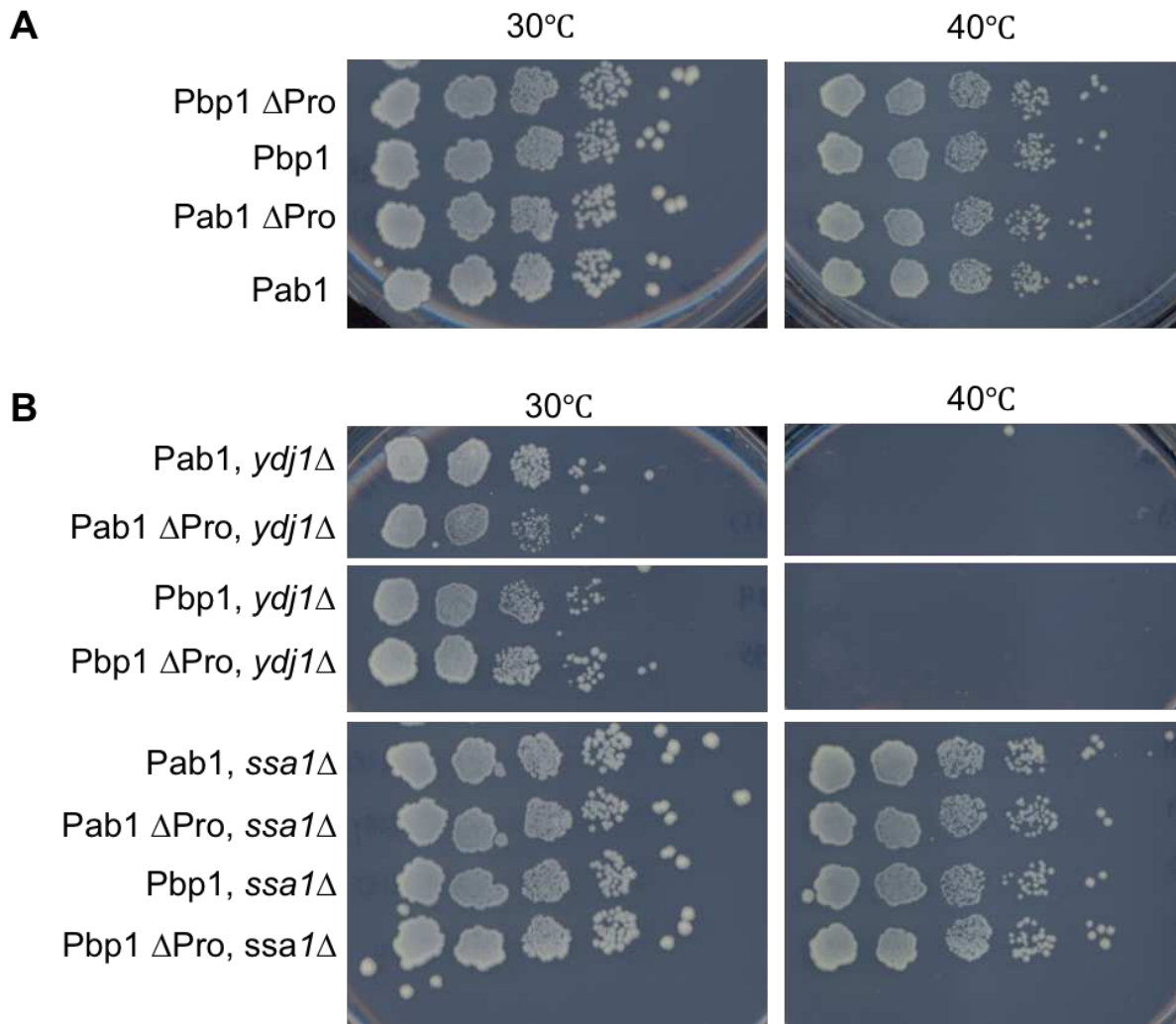


Figure 3.6: Ydj1 is required for cell survival during chronic stress. Each indicated Pab1 or Pbp1 construct was fused to mCherry and expressed from its endogenous promoter in a standard strain background (**A**) or in strains with deletions of either Ydj1 or Ssa1 (**B**). Cells were plated as serial dilutions and incubated at 30°C and 40°C (chronic heat shock) for 4 days before assessing growth.

these mutations promote disease. Evidence suggests that mutation of non-aggregation-prone amino acids to aggregation-promoting amino acids might be causative of the disease-associated aggregation of many disease-associated proteins (11). Using this reasoning, I introduced aggregation-promoting mutations into the PrLDs of two core stress granule proteins, Pab1 and Pbp1, hypothesizing that the mutations would disrupt stress granule disassembly, causing stress granules to persist as aberrant aggregates, thus recapitulating the disease state. Surprisingly, although the aggregation propensity of both PrLDs was successfully increased with these mutations, stress granule disassembly was not disrupted, suggesting that other mechanisms may be acting to control the accumulation of aberrant aggregates, as may be observed in disease.

Because proline residues are disfavored for amyloid formation (21), I chose these residues as targets for mutation to increase aggregate, and perhaps amyloid, formation in the PrLDs of Pab1 and Pbp1. Deleting all of the proline residues in these PrLDs successfully increased the aggregation propensity of each PrLD in isolation (Figure 3.1A and Figure 3.2A), as well as promoting conversion to amyloid in the case of Pbp1 (Figure 3.2B). These data suggest that these proline residues serve to maintain these PrLDs in a more soluble state, preventing conversion to an insoluble, aggregated state, at least at higher concentrations when these domains are highly prone to aggregation.

Remarkably, despite significantly increasing the aggregation propensity of these stress granule protein PrLDs, the proline deletion mutations did not cause stress granules to persist after stress and subsequent recovery (Figure 3.4). One possible explanation for this unexpected result is that these mutations were not aggressive enough to promote stress granule persistence. Deleting proline residues eliminates residues that can prevent aggregation, but no additional mutations were made to further promote aggregation of these PrLDs. Instead of deleting proline

residues, they could be substituted with residues that are more aggregation-promoting, to further enhance the ability of these domains to aggregate within stress granules.

Another explanation for the failure of these mutations to disrupt stress granule dynamics is that the proline deletions in these PrLDs do increase aggregation propensity of Pab1 and Pbp1, but are not sufficient to cause these proteins to remain in an aggregated state upon heat shock. The findings from the Pbp1 PrLD overexpression in addition to heat shock experiment are consistent with this hypothesis. The Pbp1 PrLD Δ Pro construct is sufficient to aggregate upon overexpression, but does not nucleate stress granule formation (Figure 3.3A); however, upon heat shock, these Pbp1 PrLD Δ Pro aggregates appear to relocalize to stress granules. It remains possible that these aggregates are nucleating stress granules, but this explanation seems unlikely because stress granules were not nucleated by these aggregates normally, without heat shock. The fact that these aggregates appear to relocalize to assemble into stress granules suggests that they may not be forming very stable aggregates. Perhaps these aggregates are especially more susceptible to relocalization upon heat shock, since protein denaturation and misfolding are widespread during this particular stress (38). One way to test this hypothesis would be to apply different stresses after overexpression to see if the same relocalization of aggregates occurs.

An alternative explanation for why these proline deletion mutants do not induce persistent aggregation of stress granules is that the PQC machinery might be effective enough at promoting stress granule disassembly to prevent the accumulation of aberrant aggregates. A previous study found that the chaperone Ydj1 and the Hsp70 (Ssa1-4) chaperone machinery both promote stress granule disassembly (37), so I targeted these genes for deletion to assess effects on stress granule disassembly. Deletion of the chaperone proteins Ydj1 and Ssa1 slightly delayed stress granule disassembly when the mutant proteins were expressed, but did not prevent disassembly (Figure 3.5), consistent with previous results demonstrating redundant mechanisms

for this process (36). While deletions of these chaperones individually did not prevent stress granule disassembly, I have not investigated the effects of deleting both chaperones in the same strain. Perhaps a double deletion of Ydj1 and Ssa1 would be more successful at inhibiting stress granule disassembly, rather than each chaperone alone. Ydj1 is connected to stress granule disassembly by promoting reentry to translation; however, another Hsp40 chaperone, Sis1, is also involved in stress granule clearance, by targeting stress granules for autophagy (37). Perhaps inhibition of Sis1 would prevent disassembly of these aggregation-prone mutant-containing stress granules since they are presumably more aberrant. Unfortunately, Sis1 is an essential protein, so it cannot be deleted from the genome, but a knockdown of Sis1 might still provide insight to this question. It should also be noted that the study examining the effects of Ydj1, the Hsp70 chaperones, and Sis1 investigated stress granules induced by sodium azide treatment (37), whereas this study examines stress granules induced by heat shock. Given that stress granule composition varies depending on the type of stress (31), it remains possible that stress granules are disassembled by different mechanisms depending on the stress applied to induce formation.

The heat shock experiments performed throughout this study used only a short, but acute, heat shock at 46°C, because cells cannot survive prolonged acute stress. In contrast, cells can survive a more mild, chronic heat shock, lasting several days (14). Chronic, but mild, stress is also likely more closely related to the types of stress that produce persistent aggregates in disease (39), which might make it a better system to study these aggregation-promoting mutants. I reasoned that cells forming more stably aggregated stress granules (those expressing the proline deletion mutants) would have reduced survival upon chronic heat shock; however, cell survival was not inhibited for the strains containing Pab1 or Pbp1 with proline deletions in the PrLDs (Figure 3.6). This result suggests that these mutations are not having any sort of aberrant effect on cell survival and perhaps these mutations are reasonably benign in an endogenous system.

Interestingly, the molecular chaperone Ydj1 appears to be required for survival during chronic heat shock in yeast. No cells containing a deletion of this chaperone can survive prolonged incubation at 40°C (Figure 3.6B), indicating an essential role for this chaperone. This result highlights differences between acute and chronic stress conditions, because cells are able to recover from acute stress, even without Ydj1 (Figure 3.5A-B), whereas this chaperone is required for survival upon chronic stress (Figure 3.6B). Different mechanisms are likely employed to promote cell survival in response to acute and chronic stress situations.

Understanding these differences might help us to further determine how certain aberrantly aggregated stress granules can persist, while others can be cleared.

It is worth noting that although the stress granules containing the aggregation-prone mutants of Pab1 and Pbp1 are assumed to be more stably aggregated, their dynamics have not been tested. One way to determine the dynamic state of these foci would be to perform Fluorescence Recovery After Photobleaching (FRAP) experiments. Determining whether these proline deletion mutations have any effect on stress granule dynamics specifically will be essential to conclude that stress granules containing these mutant proteins are actually in a more rigid material state than normal stress granules.

These results suggest that although we have a good understanding of how to increase the aggregation propensity of different PrLDs, we do not understand how these domains contribute to the structure and stability of stress granules. Strikingly, despite increasing the aggregation propensity of these PrLDs, stress granule disassembly remains unperturbed, suggesting that these domains might not aggregate in stress granules via canonical prion-like aggregation. Increasing the effective local concentration of aggregation-prone species, as occurs with stress granules, should cause stable aggregation, yet this aggregation remains reversible upon introduction of highly aggregation-prone species. This raises questions as to what types of interactions PrLDs

contribute to these granules and what sequence features in PrLDs promote formation and disassembly of these granules. Understanding how PrLDs normally contribute to stress granule dynamics will help us to gain insight into how certain mutations in these PrLDs can promote conversion of stress granules to more highly aggregated states, as observed in disease.

This study indicates that stress granule dynamics are not easily modulated by increasing the aggregation propensity of constituent proteins, but we have yet to determine exactly why that is. One possibility is that Pab1 and Pbp1 are not important contributors to the dynamic state of stress granules. Perhaps other proteins play a more important role in stress granule dynamics and thus might be better candidates for mutations. Another possibility is that certain components of PQC machinery are successfully removing aberrant stress granules; however, substantial further research into this area would be necessary to determine exactly what mechanism or mechanisms might be responsible. Finally, yeast may have a more robust system in place to remove aberrant aggregates than human cells, indicating that yeast may not be well suited for investigation of aberrantly aggregated stress granules. Altogether, this project turned out to be much more complex than originally thought and would require extensive further research to order reach more concrete conclusions about the mechanisms underlying aberrant stress granule dynamics.

REFERENCES

1. Buchan JR & Parker R (2009) Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* 36(6):932-941.
2. Protter DS & Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26(9):668-679.
3. Li L, McGinnis JP, & Si K (2018) Translational Control by Prion-like Proteins. *Trends Cell Biol*.
4. Gilks N, *et al.* (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15(12):5383-5398.
5. Li YR, King OD, Shorter J, & Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201(3):361-372.
6. Ramaswami M, Taylor JP, & Parker R (2013) Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154(4):727-736.
7. Taylor JP, Brown RH, & Cleveland DW (2016) Decoding ALS: from genes to mechanism. *Nature* 539(7628):197-206.
8. Liu-Yesucevitz L, *et al.* (2010) Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* 5(10):e13250.
9. Dewey CM, *et al.* (2012) TDP-43 aggregation in neurodegeneration: are stress granules the key? *Brain Res* 1462:16-25.
10. King OD, Gitler AD, & Shorter J (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* 1462:61-80.
11. Kim HJ, *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495(7442):467-473.
12. Patel A, *et al.* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162(5):1066-1077.
13. Molliex A, *et al.* (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* 163(1):123-133.
14. Riback JA, *et al.* (2017) Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* 168(6):1028-1040 e1019.
15. Conicella AE, Zerze GH, Mittal J, & Fawzi NL (2016) ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure*.
16. Ryan VH, *et al.* (2018) Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell* 69(3):465-479 e467.
17. Mackenzie IR, *et al.* (2017) TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* 95(4):808-816 e809.
18. Brangwynne Clifford P, Tompa P, & Pappu Rohit V (2015) Polymer physics of intracellular phase transitions. *Nature Physics* 11(11):899-904.
19. Lin Y, Protter DS, Rosen MK, & Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60(2):208-219.

20. Murakami T, *et al.* (2015) ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* 88(4):678-690.
21. Theillet FX, *et al.* (2013) The alphabet of intrinsic disorder: I. Act like a Pro: On the abundance and roles of proline residues in intrinsically disordered proteins. *Intrinsically Disord Proteins* 1(1):e24360.
22. Sherman F (1991) Getting started with yeast. *Methods in enzymology* 194:3-21.
23. Brachmann CB, *et al.* (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14(2):115-132.
24. Huh WK, *et al.* (2003) Global analysis of protein localization in budding yeast. *Nature* 425(6959):686-691.
25. Winzler EA, *et al.* (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285(5429):901-906.
26. Gietz RD & Sugino A (1988) New yeast-*Escherichia coli* shuttle vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites. *Gene* 74(2):527-534.
27. Shattuck JE, Waechter AC, & Ross ED (2017) The effects of glutamine/asparagine content on aggregation and heterologous prion induction by yeast prion-like domains. *Prion*:1-16.
28. Toombs JA, McCarty BR, & Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-332.
29. Buchan JR, Muhlrاد D, & Parker R (2008) P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 183(3):441-455.
30. Grousl T, *et al.* (2009) Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *J Cell Sci* 122(Pt 12):2078-2088.
31. Buchan JR, Yoon JH, & Parker R (2011) Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* 124(Pt 2):228-239.
32. Jain S, *et al.* (2016) ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164(3):487-498.
33. Bagriantsev SN, Kushnirov VV, & Liebman SW (2006) Analysis of amyloid aggregates using agarose gel electrophoresis. *Methods in enzymology* 412:33-48.
34. LeVine H, 3rd (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods in enzymology* 309:274-284.
35. Kroschwald S, *et al.* (2015) Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife* 4:e06807.
36. Alberti S, Mateju D, Mediani L, & Carra S (2017) Granulostasis: Protein Quality Control of RNP Granules. *Front Mol Neurosci* 10:84.
37. Walters RW, Muhlrاد D, Garcia J, & Parker R (2015) Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in *Saccharomyces cerevisiae*. *RNA* 21(9):1660-1671.
38. Verghese J, Abrams J, Wang Y, & Morano KA (2012) Biology of the heat shock response and protein chaperones: budding yeast (*Saccharomyces cerevisiae*) as a model system. *Microbiol Mol Biol Rev* 76(2):115-158.
39. Alberti S & Carra S (2018) Quality Control of Membraneless Organelles. *J Mol Biol.*

CHAPTER FOUR: STRESS-INDUCED ASSEMBLY OF PRION-LIKE DOMAINS IS STRONGLY INFLUENCED BY COMPOSITION¹

Introduction

Stress granules are cytoplasmic, membraneless ribonucleoprotein (RNP) assemblies containing mRNAs stalled in translation initiation (1). Stress granules form in response to various stresses and dissipate once the stress-inducing conditions are eliminated (1). Many of the RNA-binding proteins found in stress granules contain prion-like domains (PrLDs), which are glutamine/asparagine (Q/N) rich domains that compositionally resemble yeast prion domains (2). Recently, mutations in a number of these PrLD-containing RNA-binding proteins have been implicated in various neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (3).

PrLDs have attracted recent interest for a few reasons. First, some of these domains have been shown to help target proteins to stress granules (4). Second, intrinsically disordered regions (IDRs), of which PrLDs are a sub-class, are thought to provide promiscuous, but potentially stabilizing interactions in RNP granules (5, 6). Finally, and perhaps most interestingly, several of the disease-causing mutations in stress granule-associated proteins occur within PrLDs and are associated with irreversible aggregation of the PrLD-containing protein in affected individuals (7, 8). These observations have led to the idea that these PrLDs normally contribute to the functional, reversible aggregation of stress granules, but that aggregation-promoting mutations in these domains can negatively affect their normal dynamic behavior (3, 7, 9).

¹ This chapter is a manuscript in preparation for publishing. Jenifer E. Shattuck conceived the project and designed the original heat shock assay. Jenifer E. Shattuck, Kacy R. Paul, and Andrew Lamb built GFP-PrLD fusions for the original dataset (listed in Table 4.1) and performed preliminary heat shock experiments with these constructs. Sean M. Cascarina performed PrLD composition analysis (Figure 4.8-4.9 and Table 4.2). Eric D. Ross generated the stress assembly predictor for the heat shock PrLD dataset (Figure 4.10A).

Some constituent proteins of stress granules are able to undergo liquid-liquid phase separation (LLPS) *in vitro* (10-14), and these proteins seem to retain liquid-like qualities when they assemble *in vivo* (11, 12, 15). These results have led to a model in which stress granules are liquid-like RNP compartments (herein referred to as “assemblies”) resulting from LLPS of their protein and RNA components. Many PrLDs can also phase separate *in vitro*, with amino acid sequence and composition affecting LLPS ability (10, 14, 16-18). Interestingly, upon introduction of disease-associated mutations, PrLDs are more prone to form stable, solid-phase aggregates (11, 12, 15, 19). Substantial research has provided insight into the sequence features that promote these transitions to amorphous aggregates and more structured amyloids (20-22); however, the sequence features that govern the formation of reversible, liquid-like assemblies *in vivo* have not been fully defined.

To gain more insight into the sequence requirements for stress granule recruitment, we examined the response of various PrLDs to stress. Using the prion prediction algorithms PAPA and PLAAC (23, 24), we identified a variety of PrLDs in *Saccharomyces cerevisiae*, and tested these PrLDs for assembly into foci in response to heat stress. We found that many of these PrLDs reversibly assemble into foci that colocalize with stress granule markers. Additionally, most PrLDs showed similar assembly activity under two other stresses: sodium azide (NaN₃) treatment (oxidative stress), and sorbic acid treatment (pH stress). PrLDs that assembled into foci in response to stress showed substantial compositional biases, including a significant over-representation of both charged and hydrophobic amino acids. These compositional biases were sufficient to predict with reasonable accuracy which PrLDs would localize to stress granules response to heat shock, and to design mutations to modulate assembly activity. Interestingly, PrLD composition, rather than primary sequence, appears to be the determining factor dictating stress-induced assembly, including recruitment to stress granules.

Materials and Methods

Strains and Growth Conditions

All experiments were performed in *Saccharomyces cerevisiae* using standard yeast media and growth conditions (25). All plasmids were transformed into either YER1405 (*MATa his3D1 leu2D0 met15D0 ura3D0 Pab1-mcherry::URA3*) or YER1997 (*MATa his3D1 leu2D0 met15D0 ura3D0 Pbp1-HA-mCherry*) using standard yeast transformation methods (25). Both strains are derivatives of BY4741 (26). Yeast were grown at 30°C unless otherwise specified.

Cloning PrLDs

Amino acid coordinates for each PrLD are listed in Table S1. Var1, AI3, YML053C, Cdc39, and Fab1 all have slight sequence changes relative to the reference strains in the *Saccharomyces* Genome database, as described in (27). PrLDs were first amplified from yeast genomic DNA and then amplified a second time with primers containing tails for cloning into pER843. pER843 was built from pJ526 (28) by inserting GFP between the HindIII and BamHI restriction sites following the *SUP35* promoter. PrLDs were cloned into this plasmid after GFP (in-frame) between the BamHI and BglII sites.

Mutation Design

For each PrLD that was mutated from a negative to a positive (Mfg1, Pub1, and Sro9), Q, N, A, P, M, and T residues within the PrLD were randomly selected using an Excel random number generator; these residues were replaced with randomly selected E, V, F, I, R, K, L, and D residues until the score rose above 0.14. For PrLDs that were mutated from positive to negative (Prt1, Trm1, and Rsc8), C, D, E, F, I, K, L, R, V, and W were randomly chosen, and replaced with a random mixture of Q, N, A, P (in a 2:2:1:1 ratio of Q:N:A:P) until the score dropped below -0.14. To design the scrambled PrLDs, sequences were randomly shuffled using the Excel

random number generator. Both mutant and scrambled PrLDs were built synthetically using overlapping primers and then cloned into pER843 as described above.

Stress Conditions

Cells were grown to mid-log phase ($OD_{600} \approx 0.4-0.7$) at 30°C in SC-Leu media (to select for the plasmids) prior to stress induction. For heat shock, 1mL of cells was concentrated to 50 μ L and incubated in a 46°C water bath for 30 minutes prior to imaging. For oxidative stress, NaN_3 was added to cells to a final concentration of 0.5% (v/v) and cells were incubated at 30°C with shaking for 30 minutes. Following NaN_3 treatment, cells were then concentrated to 50 μ L and imaged. For pH stress, 1mL of cells were harvested and media exchanged for 6mM sorbic acid in SC-Leu. Cells were incubated at 30°C for 30 minutes with rotation or shaking and then concentrated to 50 μ L prior to imaging.

Confocal Fluorescence Microscopy

Imaging was performed on an Olympus (IX83) Inverted Spinning Disk Confocal Microscope using a 100X objective. Images were captured as single planes.

Western Blotting

Cells were grown to mid-log phase ($OD_{600} \approx 0.45-0.6$) at 30°C in SC-Leu media (to select for the plasmids) before harvesting. 30mL of the lowest density culture was collected and the volumes collected for each other culture were normalized to the lowest density culture as assessed by OD_{600} measurements. Cells were harvested by centrifugation at 3,000rpm for 5 minutes at 4°C. Cell lysis was performed as described previously (29). 15 μ L of each sample was run on a polyacrylamide gel and then transferred onto a PVDF membrane. GFP-PrLD fusions were probed using an anti-GFP antibody (Santa Cruz Biotechnology).

PrLD Composition Analyses

PrLD compositions were analyzed using in-house Python scripts. Briefly, the percent compositions of individual amino acids (Fig. S6) or amino acid groups (Fig. 3) were calculated for each PrLD. For each amino acid or amino acid group, the composition distribution associated with PrLDs localizing to stress granules was statistically compared (two-sided Mann-Whitney U test) to the composition distribution associated with PrLDs that did not localize to stress granules, and plotted as adjacent boxplots. Composition analyses were performed independently for each stress condition. Plotting and statistical tests were performed using the Matplotlib/Seaborn and Scipy packages respectively.

Algorithm Generation

For each amino acid, an odds ratio (OR_{aa}) was determined as:

$$OR_{aa} = \left[\frac{f_a}{1-f_a} \right] / \left[\frac{f_n}{1-f_n} \right] \quad (1)$$

where f_a is the mean frequency of occurrence of the amino acid among assembly-prone PrLDs, and f_n is the mean frequency of occurrence among non-assembling PrLDs. The assembly propensity of each amino acid was defined based on its log-odds ratio. To predict the assembly propensity of a PrLD, the frequency of occurrence of each amino acid in the PrLD was multiplied by the amino acid's log-odds ratio, and these values were summed.

Results

PrLDs Form Reversible Assemblies Upon Heat Shock

To gain a greater understanding of how PrLDs might contribute to stress granule assembly we investigated the behavior of isolated PrLDs in response to stress. We chose 56 PrLDs from two pre-existing yeast datasets. The first dataset was derived from a study by Wallace *et al.*, who utilized mass spectrometry to identify proteins that reversibly assemble in response to heat stress (30); we screened this dataset with the prion prediction algorithms PAPA

and PLAAC (23, 24), and found that 19 of the proteins contained PrLDs. The second dataset was from a previous study in which the yeast proteome was screened with PAPA to identify proteins with high-scoring PrLDs (27).

To examine PrLD assembly upon stress, each PrLD was fused to the C-terminus of GFP and expressed from the constitutive, intermediate-strength *SUP35* promoter. Localization of these PrLD fusions was examined before and after 30 minutes of heat shock at 46°C. About one-fifth of the GFP-PrLD fusions were not detectable on a western blot, and were therefore excluded from further analysis (Figure 4.1 and Table 4.1). Additionally, PrLDs that were membrane-associated or localized to the nucleus were also excluded (Table 4.1). Of the remaining 38 PrLDs, 13 assembled into distinct foci in >60% of cells, 20 rarely or never assembled into foci, and 5 formed foci in a moderate fraction of cells (Figure 4.2A, Figure 4.3, and Table 4.1). Although a few of the foci-forming PrLDs showed a small number of foci during normal growth at 30°C (including *AI3*, *Vac14*, and *Cdc73*), each had striking increase in foci upon stress. It should be noted that many of the PrLDs expressed poorly or showed multiple bands by western blot analysis (Figure 4.1); however, despite degradation and low protein levels, many PrLDs are still capable of assembling into foci in response to stress.

Many stress-induced assemblies are able to disassemble upon stress relief (31). We therefore assessed the ability of foci formed by each assembly-prone PrLD to dissipate post-stress. After heat shock, cells were incubated at 30°C for two hours to allow recovery. Of the PrLDs that were able to assemble upon heat shock, most were able to revert back to their soluble form when the cells were allowed to recover (Figure 4.2A and Figure 4.3A). Only *AI3* and *Cdc73* showed incomplete dissolution of stress-induced foci. These data suggest that these PrLDs are not forming insoluble, irreversible aggregates, but rather dynamic, reversible assemblies that are more analogous to RNP granules.

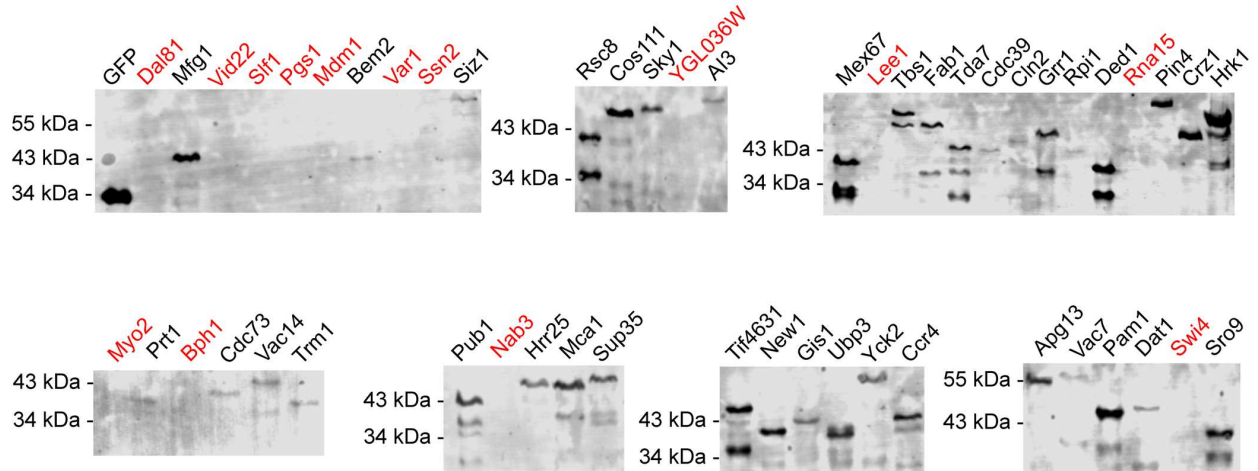


Figure 4.1: Western blot analysis of PrLDs tested. PrLDs in red were eliminated from further analysis due to lack of GFP signal via microscopy and lack of a detectable band on western blot. Despite the presence of weak bands or multiple bands, possibly indicating lower protein levels, many of these PrLDs are still capable of forming distinct foci upon stress.

Table 4.2: List of PrLDs analyzed in this study with assembly results from each stress.

Protein	PrLD Amino Acid Range	Heat Shock	NaN₃	Sorbic Acid	Heat Shock Assembly Score
Myo2	417-497	N. E.	N. E.	N. E.	0.27
Bph1	1113-1243	N. E.	N. E.	N. E.	0.26
Bem2	1800-1880	+/-	+/-	+/-	0.22
Izh3	176-492	M. L.	M. L.	M. L.	0.22
Rsc8	232-312	+	-	-	0.22
Cdc73	253-378	+	+	+	0.21
Prt1	193-273	+	+	+	0.2
Vac14	690-818	+	+	+	0.2
Mdm1	745-864	N. E.	N. E.	N. E.	0.19
Cos111	336-465	+	+	+	0.18
Trm1	286-366	+	+	+	0.18
Sky1 [‡]	353-491	+	-	+/-	0.17
Nte1	1-169	M. L.	M. L.	M. L.	0.16
Pgs1	158-277	N. E.	N. E.	N. E.	0.15
YGL036W	270-478	N. E.	N. E.	N. E.	0.15
YML053C	34-148	N. L.	N. L.	N. L.	0.15
Siz1	390-554	+	+/-	+	0.12
Vid22	641-805	N. E.	N. E.	N. E.	0.11
Al3	228-387	+	+	+	0.1
Ssn2	1025-1211	N. E.	N. E.	N. E.	0.08
Mex67	1-95	-	-	-	0.07
Rpi1	192-306	+	-	+	0.06
Var1	191-349	N. E.	N. E.	N. E.	0.05
Fab1	427-552	+/-	-	+/-	0.05
Lee1	151-301	N. E.	N. E.	N. E.	0.04
Cln2	362-503	+/-	-	+/-	0.04
Tbs1	898-1062	-	-	-	0.03
Tda7	513-636	-	-	+/-	0.03
Grr1	3-167	+/-	-	-	0.03
Slf1	183-311	N. E.	N. E.	N. E.	0.02
Ded1	1-97	+	-	+/-	0.02
Rna15	39-169	N. E.	N. E.	N. E.	0.01
Cdc39	966-1092	+	+	+	0.01
Pin4	169-492	-	-	-	-0.01
Vac7	377-541	-	-	-	-0.01
Npl3*	276-414	N. L.	N. L.	N. L.	-0.02
Pam1	617-756	-	-	-	-0.02
Hrk1	483-647	-	-	-	-0.02
Swi4	177-380	N. E.	N. E.	N. E.	-0.03
Crz1	15-179	-	-	-	-0.04
Gis1	454-584	-	-	-	-0.05
Apg13	250-414	+	+	+	-0.05
Dal81	4-168	N. E.	N. E.	N. E.	-0.05
Dat1	102-236	+/-	-	+/-	-0.07

Sro9	160-256	-	-	-	-0.14
Tif4631	1-131	-	-	-	-0.15
New1	1-118	-	-	-	-0.16
Ccr4	1-147	-	-	-	-0.17
Ubp3	1-97	-	+/-	-	-0.22
Yck2	369-533	-	-	-	-0.24
Pub1	243-327	-	-	-	-0.26
Mfg1	1-96	-	-	-	-0.3
Hrr25	395-494	-	-	-	-0.33
Nab3	559-802	N. E.	N. E.	N. E.	-0.36
Mca1	1-104	-	-	-	-0.39
Sup35	1-123	-	-	-	-0.4

A “+” indicates that the PrLD formed foci in $\geq 60\%$ of cells, a “-“ indicates that the PrLD formed foci in 0-25% of cells, a “+/-“ indicates that the PrLD formed foci in 26-59% of cells, N. E. indicates no expression of the GFP-PrLD fusion, N. L. indicates that the PrLD appeared to localize to the nucleus, and M. L. indicates that the PrLD appeared to localize to a membrane, likely precluding the formation of cytoplasmic foci.

*Npl3 did relocate from the nucleus to cytoplasmic foci occasionally during NaN_3 stress, but was left out of further analysis because it started out in the nucleus and all other proteins analyzed were initially cytoplasmic.

‡The assembly activity of the Sky1 has been separately reported as part of a study examining the role of Sky1 in stress granule dissolution (Shattuck *et al.*, submitted).

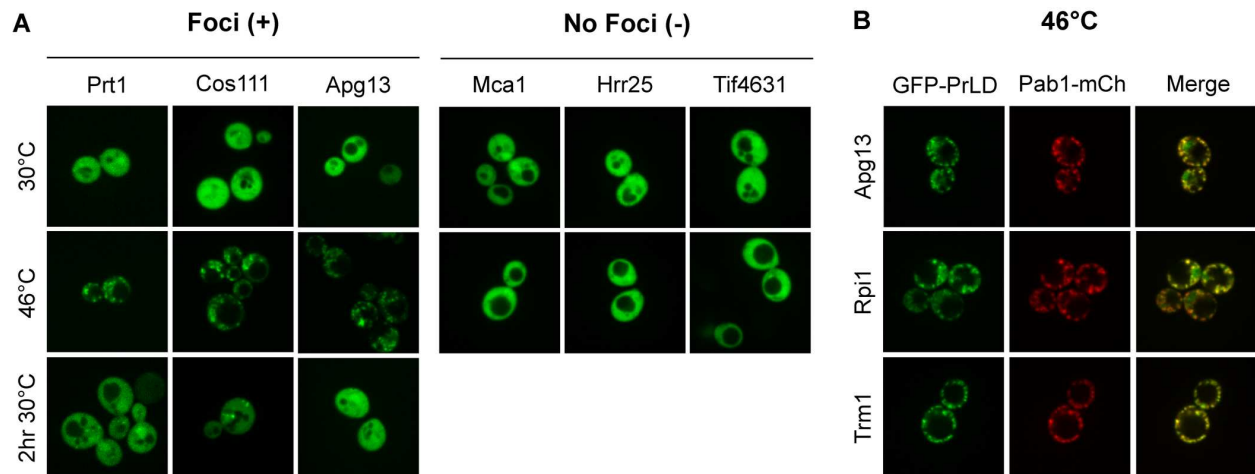


Figure 4.2: Certain PrLDs are sufficient to reversibly assemble upon stress. (A) wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of heat shock at 46°C. PrLDs that formed foci upon heat shock were also imaged after 2 hours of recovery post-heat shock at 30°C. **(B)** Foci-forming PrLDs were co-expressed with Pab1-mCh as a stress granule marker to assess colocalization with stress granules. Cells were exposed to 30 minutes of heat shock at 46°C prior to imaging.

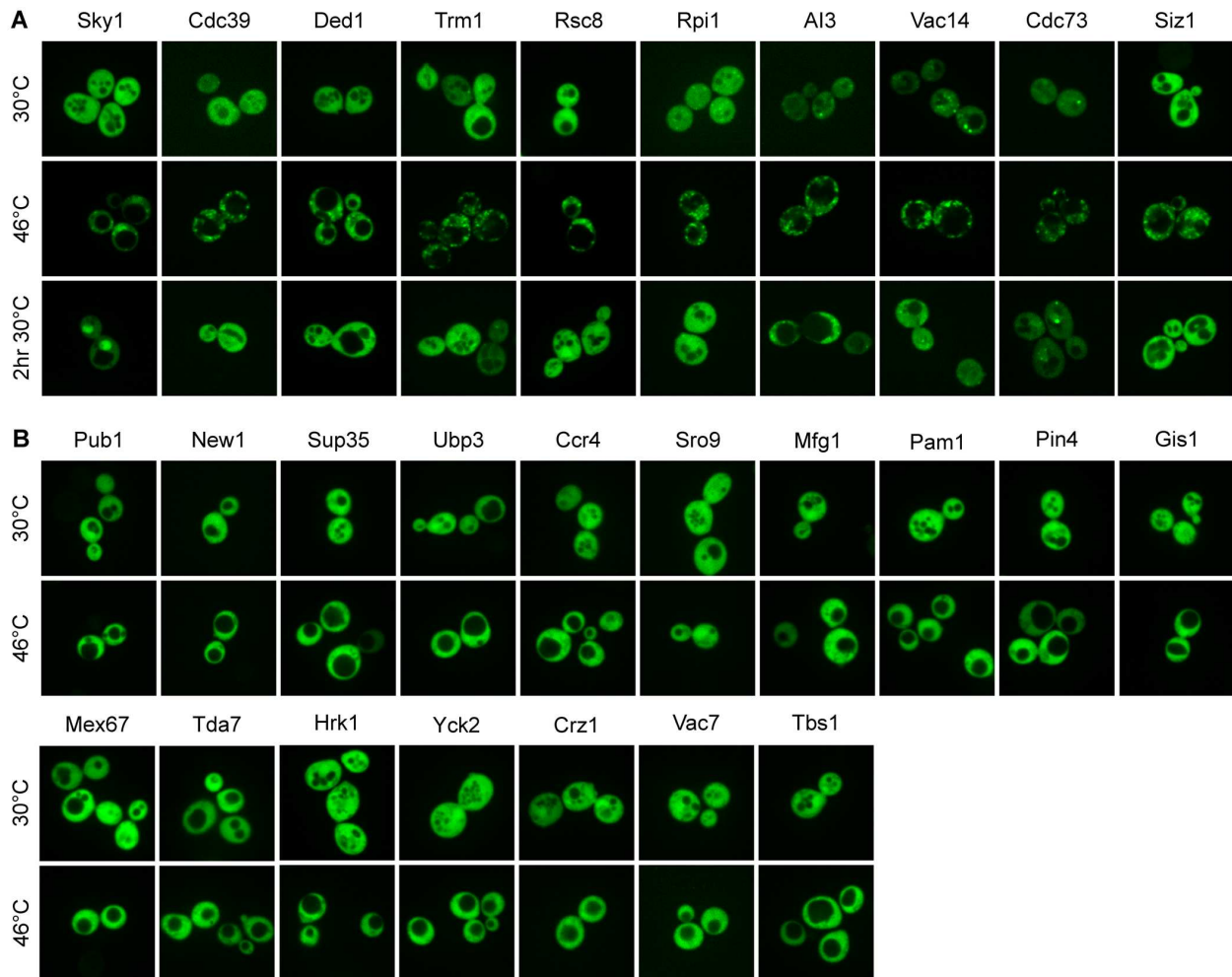


Figure 4.3: Certain PrLDs are sufficient to reversibly assemble upon stress. wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of heat shock at 46°C. **(A)** PrLDs that are sufficient to form foci upon heat shock. These PrLDs were also imaged after 2 hours of recovery post-heat shock at 30°C. **(B)** PrLDs that are not sufficient to form foci upon heat shock.

We next asked whether these PrLDs were localizing to stress granules. Each PrLD that assembled upon stress was co-expressed with mCherry-tagged Pab1, a known stress granule protein (32). Cells were exposed to heat shock at 46°C for 30 minutes and colocalization of Pab1-mCh and each PrLD was evaluated. Strikingly, all of the assembling PrLDs colocalized with Pab1-mCh (Figure 4.2B and Figure 4.4), indicating that PrLDs alone are sufficient to be recruited to stress granules. Similar results were seen with another stress granule marker, Pbp1-mCh (data not shown).

PrLDs Form Stress-Induced Assemblies in Response to Different Stresses

Different stresses result in stress granules containing different protein components (33). We therefore tested whether our PrLDs would show similar responses to other stresses. We exposed the same PrLDs to 30-minute treatments with either 0.5% NaN₃, which results in oxidative stress (33), or 6 mM sorbic acid, which causes pH stress (34). Interestingly, most PrLDs showed similar behavior across the various stresses; although slightly fewer PrLDs assembled in response to oxidative and pH stress (8 and 10 PrLDs, respectively), every PrLD that assembled under either of these two stresses also assembled in response to heat stress (Figure 4.5, Figure 4.6-4.7, and Table 4.1).

Composition is the Primary Determinant of Stress-Induced Assembly

The tested PrLDs provide a useful dataset to examine the sequence features that promote stress-induced assembly. Because there is evidence that amino acid composition contributes greatly to LLPS of PrLDs (16-18), we first examined whether each amino acid was over- or under-represented among the proteins that assembled under each stress. Strikingly, the major compositional biases were consistent across all three stresses. Charged and hydrophobic residues were overrepresented in PrLDs that were assembled in response to each stress, whereas glutamine, asparagine, proline, and alanine were overrepresented in PrLDs that did not show

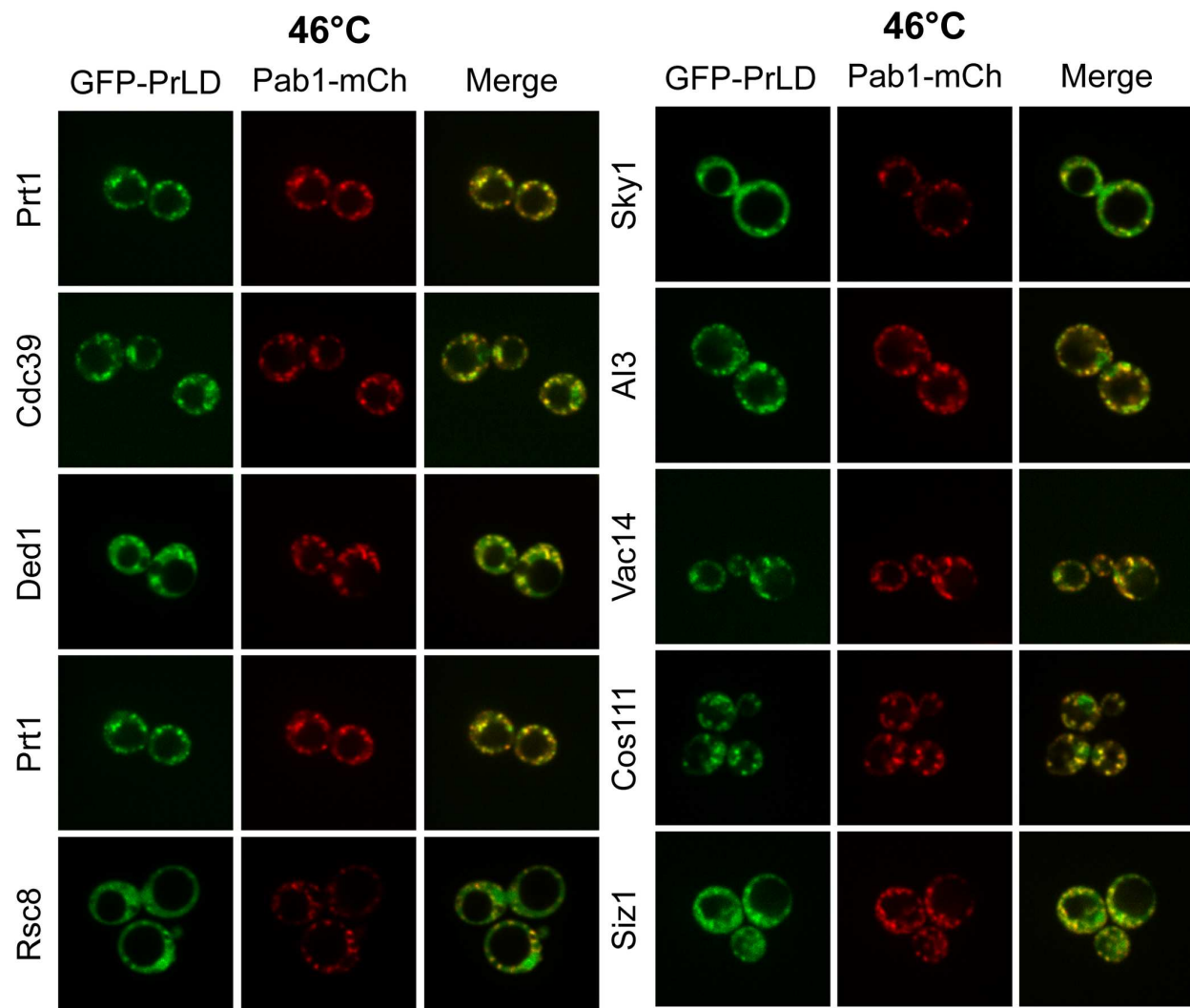


Figure 4.4: Foci-forming PrLDs are recruited to stress granules. Foci-forming PrLDs were co-expressed with Pab1-mCh as a stress granule marker to assess colocalization with stress granules. Cells were exposed to 30 minutes of heat shock at 46°C prior to imaging.

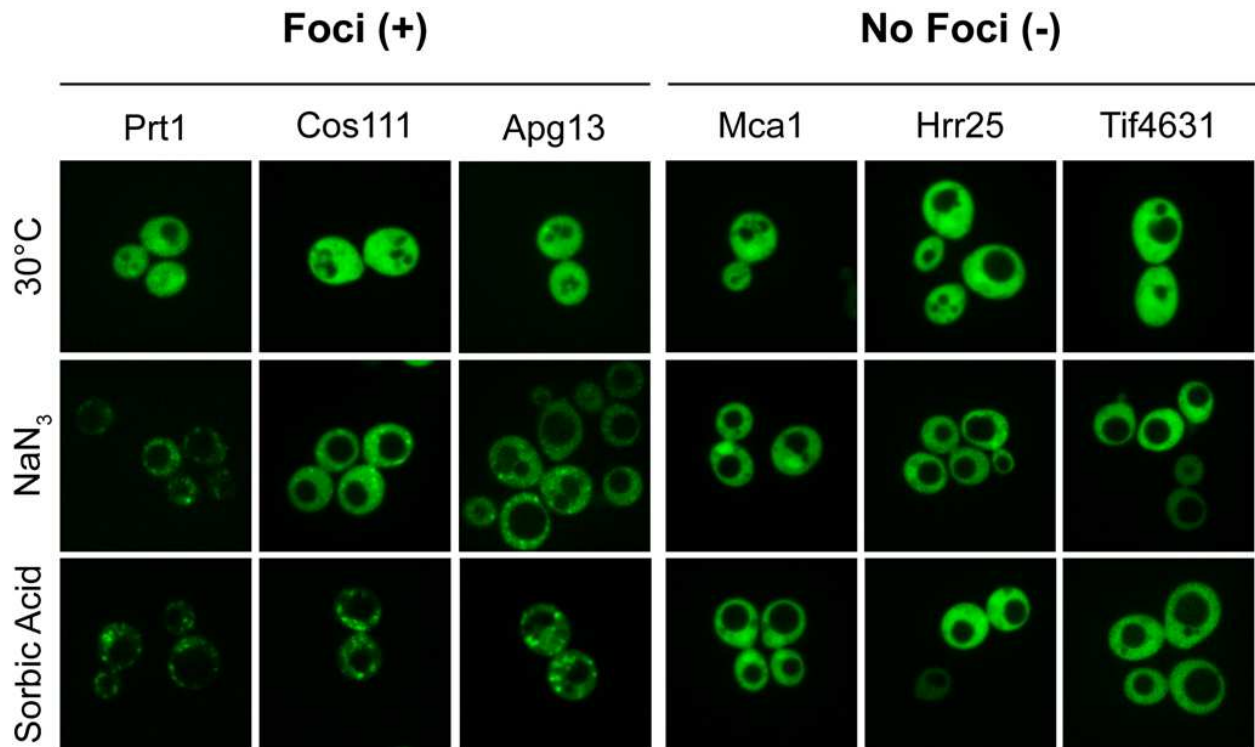


Figure 4.5: PrLDs can assemble in response to oxidative and pH stresses. wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of incubation with 0.5% NaN₃ or 30 minutes of incubation with 6mM sorbic acid.

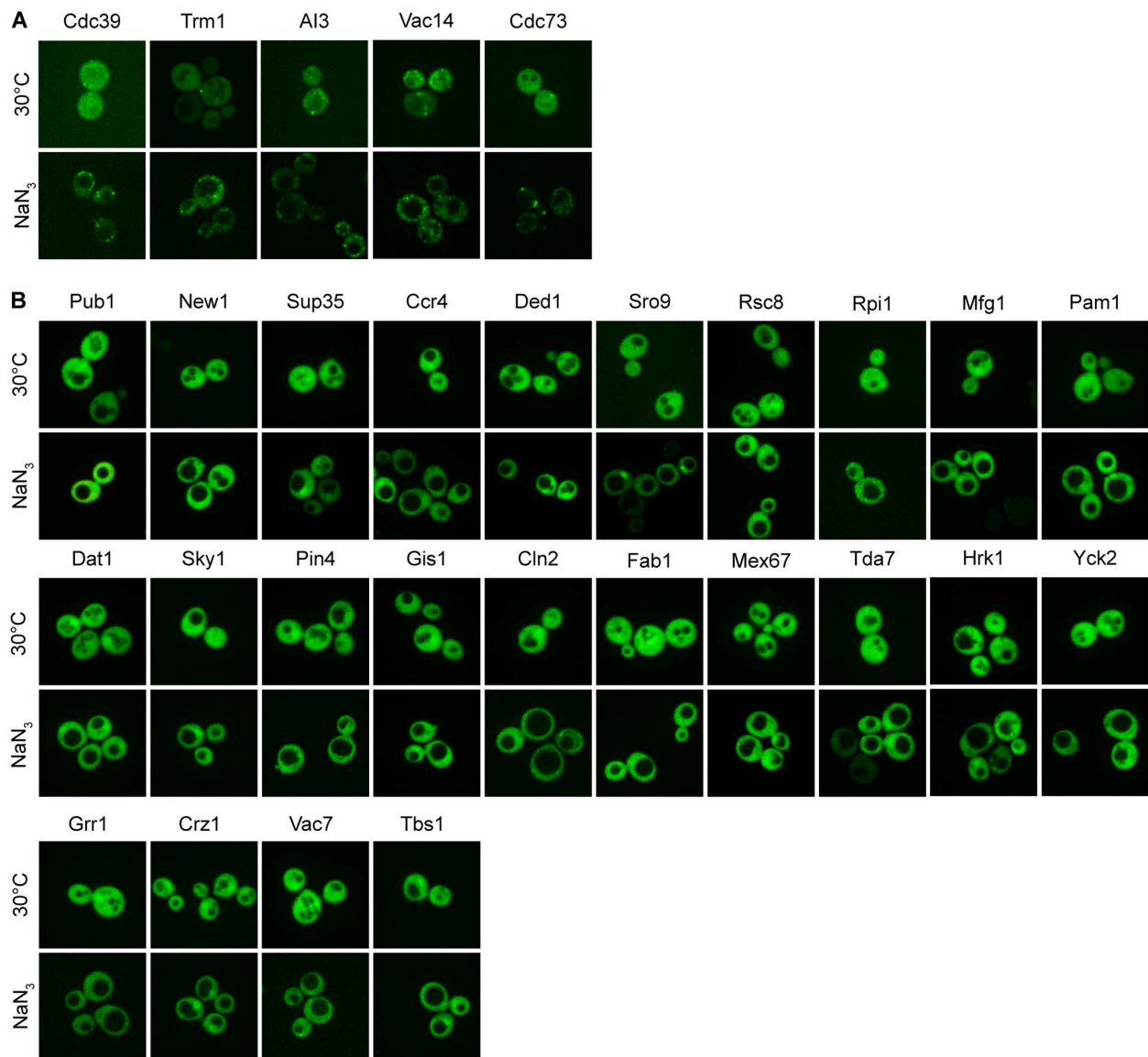


Figure 4.6: PrLDs can assemble in response to oxidative stress. wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of incubation with 0.5% NaN₃. **(A)** PrLDs that are sufficient to form foci upon oxidative stress. **(B)** PrLDs that are not sufficient to form foci upon oxidative stress.

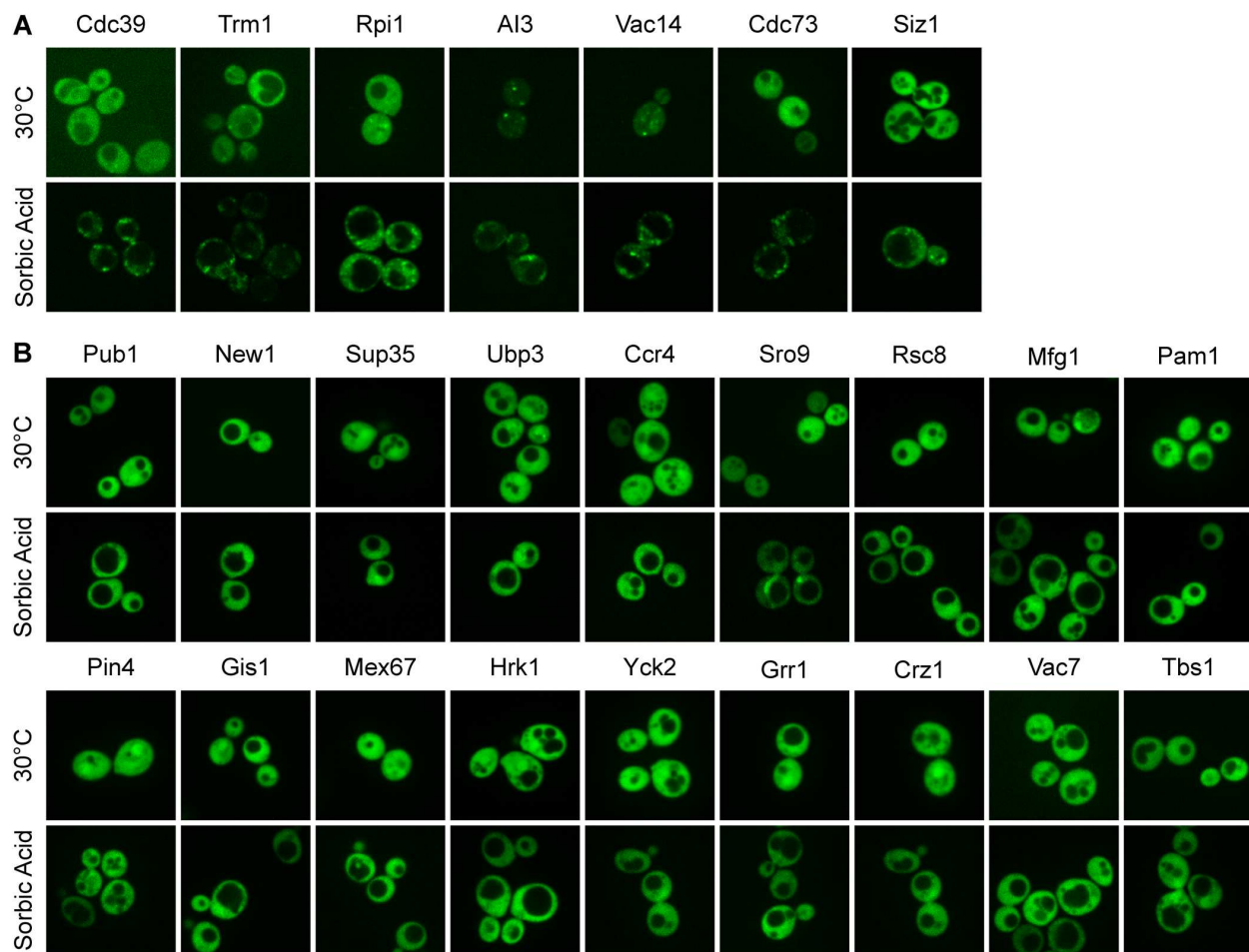


Figure 4.7: PrLDs can assemble in response to pH stress. wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of incubation with 6mM sorbic acid. **(A)** PrLDs that are sufficient to form foci upon pH stress. **(B)** PrLDs that are not sufficient to form foci upon pH stress.

stress-dependent assembly (Figure 4.8, Figure 4.9, and Table 4.2). These compositional biases are surprising since they contradict the features that make PrLDs prion-like. Yeast prion domains are generally enriched in Q/N residues and depleted in charged and hydrophobic residues (23, 35). These results strongly suggest that among PrLDs, distinct sequence features maximize reversible stress-induced assembly versus stable prion aggregate formation.

Given the compositional biases observed among assembly-prone PrLDs, we examined whether these biases were sufficient to predict which PrLDs would show stress-induced assembly. Because the compositional biases were similar across all three stresses, we focused our follow-up investigations on heat stress. For each amino acid, we calculated the mean frequency of occurrence among assembly-forming (positive) PrLDs and non-assembly-forming (negative) PrLDs. These values were used to calculate a log-odds ratio for each amino acid, representing the degree of over or underrepresentation of that amino acid in PrLDs that form assemblies (Table 4.3). To score the predicted assembly activity of each PrLD, for each amino acid we multiplied the frequency of occurrence of the amino acid in the PrLD by the log-odds ratio for that amino acid, and then summed these values (Table 4.1).

Strikingly, the 9 highest scoring PrLDs all formed stress-induced assemblies, while the 11 lowest scoring all failed to assemble. To evaluate the accuracy of the predictor, we performed an iterative leave-one-out analysis for all of the PrLDs and plotted the resulting receiver operating characteristic (ROC) curve (Figure 4.10A). The area under the curve (AUC) of 0.86 for our predictor indicates a reasonably good predictive ability. By contrast, traditional prion prediction algorithms were not effective at predicting stress-induced assembly of PrLDs, with PAPA (23) yielding an AUC of 0.54, and PLAAC (24) yielding an AUC of 0.24. This further confirms that the compositional requirements for prion formation and those for reversible stress-induced assemblies are distinct.

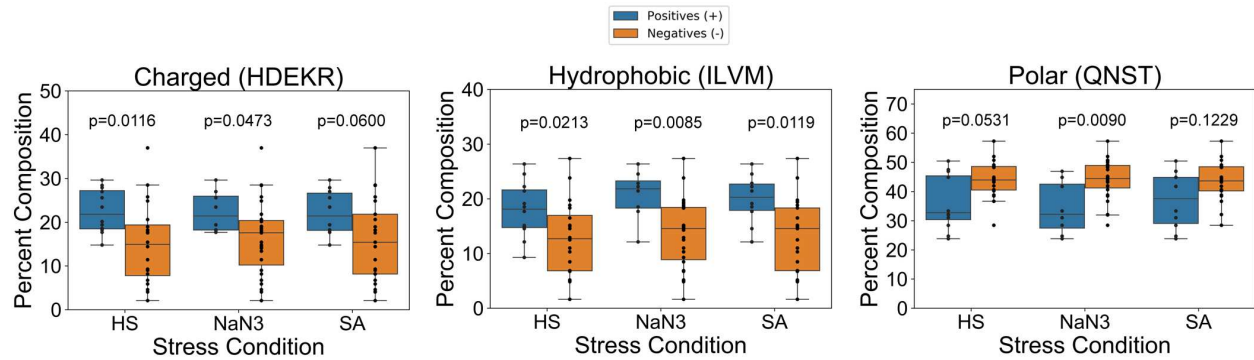


Figure 4.8: Compositional biases among assembling and non-assembling PrLDs. Box plots depicting compositional biases observed among assembly-forming (blue boxes) and non-assembly-forming (orange boxes) for each stress. HS=heat shock, NaN3=sodium azide stress, and SA=sorbic acid stress. The left-most plot depicts the percent composition of charged amino acids (H, D, E, K, and R) among the assembly and non-assembly-forming PrLDs, the middle plot depicts the percent composition of hydrophobic amino acids (I, L, V, and M) among the assembly and non-assembly-forming PrLDs, and the right-most plot depicts the percent composition of polar amino acids (Q, N, S, and T) among the assembly and non-assembly-forming PrLDs.

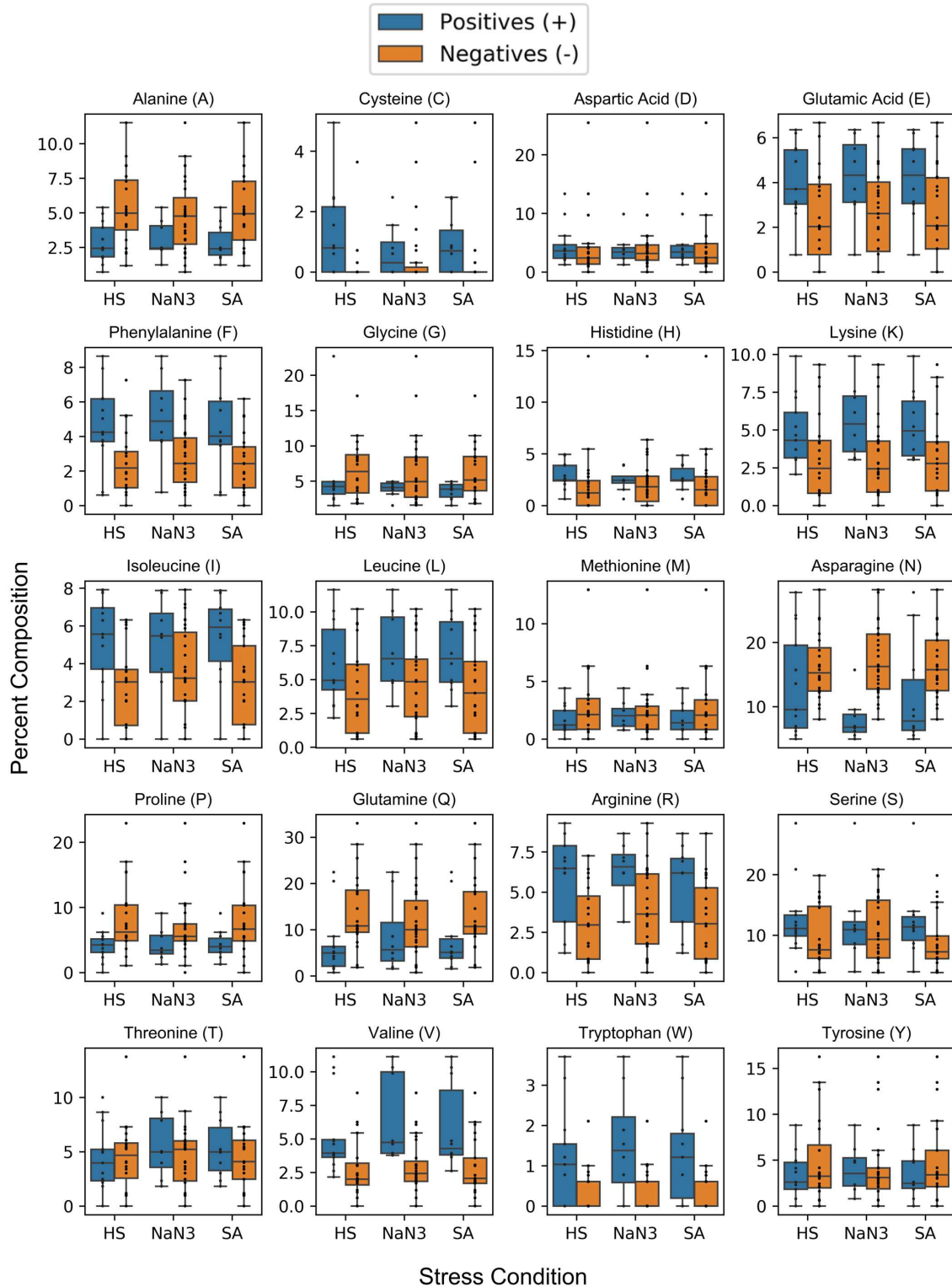


Figure 4.9: Compositional biases among assembling and non-assembling PrLDs. Box plots depicting the percent composition of each amino acid among all assembly-forming (blue boxes) and all non-assembly forming (orange boxes) PrLDs for each stress. HS=heat shock, NaN3=sodium azide stress, and SA=sorbic acid stress. P-values for each amino acid for each stress are listed in Table 4.2.

Table 4.2. P-values determined for each amino acid under each different stress.

Amino Acid	Heat Shock	NaN₃	Sorbic Acid
A	0.0071	0.1206	0.0134
C	0.0050	0.2622	0.0347
D	0.2238	0.7682	0.3746
E	0.0308	0.0536	0.0514
F	0.0089	0.0148	0.0279
G	0.1842	0.3162	0.0905
H	0.0169	0.3051	0.0890
I	0.0373	0.0148	0.0310
K	0.0083	0.1874	0.0149
L	0.0800	0.0677	0.0545
M	0.3377	0.7235	0.6877
N	0.1456	0.0002	0.0265
P	0.0089	0.0911	0.0384
Q	0.0043	0.2979	0.0295
R	0.0048	0.0307	0.0470
S	0.0937	0.7090	0.1130
T	0.7541	0.4917	0.5827
V	0.0030	0.0023	0.0056
W	0.0207	0.0037	0.0102
Y	0.6715	0.6512	0.7351

For each amino acid, the composition distribution for PrLDs that form stress-induced assemblies was compared (two-sided Mann-Whitney U test) to the composition distribution for PrLDs that do not form stress-induced assemblies. P-values from a two-sided Mann-Whitney U test are indicated for each amino acid under each different stress.

Table 4.3. Log odds ratio scores for each amino acid.

Amino Acid	Log Odds Ratio
A	-0.67
C	1.66
D	0.05
E	0.52
F	0.67
G	-0.25
H	0.33
I	0.66
K	0.47
L	0.39
M	-0.60
N	-0.23
P	-0.75
Q	-0.80
R	0.68
S	0.30
T	-0.03
V	0.69
W	1.34
Y	-0.41

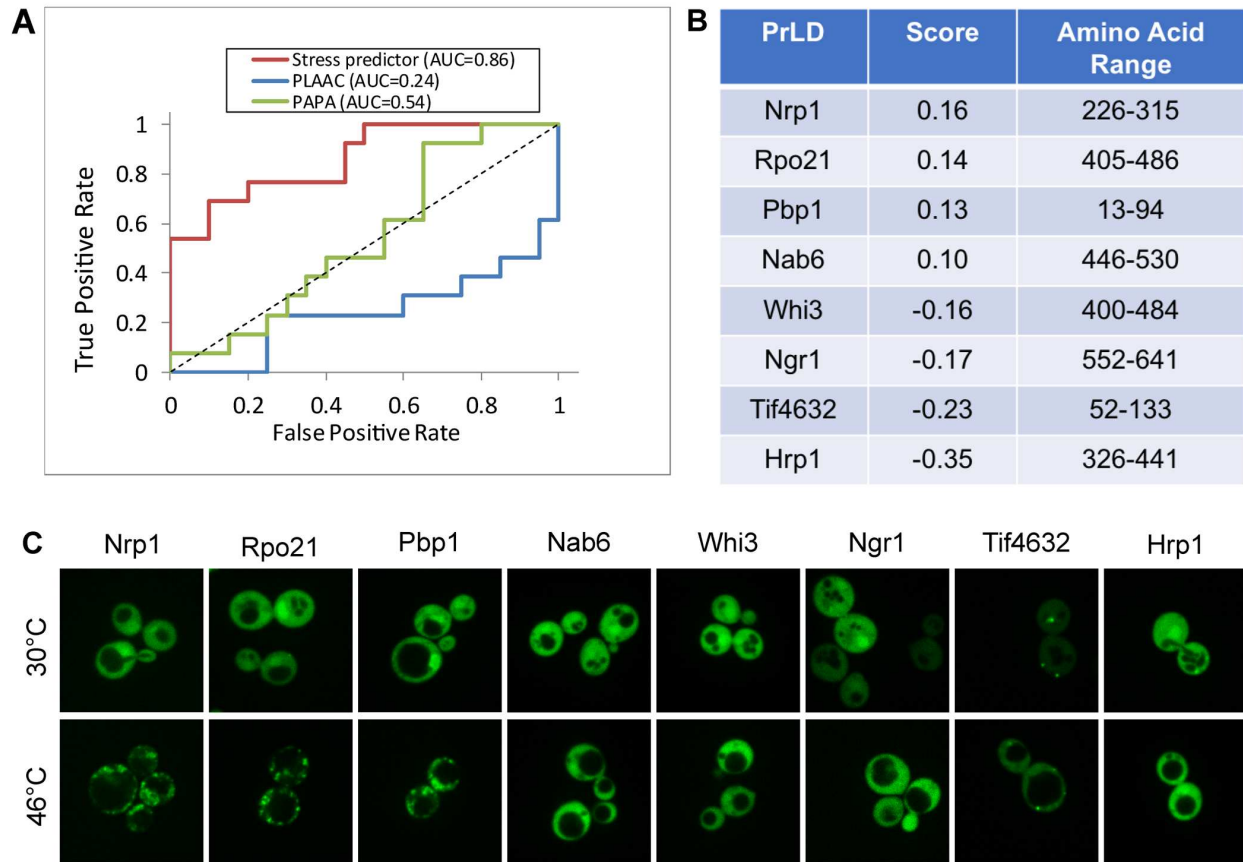


Figure 4.10: Composition is a predictor of PrLD assembly. (A) The ROC plot depicts the false positive rate versus the true positive rate for our stress-induced assembly predictor and for the prion prediction algorithms PAPA and PLAAC, along with AUC values (inset). PrLDs from the full PrLD dataset ($n=33$) were scored using an iterative leave-one-out procedure. (B) Assembly scores and domain boundaries are listed for each of the eight stress granule PrLDs that were tested. (C) The intracellular distribution of GFP-tagged PrLDs was assessed by fluorescence microscopy for cells grown at 30°C (*upper row*) and after 30 minutes of heat shock at 46°C (*bottom row*).

Many known stress granule proteins contain PrLDs that have not been tested for the ability to promote stress-induced assemblies. To test the fidelity of our predictor, we selected eight PrLDs with a range of predicted stress-induced assembly propensities, and evaluated these for assembly upon heat shock. Strikingly, the three highest-scoring PrLDs all reversibly formed stress-induced foci (Figure 4.10B, Figure 4.10C and Figure 4.11B), and these foci all colocalized with Pab1-mCh (Figure 4.11C). By contrast, four of the five lowest-scoring PrLDs failed to assemble (Figure 4.10B and Figure 4.10C). The sole exception was Tif4632, which robustly formed tiny foci under both normal and stress conditions (Figure 4.10C), indicating that it is not sufficient to undergo a stress-dependent relocalization, as predicted (Figure 4.10B). Together these results indicate that the strong compositional biases observed among the PrLDs that form stress-induced assemblies are sufficient to predict with reasonable accuracy whether a PrLD will assemble into stress-induced foci under heat stress.

Modulating Stress-Induced Assembly

The fact that we were able to predict stress-induced assembly of PrLDs based on composition suggests that we should similarly be able to modulate assembly properties by rationally changing a PrLD's amino acid composition. To test this, we selected three PrLDs that formed stress induced assemblies (Prt1, Trm1, and Rsc8) and three that did not (Mfg1, Pub1, and Sro9). These domains were selected for two reasons. First, they are all relatively short, and therefore should require fewer mutations to change their ability to assemble. Second, they all had composition scores within the ranges that our predictor was highly accurate (>10 for the positives, and <-5 for the negatives; Table 4.1, Figure 4.10A), suggesting that their compositions are sufficient to explain their assembly propensities.

None of the 10 PrLDs that had assembly scores below -0.14 formed stress-induced foci. Therefore, to prevent assembly, we randomly selected within Prt1, Trm1, and Rsc8 amino acids

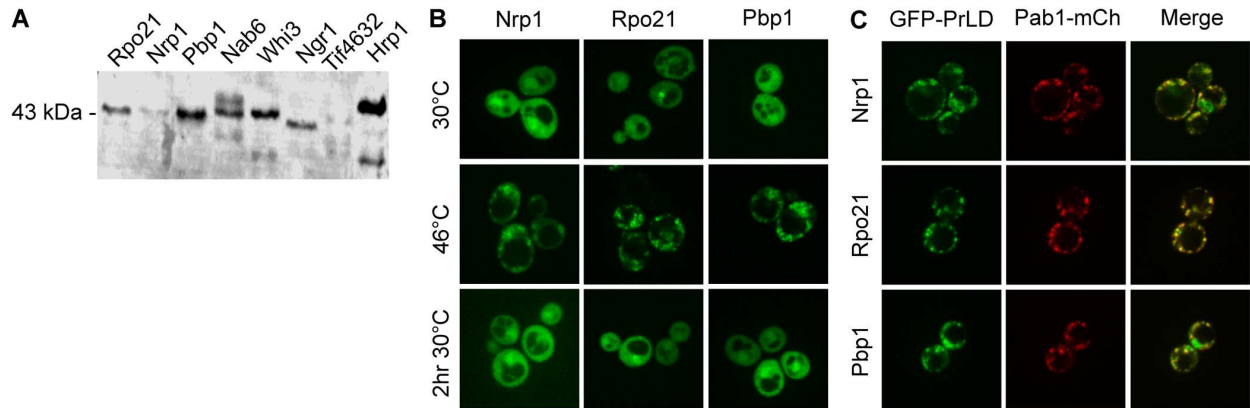


Figure 4.11: Foci formed by assembling stress granule PrLDs are reversible and also colocalize with stress granules. (A) Western blot of each of the eight PrLDs that were tested. (B) wt PrLDs from each protein were fused to the C-terminus of GFP and imaged under normal growth conditions, after 30 minutes of heat shock at 46°C, and after 2 hours of recovery at 30°C. (C) GFP-tagged PrLDs were co-expressed with Pab1-mCh as a stress granule marker to assess colocalization with stress granules. Cells were exposed to 30 minutes of heat shock at 46°C prior to imaging.

that were overrepresented among assembly-prone PrLDs (C, D, E, F, I, K, L, R, V, W), and serially replaced them with amino acids that were underrepresented among assembly-prone PrLDs (Q, N, A, and P) until the predicted assembly score was below -0.14 (Table 4.4). It should be noted that although C, D, E, F, I, K, L, R, V, and W were all more common among assembly-prone PrLDs, not all of these biases were statistically significant; however, we wanted to be conservative in targeting a range of amino acids that might contribute to assembly. Q, A, and P were chosen as replacements because each showed a reasonably strong underrepresentation among assembly-prone PrLDs. All three mutated PrLDs did not form stress-induced assemblies (Figure 4.12A).

To promote assembly, we similarly mutated Mfg1, Pub1, and Sro9, serially replacing randomly selected amino acids that were underrepresented among assembly-prone PrLDs (Q, N, A, P, M, and T) with amino acids that were overrepresented (E, V, F, I, R, K, L, and D) (Table 4.4). The mutated version of Sro9 did not express, and so could not be evaluated for stress-induced assembly (Figure 4.13A). However, while the mutated versions of both Mfg1 and Pub1 formed small foci in some cells prior to stress, both showed a substantial increase in foci formation upon heat stress (Figure 4.12B), and these foci showed partial colocalization with Pab-mCherry (Figure 4.13B). Thus, for five of the six PrLDs we were able to successfully reverse the stress-induced assembly or non-assembly phenotype through rational mutation, indicating that we can rationally modulate assembly properties. Additionally, the same sequence features that promote assembly also appear to promote recruitment of these assemblies to *bona fide* stress granules.

Finally, previous work has shown that scrambling the primary sequence of yeast prion domains does not prevent prion formation (28, 36). Because we were able to predict and alter assembly properties based on amino acid composition, we tested whether stress-induced

Table S4. Sequences of PrLD mutant and scrambled constructs.

PrLD Mutant (Assembly Score)	Sequence
Prt1 + to - (-0.17)	<u>QM</u> <u>NDA</u> <u>QP</u> <u>Q</u> <u>Q</u> <u>F</u> <u>PP</u> <u>Q</u> <u>D</u> <u>Q</u> <u>P</u> <u>Q</u> <u>T</u> <u>S</u> <u>V</u> <u>A</u> <u>Q</u> <u>N</u> <u>S</u> <u>M</u> <u>F</u> <u>N</u> <u>E</u> <u>E</u> <u>D</u> <u>S</u> <u>A</u> <u>N</u> <u>Q</u> <u>S</u> <u>R</u> <u>E</u> <u>N</u> <u>Q</u> <u>S</u> <u>T</u> <u>N</u> <u>Y</u> <u>Q</u> <u>R</u> <u>F</u> <u>S</u> <u>P</u> <u>K</u> <u>G</u> <u>T</u> <u>Y</u> <u>L</u> <u>N</u> <u>S</u> <u>Y</u> <u>H</u> <u>Q</u> <u>Q</u> <u>G</u> <u>N</u> <u>T</u> <u>A</u> <u>N</u> <u>G</u> <u>G</u> <u>P</u> <u>N</u> <u>F</u> <u>N</u> <u>N</u> <u>L</u> <u>R</u> <u>N</u> <u>A</u> <u>Y</u> <u>H</u> <u>P</u> <u>D</u> <u>V</u> <u>N</u>
Trm1 + to - (-0.16)	<u>A</u> <u>N</u> <u>Y</u> <u>N</u> <u>K</u> <u>T</u> <u>V</u> <u>E</u> <u>P</u> <u>L</u> <u>L</u> <u>S</u> <u>P</u> <u>S</u> <u>N</u> <u>A</u> <u>F</u> <u>Y</u> <u>V</u> <u>R</u> <u>V</u> <u>Q</u> <u>P</u> <u>O</u> <u>Q</u> <u>O</u> <u>N</u> <u>T</u> <u>S</u> <u>P</u> <u>P</u> <u>N</u> <u>O</u> <u>Q</u> <u>N</u> <u>V</u> <u>M</u> <u>S</u> <u>S</u> <u>T</u> <u>M</u> <u>T</u> <u>Y</u> <u>H</u> <u>P</u> <u>S</u> <u>R</u> <u>C</u> <u>G</u> <u>S</u> <u>Y</u> <u>H</u> <u>N</u> <u>Q</u> <u>P</u> <u>Q</u> <u>G</u> <u>R</u> <u>N</u> <u>S</u> <u>Q</u> <u>A</u> <u>A</u> <u>G</u> <u>O</u> <u>N</u> <u>N</u> <u>N</u> <u>T</u> <u>O</u> <u>T</u> <u>A</u> <u>Y</u> <u>S</u> <u>V</u> <u>A</u> <u>Q</u> <u>G</u> <u>P</u> <u>P</u> <u>V</u>
Rsc8 + to - (-0.15)	<u>K</u> <u>N</u> <u>A</u> <u>Y</u> <u>D</u> <u>S</u> <u>A</u> <u>Q</u> <u>D</u> <u>F</u> <u>N</u> <u>A</u> <u>L</u> <u>Q</u> <u>O</u> <u>Q</u> <u>S</u> <u>R</u> <u>N</u> <u>S</u> <u>N</u> <u>Q</u> <u>I</u> <u>H</u> <u>K</u> <u>V</u> <u>Y</u> <u>N</u> <u>O</u> <u>H</u> <u>T</u> <u>P</u> <u>G</u> <u>N</u> <u>N</u> <u>S</u> <u>P</u> <u>N</u> <u>V</u> <u>N</u> <u>Y</u> <u>H</u> <u>N</u> <u>L</u> <u>R</u> <u>A</u> <u>P</u> <u>Q</u> <u>T</u> <u>N</u> <u>N</u> <u>P</u> <u>S</u> <u>R</u> <u>Q</u> <u>Q</u> <u>P</u> <u>G</u> <u>H</u> <u>F</u> <u>G</u> <u>A</u> <u>N</u> <u>Q</u> <u>Q</u> <u>S</u> <u>S</u> <u>D</u> <u>F</u> <u>A</u> <u>N</u> <u>N</u> <u>A</u> <u>N</u> <u>N</u> <u>G</u> <u>N</u> <u>S</u> <u>V</u> <u>N</u>
Mfg1 - to + (0.16)	<u>M</u> <u>Y</u> <u>F</u> <u>G</u> <u>P</u> <u>I</u> <u>F</u> <u>R</u> <u>P</u> <u>K</u> <u>Q</u> <u>E</u> <u>V</u> <u>L</u> <u>R</u> <u>I</u> <u>Y</u> <u>I</u> <u>V</u> <u>E</u> <u>N</u> <u>L</u> <u>T</u> <u>P</u> <u>R</u> <u>Y</u> <u>P</u> <u>N</u> <u>G</u> <u>V</u> <u>I</u> <u>N</u> <u>F</u> <u>D</u> <u>P</u> <u>R</u> <u>E</u> <u>Q</u> <u>K</u> <u>V</u> <u>I</u> <u>P</u> <u>T</u> <u>V</u> <u>Y</u> <u>D</u> <u>D</u> <u>L</u> <u>V</u> <u>P</u> <u>F</u> <u>D</u> <u>G</u> <u>Q</u> <u>P</u> <u>Q</u> <u>G</u> <u>G</u> <u>Q</u> <u>F</u> <u>K</u> <u>K</u> <u>F</u> <u>S</u> <u>S</u> <u>E</u> <u>V</u> <u>V</u> <u>V</u> <u>F</u> <u>N</u> <u>F</u> <u>L</u> <u>D</u> <u>L</u> <u>V</u> <u>V</u> <u>E</u> <u>T</u> <u>F</u> <u>H</u> <u>N</u> <u>S</u> <u>Q</u> <u>I</u> <u>E</u> <u>L</u> <u>R</u> <u>K</u> <u>G</u> <u>G</u> <u>F</u> <u>G</u> <u>S</u>
Pub1 - to + (0.16)	<u>F</u> <u>E</u> <u>V</u> <u>L</u> <u>R</u> <u>K</u> <u>Y</u> <u>V</u> <u>I</u> <u>R</u> <u>R</u> <u>E</u> <u>Y</u> <u>G</u> <u>I</u> <u>N</u> <u>L</u> <u>R</u> <u>G</u> <u>G</u> <u>F</u> <u>R</u> <u>O</u> <u>Y</u> <u>E</u> <u>S</u> <u>N</u> <u>R</u> <u>N</u> <u>I</u> <u>N</u> <u>R</u> <u>N</u> <u>V</u> <u>G</u> <u>K</u> <u>D</u> <u>M</u> <u>N</u> <u>K</u> <u>F</u> <u>M</u> <u>N</u> <u>M</u> <u>R</u> <u>N</u> <u>S</u> <u>R</u> <u>G</u> <u>F</u> <u>I</u> <u>P</u> <u>S</u> <u>S</u> <u>M</u> <u>G</u> <u>M</u> <u>F</u> <u>I</u> <u>G</u> <u>A</u> <u>V</u> <u>K</u> <u>L</u> <u>D</u> <u>S</u> <u>Q</u> <u>G</u> <u>Q</u> <u>R</u> <u>O</u> <u>D</u> <u>S</u> <u>Q</u> <u>T</u> <u>I</u> <u>G</u> <u>L</u> <u>E</u> <u>E</u> <u>L</u> <u>V</u> <u>D</u> <u>F</u>
Sro9 - to + (0.22)	<u>S</u> <u>V</u> <u>E</u> <u>E</u> <u>G</u> <u>R</u> <u>S</u> <u>K</u> <u>Q</u> <u>E</u> <u>Q</u> <u>E</u> <u>P</u> <u>P</u> <u>H</u> <u>R</u> <u>N</u> <u>H</u> <u>H</u> <u>H</u> <u>S</u> <u>H</u> <u>H</u> <u>N</u> <u>S</u> <u>L</u> <u>G</u> <u>F</u> <u>F</u> <u>R</u> <u>R</u> <u>K</u> <u>F</u> <u>H</u> <u>E</u> <u>S</u> <u>N</u> <u>V</u> <u>A</u> <u>G</u> <u>E</u> <u>F</u> <u>K</u> <u>V</u> <u>Q</u> <u>G</u> <u>F</u> <u>L</u> <u>P</u> <u>L</u> <u>F</u> <u>K</u> <u>P</u> <u>Y</u> <u>Q</u> <u>G</u> <u>R</u> <u>I</u> <u>A</u> <u>R</u> <u>N</u> <u>N</u> <u>R</u> <u>D</u> <u>N</u> <u>R</u> <u>S</u> <u>K</u> <u>Y</u> <u>H</u> <u>N</u> <u>H</u> <u>F</u> <u>H</u> <u>F</u> <u>F</u> <u>I</u> <u>L</u> <u>H</u> <u>P</u> <u>Q</u> <u>D</u> <u>L</u> <u>F</u> <u>V</u> <u>K</u> <u>L</u> <u>I</u> <u>R</u> <u>D</u> <u>F</u> <u>Y</u> <u>V</u> <u>V</u> <u>R</u>
Prt1 Scramble_1	<u>G</u> <u>R</u> <u>L</u> <u>P</u> <u>N</u> <u>V</u> <u>Q</u> <u>F</u> <u>S</u> <u>K</u> <u>S</u> <u>S</u> <u>D</u> <u>F</u> <u>D</u> <u>K</u> <u>R</u> <u>E</u> <u>G</u> <u>S</u> <u>D</u> <u>P</u> <u>G</u> <u>V</u> <u>L</u> <u>W</u> <u>V</u> <u>N</u> <u>F</u> <u>H</u> <u>G</u> <u>Y</u> <u>L</u> <u>P</u> <u>V</u> <u>T</u> <u>D</u> <u>L</u> <u>N</u> <u>R</u> <u>V</u> <u>D</u> <u>T</u> <u>K</u> <u>V</u> <u>L</u> <u>D</u> <u>A</u> <u>V</u> <u>T</u> <u>M</u> <u>S</u> <u>W</u> <u>S</u> <u>H</u> <u>D</u> <u>N</u> <u>Q</u> <u>F</u> <u>S</u> <u>Y</u> <u>E</u> <u>Y</u> <u>F</u> <u>R</u> <u>E</u> <u>R</u> <u>Y</u> <u>N</u> <u>E</u> <u>W</u> <u>F</u> <u>Q</u> <u>T</u> <u>M</u> <u>V</u> <u>F</u> <u>Q</u> <u>R</u> <u>D</u> <u>R</u>
Prt1 Scramble_2	<u>R</u> <u>S</u> <u>E</u> <u>E</u> <u>D</u> <u>K</u> <u>G</u> <u>W</u> <u>T</u> <u>T</u> <u>V</u> <u>T</u> <u>R</u> <u>S</u> <u>V</u> <u>Q</u> <u>L</u> <u>F</u> <u>K</u> <u>N</u> <u>K</u> <u>S</u> <u>Y</u> <u>R</u> <u>H</u> <u>Q</u> <u>N</u> <u>E</u> <u>V</u> <u>V</u> <u>Y</u> <u>F</u> <u>M</u> <u>W</u> <u>S</u> <u>S</u> <u>Y</u> <u>L</u> <u>S</u> <u>P</u> <u>F</u> <u>R</u> <u>L</u> <u>R</u> <u>F</u> <u>R</u> <u>W</u> <u>D</u> <u>G</u> <u>Q</u> <u>F</u> <u>N</u> <u>T</u> <u>F</u> <u>G</u> <u>V</u> <u>L</u> <u>N</u> <u>Q</u> <u>P</u> <u>D</u> <u>V</u> <u>F</u> <u>N</u> <u>H</u> <u>G</u> <u>P</u> <u>L</u> <u>D</u> <u>S</u> <u>R</u> <u>V</u> <u>D</u> <u>D</u> <u>V</u> <u>A</u> <u>D</u> <u>Y</u> <u>E</u> <u>D</u> <u>M</u>
Prt1 Scramble_3	<u>Q</u> <u>D</u> <u>V</u> <u>R</u> <u>D</u> <u>F</u> <u>M</u> <u>K</u> <u>R</u> <u>V</u> <u>S</u> <u>L</u> <u>V</u> <u>P</u> <u>H</u> <u>R</u> <u>G</u> <u>T</u> <u>E</u> <u>L</u> <u>L</u> <u>E</u> <u>K</u> <u>N</u> <u>T</u> <u>S</u> <u>S</u> <u>W</u> <u>S</u> <u>V</u> <u>V</u> <u>N</u> <u>G</u> <u>Y</u> <u>P</u> <u>G</u> <u>N</u> <u>D</u> <u>H</u> <u>E</u> <u>Q</u> <u>Y</u> <u>P</u> <u>M</u> <u>L</u> <u>K</u> <u>S</u> <u>W</u> <u>V</u> <u>D</u> <u>Q</u> <u>R</u> <u>F</u> <u>A</u> <u>Y</u> <u>F</u> <u>F</u> <u>Q</u> <u>N</u> <u>D</u> <u>V</u> <u>S</u> <u>F</u> <u>V</u> <u>T</u> <u>R</u> <u>D</u> <u>W</u> <u>T</u> <u>N</u> <u>R</u> <u>R</u> <u>D</u> <u>S</u> <u>L</u> <u>G</u> <u>F</u> <u>E</u> <u>D</u> <u>F</u> <u>Y</u>
Prt1 Scramble_4	<u>S</u> <u>N</u> <u>V</u> <u>D</u> <u>D</u> <u>E</u> <u>V</u> <u>F</u> <u>R</u> <u>G</u> <u>T</u> <u>W</u> <u>G</u> <u>L</u> <u>Y</u> <u>E</u> <u>D</u> <u>V</u> <u>Q</u> <u>D</u> <u>Y</u> <u>L</u> <u>F</u> <u>R</u> <u>N</u> <u>R</u> <u>T</u> <u>M</u> <u>R</u> <u>E</u> <u>D</u> <u>V</u> <u>Y</u> <u>V</u> <u>P</u> <u>N</u> <u>S</u> <u>F</u> <u>E</u> <u>S</u> <u>S</u> <u>F</u> <u>V</u> <u>D</u> <u>S</u> <u>L</u> <u>K</u> <u>G</u> <u>T</u> <u>H</u> <u>V</u> <u>L</u> <u>W</u> <u>R</u> <u>S</u> <u>D</u> <u>F</u> <u>Y</u> <u>Q</u> <u>N</u> <u>W</u> <u>F</u> <u>H</u> <u>G</u> <u>D</u> <u>K</u> <u>T</u> <u>A</u> <u>F</u> <u>S</u> <u>V</u> <u>L</u> <u>Q</u> <u>M</u> <u>R</u> <u>R</u> <u>N</u> <u>K</u> <u>P</u> <u>P</u>
Trm1 Scramble_1	<u>S</u> <u>R</u> <u>P</u> <u>K</u> <u>N</u> <u>S</u> <u>F</u> <u>V</u> <u>K</u> <u>Y</u> <u>A</u> <u>T</u> <u>T</u> <u>V</u> <u>G</u> <u>D</u> <u>S</u> <u>S</u> <u>C</u> <u>S</u> <u>K</u> <u>N</u> <u>V</u> <u>V</u> <u>I</u> <u>G</u> <u>V</u> <u>Q</u> <u>C</u> <u>R</u> <u>Y</u> <u>A</u> <u>T</u> <u>F</u> <u>Q</u> <u>G</u> <u>K</u> <u>T</u> <u>V</u> <u>N</u> <u>F</u> <u>E</u> <u>I</u> <u>V</u> <u>Q</u> <u>S</u> <u>Y</u> <u>H</u> <u>P</u> <u>K</u> <u>R</u> <u>N</u> <u>P</u> <u>V</u> <u>H</u> <u>R</u> <u>S</u> <u>K</u> <u>S</u> <u>P</u> <u>L</u> <u>E</u> <u>Y</u> <u>S</u> <u>K</u> <u>K</u> <u>M</u> <u>L</u> <u>V</u> <u>M</u> <u>R</u> <u>T</u> <u>L</u> <u>E</u> <u>Y</u> <u>P</u> <u>G</u> <u>T</u> <u>I</u> <u>L</u> <u>T</u>
Trm1 Scramble_2	<u>S</u> <u>K</u> <u>G</u> <u>K</u> <u>V</u> <u>S</u> <u>R</u> <u>S</u> <u>Y</u> <u>E</u> <u>L</u> <u>N</u> <u>A</u> <u>M</u> <u>T</u> <u>V</u> <u>H</u> <u>C</u> <u>N</u> <u>S</u> <u>Y</u> <u>R</u> <u>T</u> <u>S</u> <u>I</u> <u>T</u> <u>Q</u> <u>T</u> <u>R</u> <u>E</u> <u>L</u> <u>L</u> <u>M</u> <u>K</u> <u>A</u> <u>V</u> <u>V</u> <u>K</u> <u>S</u> <u>P</u> <u>N</u> <u>F</u> <u>Q</u> <u>S</u> <u>V</u> <u>V</u> <u>V</u> <u>F</u> <u>G</u> <u>Q</u> <u>P</u> <u>Y</u> <u>K</u> <u>P</u> <u>K</u> <u>V</u> <u>L</u> <u>R</u> <u>G</u> <u>T</u> <u>H</u> <u>T</u> <u>S</u> <u>D</u> <u>K</u> <u>I</u> <u>I</u> <u>P</u> <u>T</u> <u>Y</u> <u>E</u> <u>V</u> <u>C</u> <u>N</u> <u>K</u> <u>R</u> <u>G</u> <u>Y</u> <u>S</u> <u>P</u>
Trm1 Scramble_3	<u>C</u> <u>R</u> <u>Y</u> <u>L</u> <u>P</u> <u>K</u> <u>T</u> <u>R</u> <u>I</u> <u>K</u> <u>R</u> <u>E</u> <u>K</u> <u>F</u> <u>M</u> <u>Y</u> <u>S</u> <u>V</u> <u>V</u> <u>N</u> <u>P</u> <u>V</u> <u>N</u> <u>Y</u> <u>V</u> <u>S</u> <u>T</u> <u>K</u> <u>Y</u> <u>N</u> <u>M</u> <u>R</u> <u>G</u> <u>K</u> <u>V</u> <u>K</u> <u>Q</u> <u>H</u> <u>T</u> <u>S</u> <u>A</u> <u>Y</u> <u>K</u> <u>S</u> <u>E</u> <u>Q</u> <u>G</u> <u>S</u> <u>T</u> <u>Q</u> <u>E</u> <u>S</u> <u>V</u> <u>C</u> <u>R</u> <u>V</u> <u>K</u> <u>T</u> <u>N</u> <u>S</u> <u>D</u> <u>I</u> <u>L</u> <u>G</u> <u>S</u> <u>T</u> <u>L</u> <u>G</u> <u>S</u> <u>H</u> <u>A</u> <u>V</u> <u>T</u> <u>P</u> <u>I</u> <u>F</u> <u>P</u> <u>L</u> <u>P</u> <u>F</u>
Trm1 Scramble_4	<u>R</u> <u>H</u> <u>L</u> <u>S</u> <u>E</u> <u>K</u> <u>C</u> <u>V</u> <u>S</u> <u>N</u> <u>P</u> <u>V</u> <u>S</u> <u>C</u> <u>K</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>Q</u> <u>G</u> <u>Y</u> <u>K</u> <u>I</u> <u>S</u> <u>T</u> <u>V</u> <u>V</u> <u>L</u> <u>S</u> <u>Q</u> <u>S</u> <u>V</u> <u>I</u> <u>G</u> <u>V</u> <u>Y</u> <u>Y</u> <u>R</u> <u>P</u> <u>I</u> <u>M</u> <u>T</u> <u>T</u> <u>N</u> <u>V</u> <u>E</u> <u>Y</u> <u>E</u> <u>P</u> <u>L</u> <u>D</u> <u>K</u> <u>R</u> <u>R</u> <u>A</u> <u>Y</u> <u>K</u> <u>N</u> <u>H</u> <u>P</u> <u>V</u> <u>G</u> <u>R</u> <u>M</u> <u>Q</u> <u>F</u> <u>T</u> <u>V</u> <u>K</u> <u>S</u> <u>K</u> <u>T</u> <u>K</u> <u>S</u> <u>F</u> <u>F</u> <u>N</u> <u>P</u> <u>L</u> <u>S</u>
Rsc8 Scramble_1	<u>H</u> <u>R</u> <u>C</u> <u>R</u> <u>G</u> <u>L</u> <u>S</u> <u>E</u> <u>Q</u> <u>I</u> <u>S</u> <u>R</u> <u>G</u> <u>V</u> <u>V</u> <u>I</u> <u>S</u> <u>F</u> <u>V</u> <u>C</u> <u>D</u> <u>S</u> <u>L</u> <u>R</u> <u>N</u> <u>N</u> <u>N</u> <u>H</u> <u>F</u> <u>N</u> <u>T</u> <u>Q</u> <u>F</u> <u>H</u> <u>S</u> <u>N</u> <u>H</u> <u>S</u> <u>G</u> <u>L</u> <u>Q</u> <u>N</u> <u>Y</u> <u>I</u> <u>D</u> <u>A</u> <u>A</u> <u>E</u> <u>V</u> <u>D</u> <u>E</u> <u>N</u> <u>S</u> <u>R</u> <u>N</u> <u>D</u> <u>N</u> <u>N</u> <u>F</u> <u>I</u> <u>D</u> <u>N</u> <u>Q</u> <u>K</u> <u>K</u> <u>R</u> <u>S</u> <u>Y</u> <u>G</u> <u>Q</u> <u>R</u> <u>F</u> <u>K</u> <u>C</u> <u>A</u> <u>E</u> <u>L</u> <u>T</u> <u>Y</u> <u>C</u> <u>A</u>
Rsc8 Scramble_2	<u>H</u> <u>S</u> <u>F</u> <u>A</u> <u>E</u> <u>H</u> <u>T</u> <u>L</u> <u>G</u> <u>T</u> <u>R</u> <u>Y</u> <u>L</u> <u>N</u> <u>D</u> <u>Q</u> <u>H</u> <u>G</u> <u>G</u> <u>N</u> <u>L</u> <u>H</u> <u>Q</u> <u>F</u> <u>N</u> <u>A</u> <u>A</u> <u>R</u> <u>E</u> <u>S</u> <u>I</u> <u>S</u> <u>I</u> <u>C</u> <u>E</u> <u>N</u> <u>R</u> <u>K</u> <u>Q</u> <u>L</u> <u>A</u> <u>S</u> <u>R</u> <u>N</u> <u>Q</u> <u>D</u> <u>D</u> <u>V</u> <u>R</u> <u>S</u> <u>N</u> <u>N</u> <u>Y</u> <u>G</u> <u>Y</u> <u>C</u> <u>C</u> <u>I</u> <u>S</u> <u>N</u> <u>I</u> <u>S</u> <u>S</u> <u>V</u> <u>V</u> <u>N</u> <u>V</u> <u>F</u> <u>D</u> <u>R</u> <u>K</u> <u>N</u> <u>N</u> <u>E</u> <u>K</u> <u>R</u> <u>D</u> <u>F</u> <u>C</u> <u>F</u> <u>Q</u>
Rsc8 Scramble_3	<u>E</u> <u>G</u> <u>C</u> <u>H</u> <u>R</u> <u>N</u> <u>I</u> <u>E</u> <u>N</u> <u>Q</u> <u>A</u> <u>R</u> <u>Q</u> <u>E</u> <u>V</u> <u>F</u> <u>R</u> <u>S</u> <u>S</u> <u>I</u> <u>L</u> <u>A</u> <u>K</u> <u>D</u> <u>F</u> <u>Q</u> <u>N</u> <u>I</u> <u>R</u> <u>S</u> <u>Y</u> <u>N</u> <u>F</u> <u>R</u> <u>K</u> <u>C</u> <u>S</u> <u>C</u> <u>L</u> <u>S</u> <u>G</u> <u>N</u> <u>N</u> <u>D</u> <u>S</u> <u>S</u> <u>H</u> <u>H</u> <u>Q</u> <u>T</u> <u>D</u> <u>Y</u> <u>D</u> <u>I</u> <u>N</u> <u>L</u> <u>K</u> <u>F</u> <u>V</u> <u>V</u> <u>Q</u> <u>L</u> <u>E</u> <u>C</u> <u>N</u> <u>G</u> <u>T</u> <u>G</u> <u>R</u> <u>A</u> <u>D</u> <u>N</u> <u>R</u> <u>N</u> <u>F</u> <u>V</u> <u>A</u> <u>H</u> <u>S</u> <u>Y</u>
Rsc8 Scramble_4	<u>S</u> <u>F</u> <u>D</u> <u>C</u> <u>L</u> <u>A</u> <u>R</u> <u>N</u> <u>R</u> <u>T</u> <u>S</u> <u>I</u> <u>F</u> <u>N</u> <u>Q</u> <u>H</u> <u>S</u> <u>S</u> <u>A</u> <u>S</u> <u>Y</u> <u>R</u> <u>Q</u> <u>V</u> <u>K</u> <u>V</u> <u>F</u> <u>C</u> <u>L</u> <u>D</u> <u>N</u> <u>G</u> <u>V</u> <u>A</u> <u>E</u> <u>L</u> <u>G</u> <u>D</u> <u>H</u> <u>R</u> <u>Y</u> <u>N</u> <u>D</u> <u>S</u> <u>C</u> <u>E</u> <u>N</u> <u>H</u> <u>V</u> <u>R</u> <u>K</u> <u>I</u> <u>G</u> <u>S</u> <u>N</u> <u>F</u> <u>L</u> <u>I</u> <u>T</u> <u>E</u> <u>F</u> <u>D</u> <u>A</u> <u>R</u> <u>G</u> <u>N</u> <u>N</u> <u>C</u> <u>Q</u> <u>N</u> <u>K</u> <u>Y</u> <u>S</u> <u>E</u> <u>R</u> <u>N</u> <u>Q</u> <u>Q</u> <u>H</u> <u>I</u> <u>N</u>
Mfg1 Scramble	<u>N</u> <u>F</u> <u>Q</u> <u>Y</u> <u>A</u> <u>T</u> <u>N</u> <u>Q</u> <u>G</u> <u>Y</u> <u>P</u> <u>G</u> <u>V</u> <u>Q</u> <u>G</u> <u>Q</u> <u>Q</u> <u>P</u> <u>P</u> <u>A</u> <u>V</u> <u>P</u> <u>N</u> <u>P</u> <u>P</u> <u>P</u> <u>G</u> <u>P</u> <u>N</u> <u>G</u> <u>Y</u> <u>P</u> <u>P</u> <u>S</u> <u>T</u> <u>Q</u> <u>P</u> <u>N</u> <u>Q</u> <u>V</u> <u>V</u> <u>F</u> <u>T</u> <u>A</u> <u>E</u> <u>N</u> <u>S</u> <u>I</u> <u>Q</u> <u>P</u> <u>I</u> <u>I</u> <u>Q</u> <u>Q</u> <u>G</u> <u>N</u> <u>Q</u> <u>V</u> <u>T</u> <u>N</u> <u>Q</u> <u>F</u> <u>T</u> <u>P</u> <u>N</u> <u>N</u> <u>M</u> <u>H</u> <u>A</u> <u>V</u> <u>Q</u> <u>P</u> <u>P</u> <u>P</u> <u>Q</u> <u>P</u> <u>Y</u> <u>P</u> <u>G</u> <u>Q</u> <u>P</u> <u>P</u> <u>P</u> <u>N</u> <u>S</u> <u>M</u> <u>S</u> <u>T</u> <u>F</u> <u>F</u> <u>Q</u> <u>T</u>

Pub1 Scramble	YNPSGMQMMQNLNMQGNNNGMGRNPYTSNIGPQNLNRNS MFSNPMGIPSRQNQNNNPRNGGGPPAMNQNQRPNMVM MYN QNSNQ
Sro9 Scramble	QTVQQNHQNPQRPRFNNSQPHPHREPPNLKGNQFHNFGSS VRRQQHQKPHSMKQYYHSHAGGQHGN YFNANRHQNHPQP NQNFKNPSHHMNQQQ

Amino acids bolded and underlined indicate residues that were mutated.

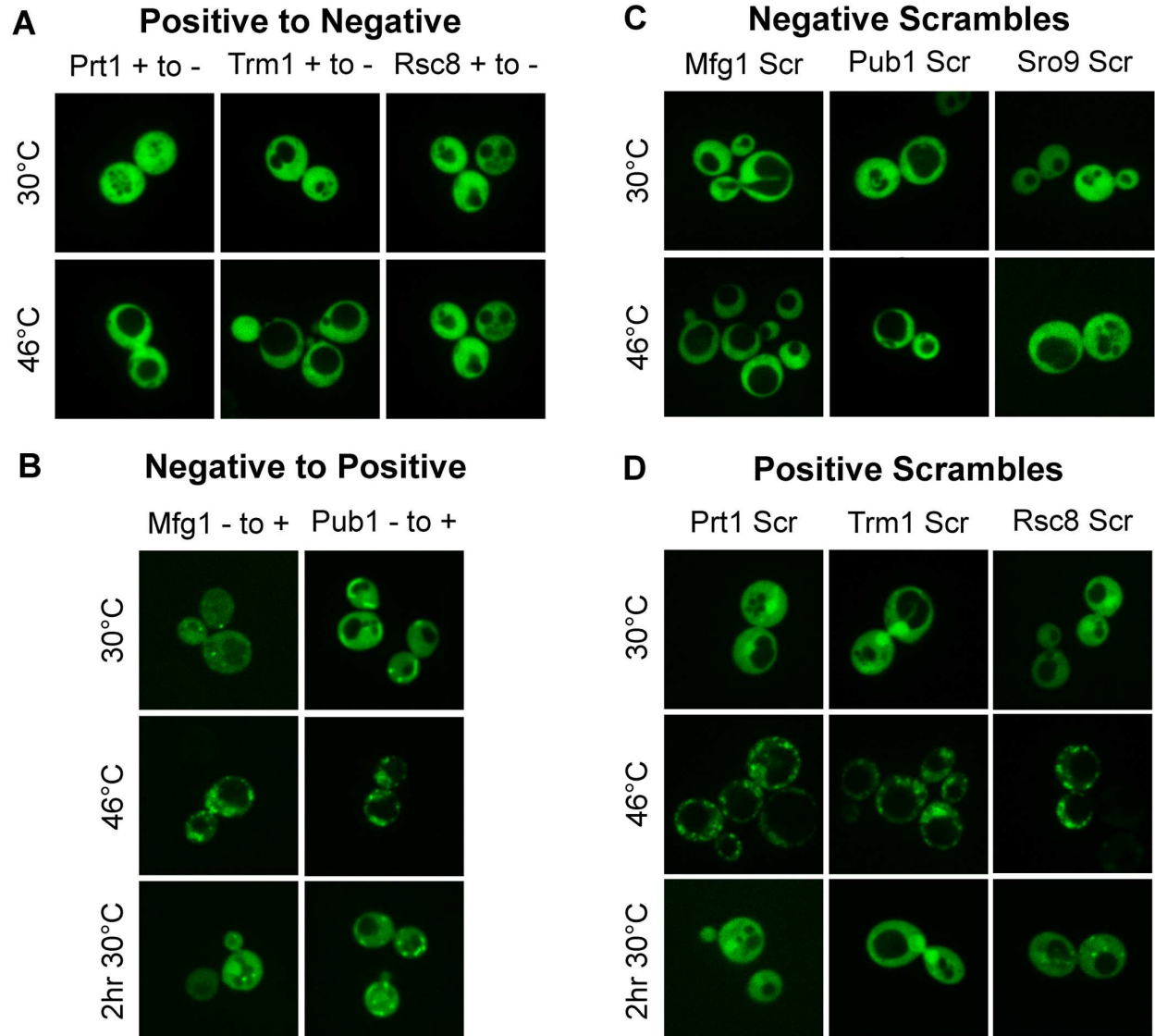


Figure 4.12: Effects of rational mutation and scrambling on PrLD behavior. PrLDs were rationally mutated to either (A) inhibit stress-induced assembly or (B) promote stress-induced assembly. Wt PrLDs from non-assembling (C) and assembling (D) constructs were scrambled to determine the effects on localization. Mutated and scrambled PrLDs were fused to the C-terminus of GFP and imaged under normal growth conditions and after 30 minutes of heat shock at 46°C. PrLDs that formed foci upon heat shock were also imaged after 2 hours of recovery post-heat shock at 30°C.

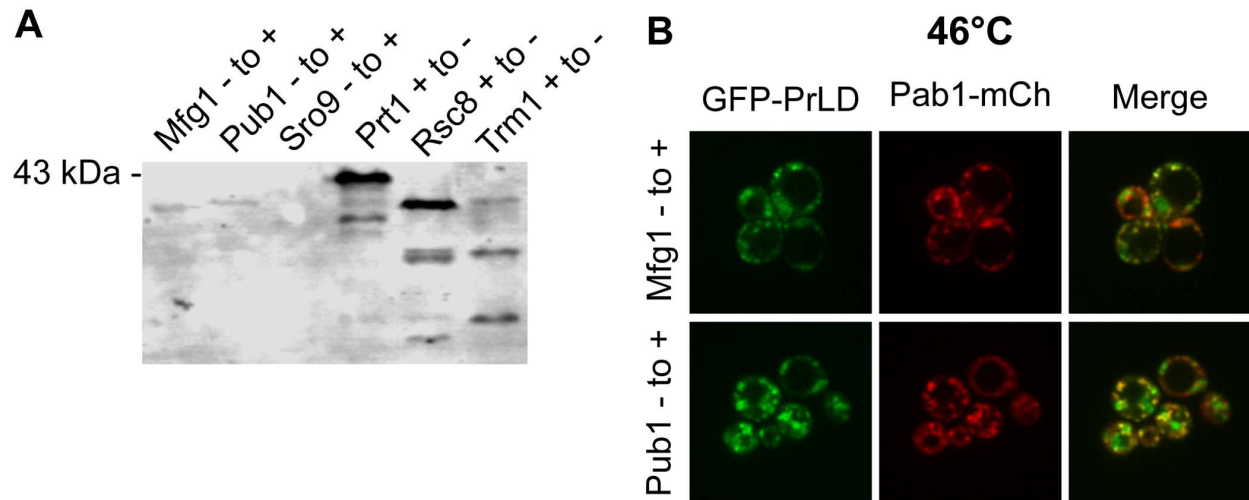


Figure 4.13: PrLD mutation does not affect recruitment to stress granules. (A) Western blot of each of the six mutated PrLDs that were tested. (B) GFP-tagged mutated PrLDs were co-expressed with Pab1-mCh as a stress granule marker to assess colocalization with stress granules. Cells were exposed to 30 minutes of heat shock at 46°C prior to imaging.

assembly showed similar insensitivity to scrambling. We scrambled the same six PrLDs that were rationally mutated in Figure 4.12A-B. All three non-assembling PrLDs retained this phenotype upon scrambling (Figure 4.12C,D). Initial scrambled versions of the three assembly-prone PrLDs were not detectable by western blot (Figure 4.14A), so multiple scrambled versions were constructed and tested (Table 4.4). Strikingly, each additional scrambled version formed foci upon heat shock (Figure 4.14B), and these foci co-localized with Pab1-mCherry (Figure 4.15). These results suggest that amino acid composition is the predominant determinant of a PrLD's recruitment to stress granules.

Discussion

PrLDs are relatively common among RNA-binding proteins, especially those that are recruited to RNP granules (8, 37, 38). Although there have been efforts to understand how these domains affect RNP granule dynamics, their role within these assemblies remains unclear. While various studies have examined specific sequence and composition features that promote phase separation *in vitro* (10, 14, 16-18), or that affect stress-induced assembly of individual proteins (14, 39), this work represents the first study to systematically examine the range of PrLDs that can assemble in response to stress *in vivo*. We have demonstrated that many PrLDs are sufficient to be reversibly recruited to stress granules in yeast, and that this recruitment occurs in a composition-dependent manner.

We were able to discern clear compositional biases among PrLDs that were sufficient to form stress-induced assemblies. Three pieces of data strongly argue that composition is the dominant feature driving stress-induced assembly: the ability of our composition-based predictor to predict which of the PrLDs will assemble into foci; the ability to predict new assembly-prone PrLDs and modulate the assembly propensity of existing PrLDs based solely on composition; and the relative insensitivity of high- and low-scoring PrLD assembly propensity to scrambling.

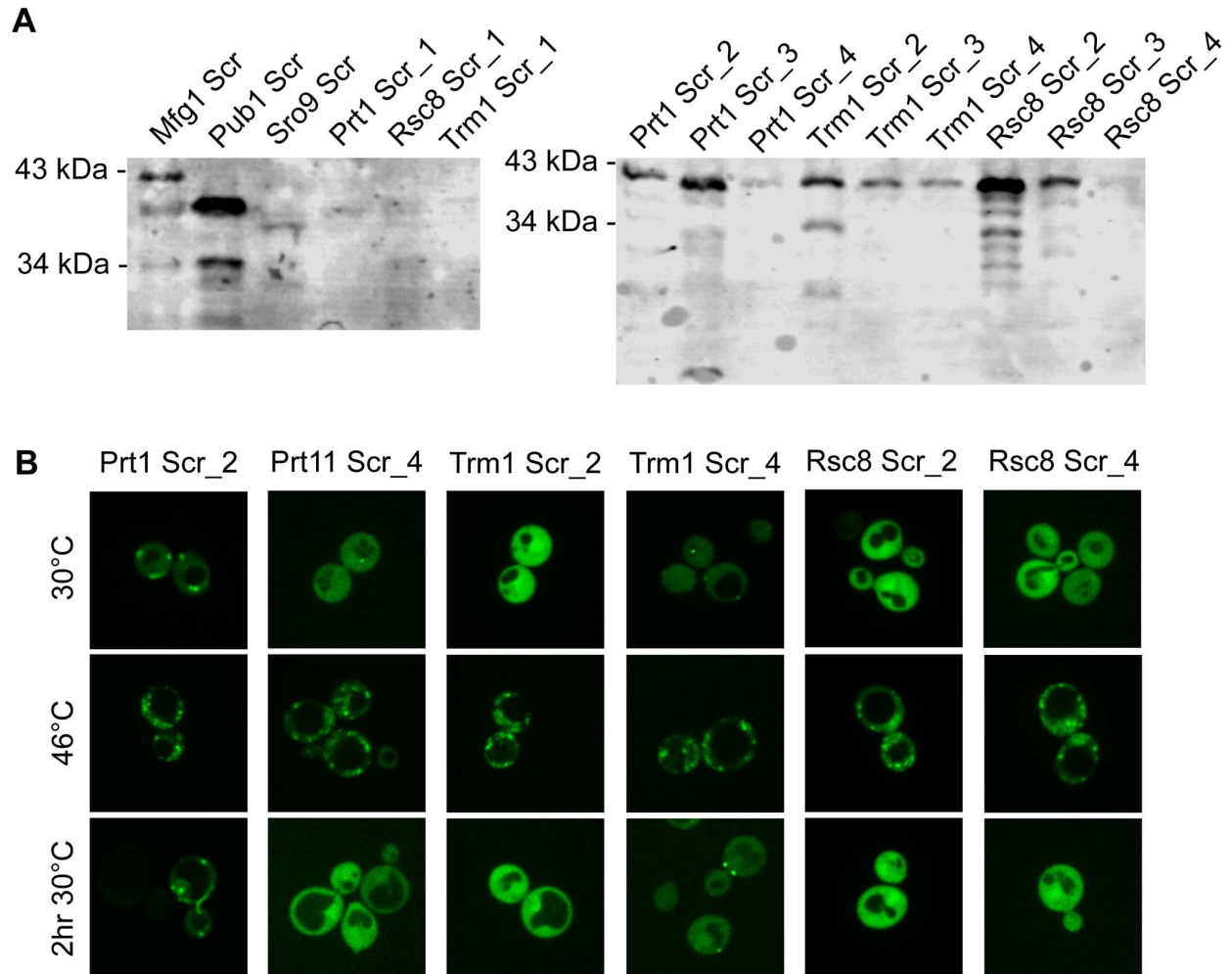


Figure 4.14: PrLD scrambling does not affect stress-induced assembly. (A) Western blot of each PrLD scramble that was tested. Version 3 of the positive scrambled PrLDs are the versions shown in Figure 4.12. (B) Different scrambled versions of each PrLD were fused to the C-terminus of GFP and imaged under normal growth conditions, after 30 minutes of heat shock at 46°C, and after 2 hours of recovery at 30°C.

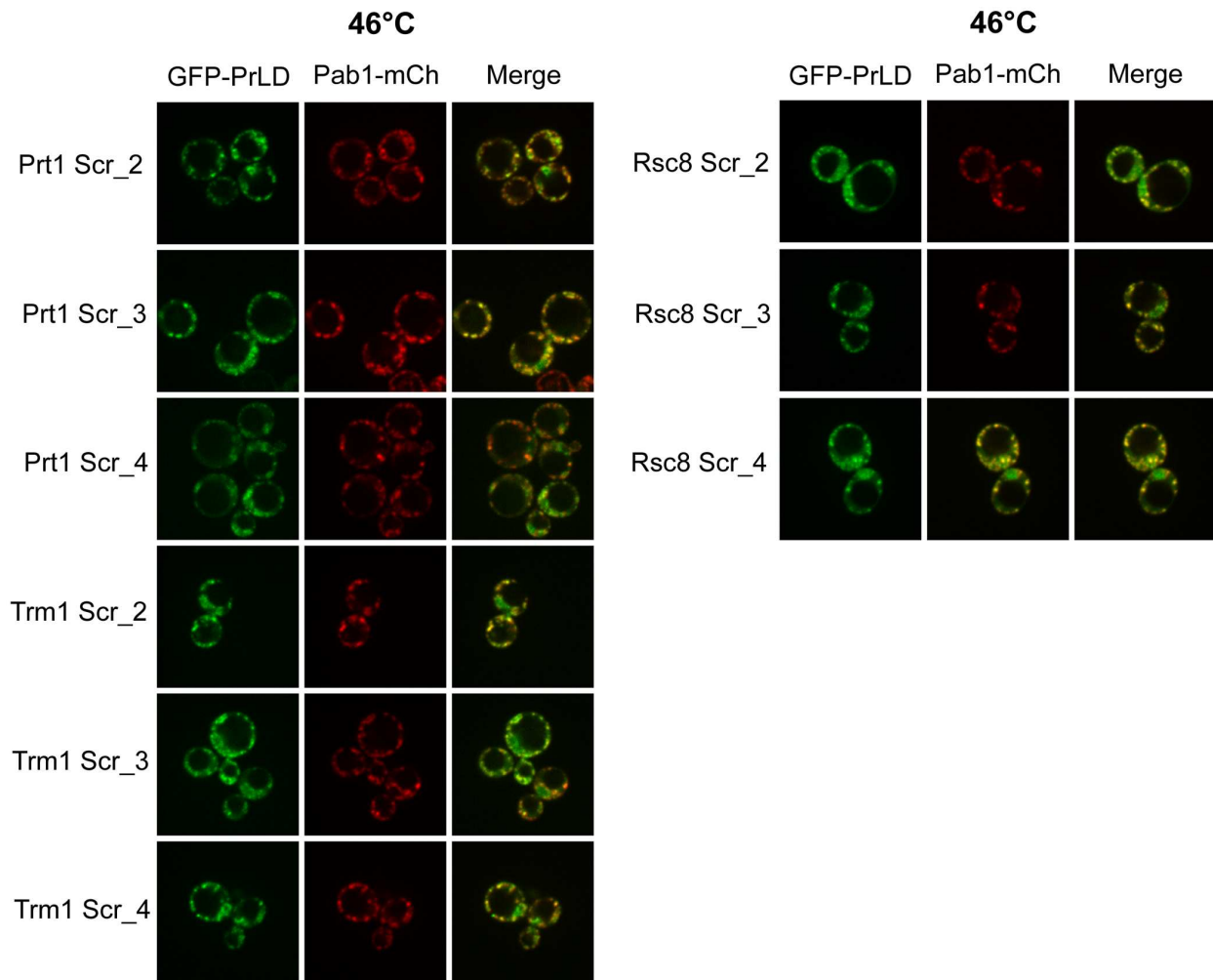


Figure 4.15: PrLD scrambling does not affect recruitment to stress granules. Different scrambled versions of each indicated PrLD were co-expressed with Pab1-mCh as a stress granule marker to assess colocalization with stress granules. Cells were exposed to 30 minutes of heat shock at 46°C prior to imaging.

Nevertheless, our data do not rule out modest effects of primary sequence. Although PrLDs with very high or low assembly scores were accurately predicted solely based on composition, our predictor was less accurate across the middle of the assembly score range. Some of this inaccuracy may simply reflect imperfections in our prediction algorithm. The assembly propensities for each amino acid have large confidence intervals, creating uncertainty in the final predictions. Alternatively, the inaccuracy of our predictions across the middle of the range may suggest that primary sequence can modulate assembly propensity. Primary sequence features could affect the intrinsic assembly propensity of PrLDs, or could promote stress granule recruitment through interactions with binding partners. Such primary sequence effects appear insufficient to overcome strong compositional effects, but may be enough to nudge moderately scored domains across the boundary of assembly, in either direction. By defining the compositional contributions to assembly into stress-induced foci, our work should facilitate the identification of contributing primary sequence motifs.

Some of the biases that we observe are consistent with previous work examining either protein aggregation or LLPS. The overrepresentation of charged residues among assembly-forming PrLDs (Figure 4.8) is consistent with previous work analyzing sequence features that promote LLPS (16, 17, 40). Charge patterning, particularly asymmetric charge distribution, has also been suggested to affect phase separation (16, 17). While our results suggest that charge patterning does not exert a dominant effect on stress-induced assembly, it may be a contributing factor.

Other biases that we observed were more surprising. A variety of evidence suggests that cation- π interactions, particularly between arginine and tyrosine, promote phase separation (10, 16, 18, 41, 42). However, we found that phenylalanine, but not tyrosine, was significantly overrepresented among assembly-forming PrLDs (Figure 4.9), although these data do not

exclude a positive role for tyrosine in some sequences. It is worth noting that phenylalanine can also participate in cation- π interactions and is capable of promoting LLPS, but tends to do so less-efficiently than tyrosine in otherwise equivalent PrLD sequences. It is possible that phenylalanine is favored in our assays because it provides the proper balance of solubility and assembly propensity under our experimental conditions, as suggested in principle previously (18).

Another unexpected compositional bias was the overrepresentation of Q/N residues in PrLDs that do not form assemblies. This result was surprising because PrLDs are defined in part by their high Q/N content (2). This bias against Q/N residues highlights an apparent contradiction in our data. The fact that so many of the tested PrLDs formed stress-induced assemblies clearly suggests that PrLDs are prone to form these assemblies, and indicates that some degree of Q/N enrichment may facilitate PrLD assembly. However, sequence features that made PrLDs maximally “prion-like” actually reduced assembly; although PrLDs tend to contain an overrepresentation of Q/N residues and an underrepresentation of charged and hydrophobic amino acids (23, 35), PrLDs within our set with very high relative Q/N content and very low relative hydrophobic content did not form stress-induced assemblies. These results are analogous to previous studies examining *bona fide* prion formation (23); these studies showed that while compositional similarity to known prion domains is an excellent way to identify prion candidates, such methods are relatively ineffective at ranking these candidates. For stress-induced assembly, our results go one step further; although the prediction algorithm PLAAC was good at identifying candidates for stress-induced assembly, the rankings of the highest-scoring prion candidates were actually anti-correlated with the propensity to form stress-induced assemblies. These results suggest that the sequence features that promote stress-induced assembly partially, but incompletely, overlap with features that make a sequence prion-like. For

example, it is possible that modestly elevated Q/N content promotes assembly, perhaps by increasing disorder propensity, but extremely elevated Q/N content hinders assembly.

We were also surprised to find no substantial differences in compositional bias among three different stresses. Because different stresses likely result in different changes to the cellular environment, we expected that each stress would select for distinct compositional features, potentially explaining why the composition of stress granules differs depending on the stress (33). Instead, this result suggests that common features of the stress response contribute to PrLD assembly. It may also suggest that PrLDs generally provide a consistent contribution to stress-induced assembly across a variety of stresses, while the remainder of each PrLD-containing protein might dictate localization to granules in a stress-type-dependent manner, thereby accounting for the differences in protein composition observed for stress granules under various stress conditions.

Strikingly, all of the assembly-prone PrLDs reversibly localized to stress granules in response to heat shock (Figure 4.2, Figure 4.3A and Figure 4.4). The fact that amino acid composition was the dominant determinant of reversible assembly is consistent with the theory of LLPS; this result suggests that the basic physical properties of a PrLD determine whether it will reversibly partition into the cytosolic or stress granule phase upon stress. These results also highlight the lack of specificity of stress granule recruitment.

Finally, it should be noted that the variable expression levels of the tested PrLDs conceivably could influence assembly behavior, as aggregation is a concentration-dependent process. However, the assembly-prone sequences showed a diverse range of expression levels. Furthermore, the fact that we were able to build a prediction algorithm from this dataset, despite the possibility of variable expression creating noise in the analysis, only further highlights the dominant effects of composition.

With these results we have established a groundwork for understanding the interactions that PrLDs contribute towards RNP granule dynamics. Using our foundational understanding of the effects of composition on PrLD localization, we can now begin to identify additional layers of sequence features that dictate assembly and localization upon stress. Together, these findings will help us to further clarify the role of these domains in the context of RNP granule biology.

REFERENCES

1. Protter DS & Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26(9):668-679.
2. Li L, McGinnis JP, & Si K (2018) Translational Control by Prion-like Proteins. *Trends Cell Biol*.
3. Ramaswami M, Taylor JP, & Parker R (2013) Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154(4):727-736.
4. Gilks N, *et al.* (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15(12):5383-5398.
5. Protter DSW, *et al.* (2018) Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly. *Cell Rep* 22(6):1401-1412.
6. Mittag T & Parker R (2018) Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly. *J Mol Biol*.
7. King OD, Gitler AD, & Shorter J (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* 1462:61-80.
8. March ZM, King OD, & Shorter J (2016) Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res*.
9. Li YR, King OD, Shorter J, & Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201(3):361-372.
10. Kato M, *et al.* (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149(4):753-767.
11. Patel A, *et al.* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162(5):1066-1077.
12. Molliex A, *et al.* (2015) Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* 163(1):123-133.
13. Lin Y, Protter DS, Rosen MK, & Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60(2):208-219.
14. Riback JA, *et al.* (2017) Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response. *Cell* 168(6):1028-1040 e1019.
15. Murakami T, *et al.* (2015) ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* 88(4):678-690.
16. Pak CW, *et al.* (2016) Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* 63(1):72-85.
17. Nott TJ, *et al.* (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57(5):936-947.
18. Wang J, *et al.* (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*.
19. Conicella AE, Zerze GH, Mittal J, & Fawzi NL (2016) ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure*.
20. Paul KR, Hendrich CG, Waechter A, Harman MR, & Ross ED (2015) Generating new prions by targeted mutation or segment duplication. *Proc Natl Acad Sci U S A* 112(28):8584-8589.

21. Paul KR, *et al.* (2017) Effects of Mutations on the Aggregation Propensity of the Human Prion-Like Protein hnRNPA2B1. *Mol Cell Biol* 37(8).
22. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry* 75:333-366.
23. Toombs JA, McCarty BR, & Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-332.
24. Lancaster AK, Nutter-Upham A, Lindquist S, & King OD (2014) PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* 30(17):2501-2502.
25. Sherman F (1991) Getting started with yeast. *Methods in enzymology* 194:3-21.
26. Brachmann CB, *et al.* (1998) Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* 14(2):115-132.
27. Shattuck JE, Waechter AC, & Ross ED (2017) The effects of glutamine/asparagine content on aggregation and heterologous prion induction by yeast prion-like domains. *Prion*:1-16.
28. Ross ED, Edskes HK, Terry MJ, & Wickner RB (2005) Primary sequence independence for prion formation. *Proc Natl Acad Sci U S A* 102(36):12825-12830.
29. Fredrickson EK, Gallagher PS, Clowes Candadai SV, & Gardner RG (2013) Substrate recognition in nuclear protein quality control degradation is governed by exposed hydrophobicity that correlates with aggregation and insolubility. *J Biol Chem* 288(9):6130-6139.
30. Wallace EW, *et al.* (2015) Reversible, Specific, Active Aggregates of Endogenous Proteins Assemble upon Heat Stress. *Cell* 162(6):1286-1298.
31. Buchan JR (2014) mRNP granules. *RNA Biology*:e29034.
32. Buchan JR, Muhrad D, & Parker R (2008) P bodies promote stress granule assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 183(3):441-455.
33. Buchan JR, Yoon JH, & Parker R (2011) Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *J Cell Sci* 124(Pt 2):228-239.
34. Munder MC, *et al.* (2016) A pH-driven transition of the cytoplasm from a fluid- to a solid-like state promotes entry into dormancy. *eLife* 5.
35. DePace AH, Santoso A, Hillner P, & Weissman JS (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93(7):1241-1252.
36. Ross ED, Baxa U, & Wickner RB (2004) Scrambled prion domains form prions and amyloid. *Mol Cell Biol* 24(16):7206-7213.
37. Decker CJ, Teixeira D, & Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179(3):437-449.
38. Reijns MA, Alexander RD, Spiller MP, & Beggs JD (2008) A role for Q/N-rich aggregation-prone regions in P-body localization. *J Cell Sci* 121(Pt 15):2463-2472.
39. Franzmann TM, *et al.* (2018) Phase separation of a yeast prion protein promotes cellular fitness. *Science* 359(6371).
40. Elbaum-Garfinkle S, *et al.* (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* 112(23):7189-7194.
41. Lin Y, Currie SL, & Rosen MK (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem*.

42. Qamar S, *et al.* (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation- π Interactions. *Cell* 173(3):720-734 e715.

CHAPTER FIVE: CONCLUSION

Throughout the course of this work, I have presented new findings that further our understanding of how prion-like domains contribute to RNP granule biology. Aggregation-promoting mutations in PrLDs that increase their prion-like character do not appear to disrupt the dynamics of stress granules or p-bodies, indicating that prion-like aggregation does not underlie the interactions of PrLDs in RNP granules. Instead, PrLDs that are compositionally less prion-like appear to be better suited to assemble into foci upon stress and to localize to stress granules specifically. These results prompt further questions regarding how PrLDs contribute to assembly and structure of these granules.

Aggregation-Promoting Mutations in the Lsm4 PrLD Do Not Disrupt P-body Disassembly

Although the aggregation propensity of the Lsm4 PrLD can effectively be increased using mutations designed to enhance prion-like character (Figure 2.1), these mutations have no obvious effect on p-body disassembly (Figure 2.3). There are a variety of reasons that could explain why p-body disassembly is unaffected by these mutations, one of them being that these experiments were performed in very specific strain. The strain that was used for this study, which contains a deletion of the p-body protein Edc3 as well as the Lsm4 PrLD, was one that was previously reported to have significantly reduced p-body formation (1). Because p-body formation supposedly requires the Lsm4 PrLD, I thought this system would be ideal for studying the effects of aggregation-promoting mutations in a PrLD on p-body disassembly. The system turned out to be much less robust than I had originally thought, with a reduction in p-body formation from about 60% with the Lsm4 PrLD to about 40% without the Lsm4 PrLD (Figure 2.4B). This subtle difference in p-body formation between wild-type and mutant cells suggests

that it might not be the best system to use to determine large differences in p-body formation and persistence, which was the original goal of this project.

Additionally, Lsm4 did not appear to localize to p-bodies very robustly when it was tagged in this same *edc3Δlsm4CΔ* strain (Figure 2.5A), suggesting that it may not be a good candidate protein on which to study the effects of aggregation-promoting mutations. There are a few possible explanations for why Lsm4 did not localize to p-bodies in a robust manner in this strain. First, Lsm4 (and its PrLD) may not be as important for p-body formation as originally thought. Second, Lsm4 may be localizing to p-bodies, but to such a small extent that localization is undetectable by microscopy. Finally, it is possible that Lsm4 is more strongly recruited to p-bodies in wild-type strain backgrounds and Edc3 might be important for recruiting Lsm4 to p-bodies, which could explain why Lsm4 appeared to have such a minimal level of recruitment.

Finally, only two different aggregation-promoting Lsm4 PrLD mutants were analyzed in this study, with aromatic residues added to increase aggregation-propensity. Perhaps this type of mutation was not effective enough to disrupt p-body assembly, but other ones may be. It is possible that there are only certain types of mutations that would be successful at increasing the aggregation propensity of the Lsm4 PrLD to an extent at which p-body disassembly is affected. Further understanding of the interactions of PrLDs within p-bodies will help to elucidate the types of mutations that might be successful at causing p-body persistence.

Lsm4 Future Directions

In future experiments for this project, I would start by using a different system. The most effective system to use would likely be one in which p-body formation was completely, or almost completely eliminated without the Lsm4 PrLD, such as a *dhh1Δedc3Δlsm4ΔC* strain in which p-body assembly is observed in approximately 15% of cells (2). This is a strain in which I would expect the presence of the Lsm4 PrLD to more strongly promote p-body formation since

other structural p-body components are eliminated. Mutations made in a PrLD that was essential for assembly might be well-suited to promote persistent p-body formation. Additionally, it would be essential to first test whether Lsm4 was even being recruited to p-bodies in order to make further conclusions about the effects of these mutations.

Another potential direction to pursue would be to examine a broader range of mutations. Although results from this study suggest that increasing the aggregation propensity of a p-body PrLD does not disrupt disassembly, this study used a very small sample size of only two aggregation-promoting mutants, which is not enough to make any broad conclusions about whether any aggregation-promoting mutations can disrupt p-body assembly. A wider variety of mutations could be tested for aggregation-promoting ability, such as substitutions of aggregation-inhibiting residues with aggregation-promoting residues or deletions of aggregation-inhibiting residues. This larger sample size would help to validate the broad conclusion that aggregation-promoting mutations in p-body PrLDs do not disrupt p-body disassembly. Additionally, PrLDs from other p-body proteins could also be mutated to see if components other than Lsm4 have a greater effect on the overall aggregation propensity of p-bodies.

While I assumed that aggregation-promoting mutations in the Lsm4 PrLD would disrupt p-body dynamics enough to prevent p-body disassembly, I never actually tested how p-body dynamics were affected by these mutations in Lsm4. Testing the dynamic state of Lsm4 within p-bodies using a photobleaching method, such as FRAP, could help determine whether these mutations made in Lsm4 are actually changing the material state of the protein when in p-bodies. Unfortunately, p-bodies are so small that FRAP-type experiments may not be possible. Given that Lsm4 does not robustly localize to p-bodies in the strain I used for these experiments, it also seems unlikely that the mutations introduced effectively disrupted p-body dynamics.

Alternatively, SDD-AGE or thioflavin T staining may be more successful to determine whether or not Lsm4 is forming more ordered aggregates within p-bodies when mutated.

Finally, eliminating components of the PQC machinery in addition to aggregation-promoting mutations in the Lsm4 PrLD together might be more effective at promoting persistent p-body formation, rather than mutation alone. Evidence points to an increasing role of defective PQC components in disease states (3, 4), so perhaps a combination of aggregation-promoting mutations in PrLDs in addition to a decline in PQC activity is required for persistence of RNP granules. These findings are consistent with my results, given that introducing aggressive mutations into the Lsm4 PrLD alone does not perturb p-body dynamics. Another possibility is that p-body disassembly is so robust that it cannot be easily perturbed. This reversible aggregation system may have redundancies in place that prevent the accumulation of aberrant aggregates. Interestingly, in contrast with stress granules, no p-body proteins have yet been implicated in disease, which further indicates that it may be an especially robust functional aggregation system.

Aggregation-Promoting Mutations in the PrLDs of Core Stress Granule Proteins Are Not Sufficient to Prevent Stress Granule Disassembly

Similar to the Lsm4 p-body system, aggregation-promoting mutations in the PrLDs of the core stress granule proteins Pab1 and Pbp1 effectively enhanced aggregation of these domains (Figure 3.1 and Figure 3.2), but were unable to perturb stress granule disassembly (Figure 3.4). I reasoned that stress granules would be a better system than p-bodies to target for mutation for a few reasons. One, unlike p-body proteins, many stress granule proteins contain disease-associated mutations (5, 6), indicating that this system might be more sensitive to mutation. Second, stress granules are only present during stress (7), whereas p-bodies are present both normally and under stress (8). Finally, stress granules have also been shown to be more solid-like

than p-bodies in yeast (9), indicating that they might have a greater propensity to persist due to aberrant aggregation of proteins within. However, despite the fact that this system appeared to be better suited to this type of mutational study, the results ended up being quite similar to the Lsm4 study. I could effectively increase the aggregation propensity of both PrLDs, this time by deleting proline residues to try to promote amyloid formation, but these mutations were not sufficient to promote persistence of stress granules, further indicating that the reversible aggregation associated with RNP granules is a very robust and tightly controlled process that is difficult to perturb.

Stress Granule PrLD Mutation Future Directions

Expanding the dataset of mutations would help to elucidate any other mutations that might be more effective at promoting hyper-aggregation of stress granules than the mutations I chose to make. I investigated only one type of mutation, deletions of proline residues that I thought were inhibiting aggregation. Another tactic would be to substitute these residues with residues that are more aggregation-promoting, an example being substitutions of proline residues for aromatic residues, or other aggregation-promoting residues. Additionally, other aggregation-inhibiting residues besides prolines could also be targeted for mutation. Increasing the number and types of mutations made would help to determine whether the types of mutations made are just not effective enough at promoting stress granule persistence, or if other factors, such as the PQC machinery, are also playing a role in destroying persistent aggregates that might otherwise form upon introduction of these mutations.

The majority of this study was performed using acute heat shock to stress the cells; however, chronic stress is likely more relevant in disease. Given that the goal of this project is to try to recapitulate disease-like aggregates, it might be worthwhile to further investigate the effects of these aggregation-prone mutants on different types of chronic stress, including aging.

It is possible that aggregation-promoting mutations do not have an effect on stress granules that are short-lived, as observed during acute stress, but these mutations could be more detrimental for stress granules that last a long time, as they might during chronic stress. Although preliminary results suggest that there is no difference between normal and mutant forms of Pab1 and Pbp1 during chronic stress (Figure 3.6), this assay only tested for growth defects over time and did not provide any information about stress granule formation that may have been occurring. To determine if chronic stress has a different effect on stress granule formation and recovery, cells could be exposed to chronic heat stress, or perhaps aging, and then imaged to assess stress granule formation. If stress granule formation is occurring, testing disassembly by returning the cells to normal growth conditions could provide insight into whether a lasting stress provides enough time for stronger interactions to develop that might prevent the granules from disassembling.

It is striking that increasing the aggregation propensity of PrLDs from both p-bodies and stress granules successfully increases aggregation, as observed through various overexpression assays; yet, when these mutated proteins are placed in an endogenous context, in the full-length protein under lower, normal expression levels, these aggregation-promoting mutations have no effect on the disassembly of RNP granules, despite being confined at higher local concentrations within these granules. Altogether, these data suggest that other factors are responsible for keeping these granules in a fluid-like state and that these systems are robust enough to deal with aberrant aggregates that may be present.

Stress-Induced Assembly of PrLDs

In this study, we found that PrLD localization to stress-induced assemblies, and to stress granules specifically in the case of heat shock, is determined by composition. Hydrophobic and charged amino acids are overrepresented among PrLDs that assemble upon stress, and glutamine

and asparagine (Q/N) are overrepresented among PrLDs that do not assemble upon stress (Figure 4.8). Additionally, we found that amino acid composition can be used to predict PrLDs that are and are not sufficient to localize to stress-induced assemblies (Figure 4.10) and that we can use these compositional biases to effectively design rational mutations to promote or prevent assembly upon stress (Figure 4.12A,B). Perhaps most interestingly, we found that scrambling the primary sequence of these PrLDs does not affect the ability of that PrLD to localize to stress-induced assemblies (Figure 4.12C,D) and also does not affect localization to stress granules (Figure 4.14B), suggesting that composition is also responsible for PrLD recruitment to specific stress-induced assemblies.

Interestingly, the amino acids that are favored for localization to stress-induced assemblies, charged and hydrophobic residues, are those that are underrepresented among prion forming domains (10). Additionally, Q/N residues, which are overrepresented among prion forming domains (11), appear to be disfavored for stress-induced assembly, further suggesting that these two processes represent different types of aggregation. Because localization of PrLDs to stress-induced assemblies seems to require different amino acids than for prion-like aggregation, this could explain why the *Lsm4* and *Pab1/Pbp1* proline deletion projects were unsuccessful. Those studies were based on the assumption that prion-like aggregation underlies localization of these domains to stress granules. Now that we know this reasoning is incorrect, perhaps we can use our newfound knowledge about PrLD recruitment to stress granules to further investigate the causes of persistent aggregate formation of stress granules as they pertain to disease.

PrLD Stress Assembly Future Directions

One of the most interesting findings from this study is that domains that are compositionally less prion-like, with fewer Q/N residues, are more prone to stress-induced

assembly than domains that are more prion-like. This result is especially puzzling because PrLDs are defined, in part, by their high Q/N content (12). Even though many PrLDs are capable of assembling into foci upon stress, this result suggests that other intrinsically disordered domains, which are not necessarily prion-like, may form stress-induced assemblies more readily than PrLDs specifically. Alternatively, because the original dataset was only composed of PrLDs specifically, all of which have elevated Q/N content to begin with, it is plausible that stress-induced assembly can only occur with PrLDs, but only those with lower Q/N content. For example, there may be a range of acceptable Q/N content for a domain to assemble, but too many or too few Q/N residues precludes the domain from assembling. Evaluating stress-induced assembly of a new dataset of intrinsically disordered domains that have lower Q/N content and are not prion-like would be a way to determine if domains other than PrLDs are more readily recruited to assemblies than PrLDs.

The initial dataset that was tested was very small, with only 35 PrLDs used for generating the predictor. Many PrLDs were eliminated based on the fact that they did not appear to express by both western blot and imaging (Table 4.1). Interestingly, 9 out of the 12 proteins that were eliminated for little to no expression scored positive by our stress assembly predictor (Table 4.1), suggesting that these domains might be more aggregation-prone due to their higher assembly propensity. One possibility is that the cell can efficiently dispose of these aggregation-prone PrLDs when expressed at this level. Testing expression of these PrLDs under stronger promoters may stabilize some of the non-expressing PrLDs enough to determine whether they are capable of localizing to stress-induced assemblies and could reveal more domains from the original dataset that are actually positives. Additionally, because the original dataset is small, expanding the dataset will be essential to making better predictions. We are successful in predicting the behavior of PrLDs with very high or very low scores, but we do not have very good predictive

capabilities for PrLDs that have intermediate scores. Expanding the dataset with PrLDs that score in this middle range will help to increase the accuracy of our predictor.

One other interesting observation from this study is that many of the assembling PrLDs that localized to stress granules upon heat shock are not from known stress granule proteins. Of the PrLDs that localized to stress granules, only Ded1 and Prt1 have been previously reported to localize to stress granules (13, 14). It is possible that the proteins of the other assembling PrLDs also localize to stress granules, which could imply that screening for assembling PrLDs may provide a new way to identify more stress granule components. Another possibility is that PrLDs in isolation can localize to stress granules based on composition, but when in the context of the full-length protein, are not sufficient to cause localization of that protein. Testing localization of the full-length proteins of each assembling PrLD would provide more insight into whether PrLD recruitment is sufficient to drive localization of full-length proteins to stress granules.

So far, all experiments for this PrLD stress assembly project have been performed *in vivo*, further analysis on purified PrLDs *in vitro* could provide insight into the physical nature of assembling and non-assembling PrLDs. Assuming that stress granule assembly proceeds through a liquid-liquid phase separation (LLPS) mechanism, the assembling PrLDs are likely also assembling via a LLPS mechanism. Purifying these proteins and testing them for LLPS characteristics *in vitro* could provide insight into whether they are assembling through this type of mechanism. If the assembling PrLDs are more prone to LLPS than the non-assembling PrLDs, that would indicate that these domains are poised for assembly based on their propensity to phase separate. These LLPS experiments could help to verify if the assembly compositions we are observing are actually compositions consistent with phase separation in general. Indeed, many of the compositional biases we have discovered for assembling PrLDs are consistent with compositional trends observed among phase separating proteins. Charged residues in particular

have been shown to play roles in phase separation, through both cation- π interactions, and charge patterning (15-17), and these residues were significantly overrepresented among the assembly-prone PrLDs (Figure 4.8). Although we have not yet examined whether sequence features such as cation- π interactions or charge patterning are contributing to PrLD assembly, those features would be interesting to evaluate in future studies.

Altogether, these results provide new insight into the compositional features of PrLDs that promote their aggregation and recruitment to RNP granules. Interestingly, the features that promote PrLD recruitment to stress-induced assemblies are different than the features promoting prion-like aggregation, which is a more traditional view of RNP granule assembly (6). My work here demonstrates that canonical prion-like aggregation is likely not responsible for PrLD association to stress granules. Instead, the compositional biases observed among assembling PrLDs suggest different mechanisms of association. This work also highlights the fact that these reversible aggregation systems are robust, and the cell can very tightly control the aggregation associated with them, even upon introduction of highly aggregation-prone species. Together, these findings contribute further understanding of the functional, reversible aggregation process underling RNP granule biology.

REFERENCES

1. Decker CJ, Teixeira D, & Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179(3):437-449.
2. Rao BS & Parker R (2017) Numerous interactions act redundantly to assemble a tunable size of P bodies in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*.
3. Alberti S & Carra S (2018) Quality Control of Membraneless Organelles. *J Mol Biol*.
4. Alberti S, Mateju D, Mediani L, & Carra S (2017) Granulostasis: Protein Quality Control of RNP Granules. *Front Mol Neurosci* 10:84.
5. Li YR, King OD, Shorter J, & Gitler AD (2013) Stress granules as crucibles of ALS pathogenesis. *J Cell Biol* 201(3):361-372.
6. Ramaswami M, Taylor JP, & Parker R (2013) Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell* 154(4):727-736.
7. Protter DS & Parker R (2016) Principles and Properties of Stress Granules. *Trends Cell Biol* 26(9):668-679.
8. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, & Parker R (2005) Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 11(4):371-382.
9. Kroschwald S, *et al.* (2015) Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife* 4:e06807.
10. Toombs JA, McCarty BR, & Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-332.
11. DePace AH, Santoso A, Hillner P, & Weissman JS (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93(7):1241-1252.
12. Li L, McGinnis JP, & Si K (2018) Translational Control by Prion-like Proteins. *Trends Cell Biol*.
13. Jain S, *et al.* (2016) ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164(3):487-498.
14. Grousl T, *et al.* (2009) Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *J Cell Sci* 122(Pt 12):2078-2088.
15. Wang J, *et al.* (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*.
16. Nott TJ, *et al.* (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* 57(5):936-947.
17. Pak CW, *et al.* (2016) Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. *Mol Cell* 63(1):72-85.

APPENDIX ONE: THE EFFECTS OF MUTATIONS ON THE AGGREGATION PROPENSITY OF THE HUMAN PRION-LIKE PROTEIN HNRNPA2B1²

Introduction

Amyloid fibrils are ordered, self-propagating, β -sheet-rich protein aggregates (1, 2). In *Saccharomyces cerevisiae*, numerous prions (infectious proteins) have been identified that result from the conversion of proteins to an infectious amyloid form (3, 4). Most of the yeast prion proteins contain low-complexity, glutamine/asparagine (Q/N) rich prion domains (5). Hundreds of human proteins contain similar prion-like domains (PrLDs), defined as protein segments that compositionally resemble yeast prion domains (6, 7). PrLDs are a subset of low complexity sequence domains (LCDs) that are found in about one third of the human proteome, and which are generally predicted to be intrinsically disordered (7, 8). PrLDs are particularly enriched in RNA-binding proteins (7). Mutations in various PrLD-containing RNA-binding proteins have been linked to degenerative disorders, including amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (7, 9).

A number of these PrLD-containing RNA binding proteins are components of RNA-protein granules, such as P-bodies and stress granules (9, 10), and the PrLDs are thought to mediate interactions that are involved in the formation of these granules (11-13). These PrLD-containing RNA binding proteins can form a range of assemblies, which differ in the degree of order in the structure, and possibly in the nature of the underlying interactions (14-17). These range from highly dynamic liquid-liquid phase separations, in which liquid droplets are formed in a temperature- and concentration-dependent manner (15-17); to hydrogels, consisting of metastable amyloid-like fibers (14); to more stable, ordered amyloid aggregates (16, 17).

² This chapter has been reformatted from the following publication: Paul KR, Molliex A, Cascarina S, Boncella AE, Taylor JP, Ross ED. *Mol. Cell. Biol.* (2017) 37:e00652-16. My contribution consisted of the ThT assays in Figure A1.7 as well as assisting with TEM imaging of the endpoint fibrils.

Disease-associated mutations appear to specifically shift these proteins towards the amyloid state (16-18). This raises the intriguing hypothesis that these PrLDs are evolved to mediate the weak, dynamic interactions involved in formation of dynamic RNA-protein granules, but disease-associated mutations promote conversion of the PrLDs to more stable structures (9, 10, 19).

The human heterogeneous nuclear ribonucleoprotein hnRNPA2B1 provides a useful model to examine this hypothesis. hnRNPA2B1 is a ubiquitously expressed RNA binding protein that has two alternatively spliced forms, A2 and B1, which differ by 12 amino acids at the N-terminus. The shorter hnRNPA2 is the predominant isoform in most tissues. hnRNPA2B1 contains a PrLD (Figure A.1A), and a single point mutation (D290V in hnRNPA2) in this PrLD causes multisystem proteinopathy (18). Interestingly, mutations at the corresponding position of a paralogous heterogeneous nuclear ribonucleoprotein, hnRNPA1, can cause either multisystem proteinopathy or familial ALS (18). The mutations in both proteins promote incorporation into stress granules, and in *Drosophila* cause formation of cytoplasmic inclusions. In vitro, the mutations accelerate formation of amyloid fibrils. In yeast, the core prion-like domain is able to support prion formation when inserted in the place of the portion of the prion domain of Sup35 that is responsible for nucleating prion formation (18, 20). Thus, hnRNPA2 provides a range of experimental systems to monitor the effects of mutations on protein aggregation.

Intriguingly, PAPA and ZipperDB, two algorithms designed to predict amyloid or prion propensity, both correctly predict the effects of the three known disease-associated mutations in hnRNPA2B1 and hnRNPA1 (18). These results suggest that it might be possible to predict the effects of other mutations in these proteins, and to rationally design mutations to alter aggregation propensity. However, this prediction success is currently based on a very small sample size: just one mutation in hnRNPA2, and two in hnRNPA1. Additionally, PAPA and

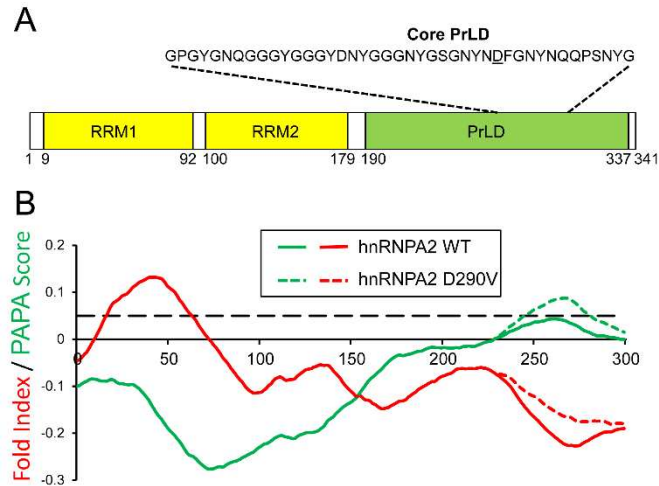


Figure A.1: hnRNPA2 contains a predicted prion-like domain. (A) Schematic of the hnRNPA2 domain architecture. **(B)** The disease-associated D290V mutation increases predicted prion-proneness. PAPA scores (green) and FoldIndex scores (red) were calculated for hnRNPA2 wild-type (solid) and D290V (dashed). The black dotted line indicates a PAPA score of 0.05, the threshold that was most effective at separating prion-like domains with and without prion activity (21). Regions with high PAPA scores and negative FoldIndex scores are predicted to be prion-prone. Adapted from Kim et al. (18).

ZipperDB use very different features to score aggregation propensity, so it is unclear which of these features is most predictive.

Specifically, PAPA was derived by replacing an 8-amino-acid segment from a scrambled version of Sup35 with a random sequence to build a library of mutants, and then screening this library for prion formation (22). A prion propensity score was then derived for each amino acid by comparing its frequency among the prion-forming isolates relative to the starting library. PAPA predicts prion activity by first using FoldIndex to identify regions of proteins that are predicted to be intrinsically disordered, and then scanning these regions with a 41-amino acid window size, adding up the prion propensity scores of each amino acid across the window (21, 23). By contrast ZipperDB is a structure-based algorithm designed to look for short peptide fragments with a high propensity to form steric zippers (24). ZipperDB was developed by first solving the structure of a 6-amino-acid peptide from Sup35 in its amyloid conformation (25). The peptide was found to form a cross- β -sheet structure, with tight steric zipper interactions between the sheets. ZipperDB predicts amyloid propensity by threading 6-amino-acid peptides into this structure *in silico*, and using Rosetta to determine the energetic fit.

Thus, PAPA solely considers amino acid composition, and uses a large window size, while ZipperDB uses a much smaller window, and is sensitive to primary sequence. Despite these differences, both accurately predicted the effects of the hnRNP mutations. PAPA predicts that the aggregation propensity of the wild-type hnRNPA2 PrLD falls just below the threshold for prion-like aggregation, while the mutation increases aggregation propensity well beyond this threshold (Figure A.1B). ZipperDB predicts that the disease-associated mutations should create a strong steric zipper (18).

Here, to define the sequence features that drive aggregation, we designed a variety of mutations in the hnRNPA2 prion-like domain. Both in yeast and *in vitro*, the effects of mutations

could be predicted entirely based on amino acid composition. By contrast, while the original disease-associated mutations created predicted steric zipper motifs, such motifs were neither necessary nor sufficient for aggregation in yeast. Although composition alone accurately predicted the effects of mutations on isolated prion-like domain fragments both in yeast and *in vitro*, it was less accurate at predicting their effects in the context of the full-length protein in *Drosophila*. This highlights a critical limitation of our current prediction methods. While these methods can predict the effects of mutations on intrinsic aggregation propensity, other factors (including interactions with other parts of the protein, interacting proteins, and localization) are currently much more challenging to predict.

Materials and Methods

Yeast Strains and Media

Standard yeast media and methods were as previously described (26). All experiments were performed in strain YER635/pJ533 ((27); α *kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 ppq1::HIS3 sup35::KanMx*). pJ533 (*URA3*) expresses *SUP35* from the *SUP35* promoter. Yeast were grown at 30°C for all experiments.

Prion Formation in Yeast

Plasmids pER599 and pER600 (cen, *LEU2*) expressing wild-type and D290V hnRNPA2-Sup35 fusions, respectively, were previously described (18). All additional mutations were made by PCR, and confirmed by DNA sequencing. Plasmids were transformed into YER635/pJ533, selected on medium lacking leucine, and then transferred to 5-fluoroorotic-acid-containing medium to select for loss of pJ533.

To construct plasmids to transiently overexpress the PrLDs fused to GFP, the NM domain (the prion domain, plus the adjacent middle domain; see Figure A.2A) of each hnRNPA2-Sup35 fusion was amplified with oligonucleotides EDR1624

(GAGCTACTGGATCCACAATGTCAGGACCTGGATATGGCAACCAG) and EDR1924 (GTCGATGCTACTCGAGTCGTTAACAACCTTCGTCATCCACTTC). The resulting PCR products were digested with BamHI and XhoI and inserted into BamHI/XhoI cut pER760, a TRP1 plasmid that contains GFP under control of the *GALI* promoter (28).

Prion formation assays were performed as previously described (29). Briefly, cells expressing a given hnRNPA2-Sup35 fusion as the sole copy of Sup35 were transformed with either an empty vector (pKT24; (30)) or a plasmid expressing the matching PrLD-GFP fusion under control of the *GALI* promoter. Cells were grown for 3 days in galactose/raffinose dropout medium lacking tryptophan. Ten-fold serial dilutions were then plated onto synthetic complete medium lacking adenine to select for [*PSI*⁺] cells and onto medium with adenine to test for cell viability.

Western Blot

To probe PrLD-GFP expression levels in yeast, TRP1 plasmids expressing the PrLD-GFP fusion were transformed into the corresponding mutant strain. Low density cultures were pre-grown in raffinose dropout medium overnight, diluted to an OD of 1.0 in 10mL of 3% galactose/raffinose dropout medium, and grown for 4 h. Cells were harvested by centrifugation. Cell pellets were lysed as previously described (31), with protease inhibitor cocktail (Gold Biotechnology) included in the lysis buffer. Lysates were normalized based on total protein concentration, as determined by Bradford assay (Sigma). Proteins were separated on SDS/12% PAGE gels and transferred to a PVDF membrane, and immunoblotted, using a monoclonal anti-GFP primary antibody (Santa Cruz Biotechnology), and Alexa Fluor IR800 goat anti-mouse secondary antibody (Rockland).

To probe endogenous expression levels, log-phase cultures were harvested by centrifugation, and cells lysed as above. Proteins were separated by SDS-PAGE and analyzed by

western blot, using a monoclonal antibody against the Sup35C domain (BE4 (32), from Cocalico Biologicals, kindly made available by Susan Liebman) as the primary antibody, and Alexa Fluor IR800 goat anti-mouse (Rockland) as the secondary antibody.

Fly Stocks and Culture

Mutagenesis using the QuickChange Lightning kit (Agilent) was performed on pUASTattB-wild type hnRNPA2 construct as previously described (18). Flies carrying transgenes in pUASTattB vectors were generated by performing a standard injection and ϕ C31 integrase-mediated transgenesis technique (BestGene Inc.). To express a transgene in muscles, Mhc-Gal4 was used (from G. Marqués). All *Drosophila* stocks were maintained in a 25°C incubator with a 12 h day/night cycle and a standard diet.

Preparation of Adult Fly Muscle for Immunofluorescence

Adult flies were embedded in a drop of OCT compound (Sakura Finetek) on a glass slide, frozen in liquid nitrogen and bisected sagittally by using a razor blade. After being fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS), fly tissues were permeabilized in PBS containing 0.2% Triton X-100, and indiscriminant binding was blocked by adding 5% normal goat serum in PBS. The hemithoraces were stained with anti-hnRNPA2B1 (EF-67) antibody (Santa Cruz Biotechnology) followed with Alexa-488-conjugated secondary antibody (Invitrogen), Texas Red-X phalloidin (Invitrogen) and DAPI according to manufacturer's instructions. Stained hemithoraces were mounted in 80% glycerol and the muscles were imaged with a Marianas confocal microscope (Zeiss, x63).

Fly Thoraces Fractionation Protocol

Thoraces of at least 15 adult flies were dissected, homogenized in RIPA buffer, and lysed on ice for 15 min. The cell lysates were sonicated and then cleared by centrifugation at 100,000 \times g for 30 min at 4 °C to generate the RIPA soluble samples. To prevent carry-overs, the

resulting pellets were washed with RIPA buffer. RIPA insoluble pellets were then extracted with urea buffer (7 M urea, 2 M thiourea, 4% CHAPS, 30 mM Tris, pH 8.5), sonicated, and centrifuged at $100,000 \times g$ for 30 min at 22 °C. Protein concentration was determined by bicinchoninic acid method (Pierce), and samples were boiled for 5 min and analysed by the standard western blotting method provided by Odyssey system (LI-COR) with 4–12% NuPAGE Bis-Tris Gel (Invitrogen) and anti-hnRNPA2B1 (DP3B3) antibody (Abcam, 1:2000) and anti-actin antibody (Santa Cruz, 1:10000).

In Vitro Aggregation Assays

A 96-well plate was treated with 5% casein solution for 5 minutes at room temperature, and then rinsed with DI water and allowed to dry. Synthetic peptides (GenScript) were dissolved at 2.5 mM in 6M guanidine HCl. Peptides were then diluted approximately 100-fold to a final concentration of 25 μ M in 10 mM sodium phosphate, 150 mM NaCl, 12.5 μ M thioflavin T, 0.02% casein, pH 7.4 in the 96-well plate to initiate aggregation. Fluorescence was monitored in a Victor3 Perkin Elmer fluorescence plate reader, with excitation and emissions wavelengths of 460 and 490 nm, respectively. Reactions were monitored for 48 h. Between readings, reactions were incubated without agitation for 3 minutes, and then shaken for 10 sec. The fraction aggregated was calculated by normalizing relative to the final fluorescence of the well.

For electron microscopy of *in vitro* aggregation reactions, 10-20 μ l of sample was incubated on carbon copper grids for 5 min, and then rinsed with distilled water. Grids were stained with 1% uranyl acetate for 30 sec, and observed on a JEOL JEM-1400 TEM, imaging with a Gatan Orius 832 Camera.

Results

Hydrophobic and Aromatic Residues Promote Aggregation

We previously developed a yeast system to monitor the prion-like activity of hnRNPA2 (18). The yeast prion [*PSI*⁺] is the prion form of the translation termination factor Sup35 (33, 34). Sup35 has three functionally distinct domains: an N-terminal prion domain that is required for prion aggregation; a C-terminal functional domain that is necessary and sufficient for Sup35's normal function in translation termination; and a highly charged middle domain that is not required for either prion formation or Sup35's translation termination activity, but which stabilizes prion fibers (Figure A.2A; (35-37)). Yeast prion domains are generally modular, meaning that they maintain prion activity when attached to other proteins (38). Because simple assays are available to detect [*PSI*⁺] formation, substitution of the prion domain of Sup35 with fragments from other prion-like proteins has been widely used to probe for prion activity (39-41). The first 40 amino acids of the Sup35 prion domain are required for prion formation, while the remainder of the prion domain, which is composed of a series of imperfect oligopeptide repeats, is predominantly involved in prion maintenance (Figure A.2A; (20, 42, 43)). Therefore, substitution of fragments in the place of the first 40 amino acids of Sup35 can be used to probe aggregation propensity of these domains, and to examine the effects of mutation on aggregation propensity (20). The core PrLDs from mutant hnRNPA2 can support prion activity when substituted in place of the first 40 amino acids of Sup35, while the wild-type prion domain cannot (18). Thus, these fusion proteins provide a convenient system for examining how amino acid sequence affects PrLD aggregation propensity.

The prion prediction algorithm PAPA predicts that within PrLDs, charged amino acids and proline should strongly inhibit prion formation, while aromatic and hydrophobic amino acids promote prion formation (28, 44). The disease-associated mutations in hnRNPA2B1 and

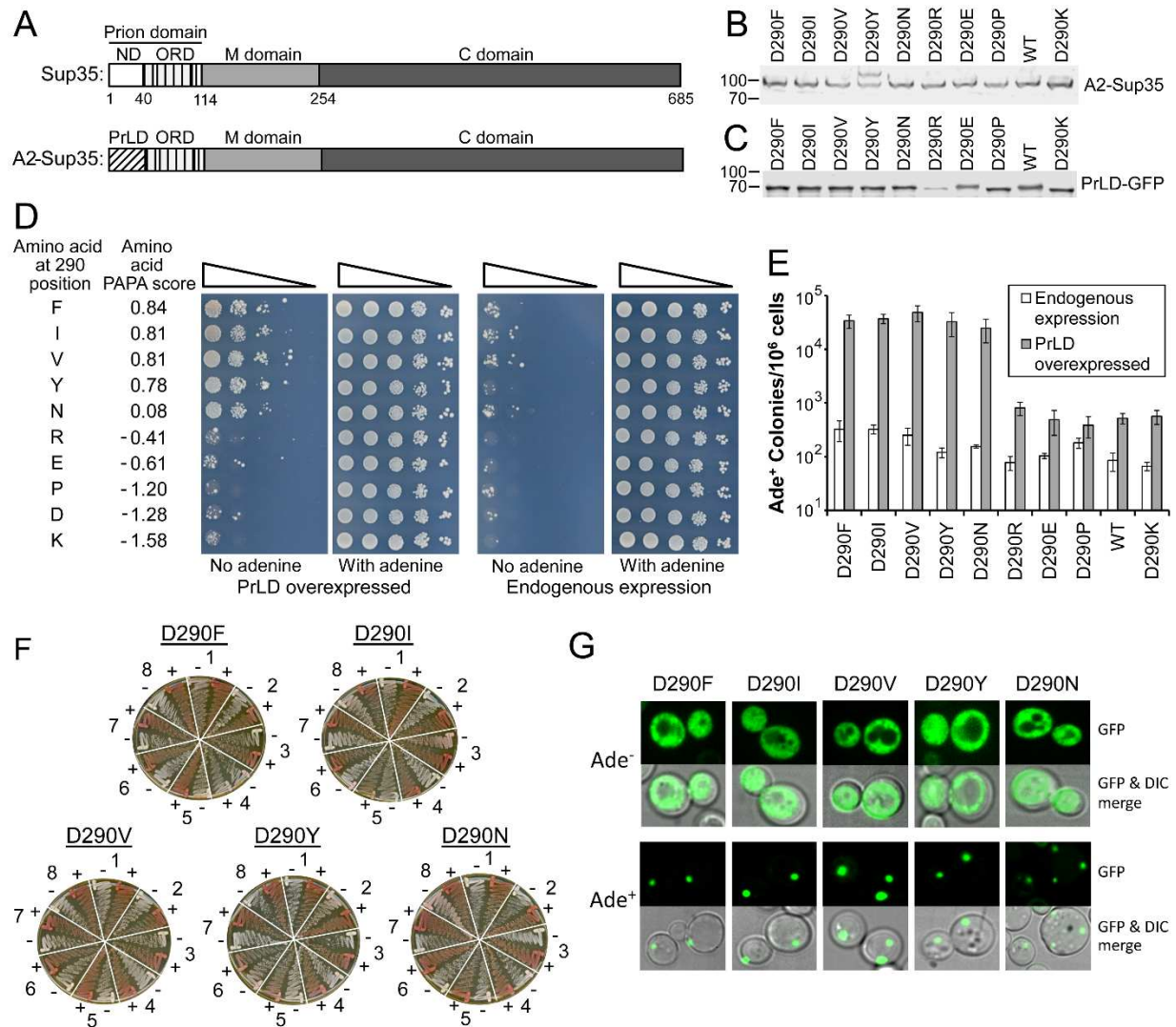


Figure A.2: PAPA accurately predicts prion-promoting mutations at the 290 position. (A) Schematic of wild-type Sup35 and the hnRNPA2-Sup35 chimeric protein (18). Sup35 contains three domains: an N-terminal prion domain (N), a highly charged middle (M) domain, and a C-terminal domain that is responsible for Sup35's translation termination function. The prion domain contains two parts: a nucleation domain (ND) that is required for prion formation, and an oligopeptide repeat domain (ORD) that is dispensable for prion nucleation, but is required for prion propagation. In the hnRNPA2-Sup35 fusion, the ND (amino acids 3-40) of Sup35 was replaced with the core PrLD (amino acids 261-303) from hnRNPA2B1. **(B)** Western blot analysis of endogenous expression of full-length wild-type (WT) and mutant hnRNPA2-Sup35 chimeric proteins, using an antibody to the Sup35 C-terminal domain. **(C)** Western blot analysis of overexpression of PrLD-GFP fusions, using an antibody to GFP. The NM domain of each hnRNPA2-Sup35 chimera was fused to GFP and expressed from the *GAL1* promoter. **(D)** Effects of different amino acids at the D290 position. [*psi*⁻] strains were generated that expressed hnRNPA2-Sup35 fusion proteins with the indicated substitution at the D290 position as the sole copy of Sup35 in the cell. The strains were transformed with either an empty vector (endogenous expression) or a plasmid expressing the matching PrLD-GFP mutant under control of the *GAL1* promoter (PrLD overexpression). Cells were grown in galactose dropout medium for 3 days, and

then 10-fold serial dilutions were plated onto medium lacking adenine to select for [*PSI*⁺] and medium containing adenine to test for cell viability. PAPA scores for each amino acid are indicated. Wild-type (D290) and D290V were previously reported (18). **(E)** Quantification of Ade⁺ colony formation. Serial dilutions of the galactose cultures from Panel D were plated onto full plates containing medium with and without adenine. The frequency of Ade⁺ colony formation was determined as the ratio of colonies formed with and without adenine. Data represent mean \pm s.d.; n \geq 3. **(F)** Curability of Ade⁺ colonies. For each mutant, eight individual Ade⁺ isolates were grown on YPD (-) or YPD plus 4 mM guanidine HCl (+). Cells were then restreaked onto YPD to test for loss of the Ade⁺ phenotype. **(G)** The Ade⁺ phenotype is associated with protein aggregation. For the indicated mutants, Ade⁻ and Ade⁺ cells were transformed with a plasmid expressing the matching PrLD-GFP mutant under control of the *GALI* promoter. Cells were grown for 1 h in galactose dropout medium and visualized by confocal microscopy.

hnRNPA1 each involve substitution of a strongly prion-inhibiting amino acid (aspartic acid) with a neutral (asparagine) or prion-promoting (valine) amino acid. We therefore hypothesized that replacing the aspartic acid with any predicted prion-promoting amino acid would have a similar effect.

To test this hypothesis, we replaced the aspartic acid at the disease-associated position in hnRNPA2 with predicted prion-promoting amino acids (phenylalanine, isoleucine, and tyrosine), a prion-neutral amino acid (asparagine), and prion-inhibiting amino acids (arginine, glutamic acid, proline, or lysine). We tested these mutations in the hnRNPA2-Sup35 chimeric protein (Figure A.2A). $[PSI^+]$ formation can be assayed by monitoring nonsense suppression of *ade2-1* allele (45). *ade2-1* mutants are unable to grow in the absence of adenine, and turn red on limited adenine due accumulation of a pigment derived from the substrate of the Ade2 protein. In $[PSI^+]$ cells, Sup35 is sequestered into prion aggregates, resulting in occasional read through of the *ade2-1* premature stop codon; therefore, $[PSI^+]$ are able to grow in the absence of adenine, and form white or pink colonies on limiting adenine. One hallmark of prion activity is that increasing protein concentration should increase the frequency of prion formation (34). We therefore monitored the frequency of Ade⁺ colony formation with and without overexpression of the matching prion domain fused to GFP.

The full-length fusions showed only modest differences in protein expression, although the D290Y mutant showed two bands, suggesting a possible post-translational modification (Figure A.2B). Likewise, all of the PrLD-GFP fusions, except the one from the D290R mutant, showed similar levels of overexpression (Figure A.2C). The single point mutations had profound effects on Ade⁺ colony formation, with the mutants showing multiple orders-of-magnitude differences upon PrLD overexpression (Figure A.2D,E). Strikingly, there was a strong correlation between the predicted effect of each mutation and the observed frequency of Ade⁺

colony formation. Each mutation predicted to enhance prion activity (D290F, I, V, Y, and N) showed statistically significant increases in Ade⁺ colony formation upon PrLD overexpression (P<0.001 by t test) relative to the wild-type fusion.

Ade⁺ colony formation can result from either prion formation or from a nonsense suppressor mutation. For each of the prion-promoting mutations (D290F, I, V, Y, and N), the fact that the frequency of Ade⁺ colony formation showed a multiple orders-of-magnitude increase upon PrLD overexpression strongly suggests that the Ade⁺ phenotype is a result of prion formation, as the frequency of DNA mutation should be insensitive to expression levels (34). Two assays were used to further confirm that these mutants were forming prions. First, we tested whether the Ade⁺ phenotype could be cured by low concentrations of guanidine HCl. Guanidine HCl cures [*PSI*⁺] (46) by inhibiting Hsp104 (47, 48). For the D290F, I, V, Y, and N mutants, almost all tested Ade⁺ colonies formed upon PrLD overexpression maintained a white phenotype in the absence of guanidine HCl, but turned red after treatment with guanidine HCl (Figure A.2F), consistent with the Ade⁺ phenotype resulting from prion formation. By contrast, none of the tested Ade⁺ colonies formed by the D290R, E, P, and D mutants were not curable by guanidine HCl (data not shown), suggesting that the Ade⁺ phenotype is likely a result of DNA mutation. The D290K mutant did have a small number of stable, curable Ade⁺ colonies, although these occurred less frequently than for any of the aggregation-promoting mutations (data not shown). Second, we used a GFP assay (49) to confirm that the fusion proteins were aggregated in curable Ade⁺ cells. When Sup35N-GFP is transiently overexpressed in [*psi*⁻] cells, it initially shows diffuse cytoplasmic localization; by contrast, in [*PSI*⁺] cells, Sup35N-GFP rapidly joins existing prion aggregates, and coalesces into foci (49). Therefore, to test for the presence of prion aggregates, we transiently overexpressed PrLD-GFP fusions in Ade⁺ and Ade⁻ cells for each

predicted prion-promoting mutant. In the Ade⁻ cells, the GFP fusions remained diffuse, while in Ade⁺ cells the fusions rapidly coalesced into foci (Figure A.2G).

Additive and Compensatory Mutations

Each of the disease-associated mutations in hnRNPA2B1 and hnRNPA1 target a highly conserved aspartic acid within a motif that is conserved across much of the hnRNP A/B family (18), suggesting that this position may be a critical determinant of aggregation propensity; however, composition-based algorithms like PAPA predict that there is nothing unique about this specific aspartic acid, and that similar mutations at other positions should exert a similar effect. The hnRNPA2 PrLD contains very few predicted prion-inhibiting amino acids, but a second aspartic acid is found at amino acid 276 (Figure A.1). An aspartic acid to valine substitution at this position also promoted Ade⁺ colony formation, although to a lesser extent than the disease-associated mutations (Figure A.3A,B). Additionally, combining mutations at both positions had an additive effect, generating a mutant that formed Ade⁺ colonies efficiently even in the absence of PrLD overexpression (Figure A.3A,B). For both the D276V mutant and the double mutant, the majority of Ade⁺ colonies formed upon PrLD overexpression were curable by guanidine HCl, consistent with prion formation (data not shown). The strong additive effect of the mutations was not due to differences in protein expression; the double mutant actually had slightly lower levels of expression for the full-length hnRNPA2-Sup35 chimeric protein, and its PrLD-GFP fusion showed similar levels of expression to the wild-type and D290V mutant (Figure A.3C). PAPA also predicts that it should also be possible to design compensatory mutations that offset the effects of the disease-associated mutations. Tyrosines are predicted to be strongly prion-promoting (44). As predicted, replacing Y283 with various prion-inhibiting amino acids (R, E, P, D, K) partially or completely offset the effects of the D290V mutation, while replacing this

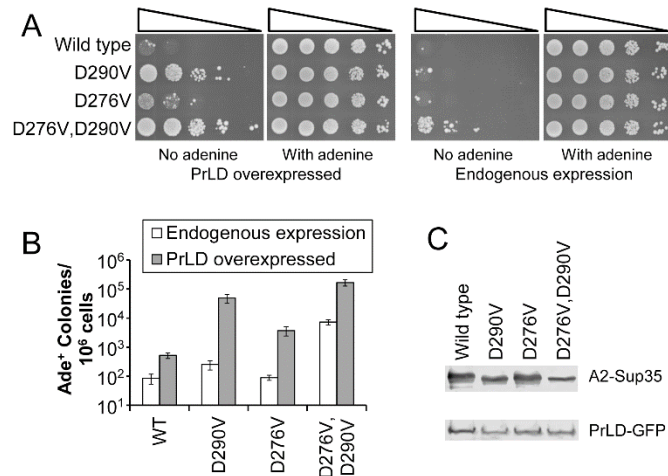


Figure A.3: Additive mutations. (A) The indicated hnRNPA2-Sup35 mutants were tested for prion formation. D276V enhances Ade⁺ colony formation, albeit less than the D290V mutation. The D276V/D290V double mutant shows substantially higher levels of Ade⁺ colony formation than D290V alone, even forming Ade⁺ colonies in the absence of PrLD overexpression. (B) Quantification of Ade⁺ colony formation. Data represent mean \pm s.d.; $n \geq 3$. (C) Western blot analysis of endogenous expression of full-length hnRNPA2-Sup35 chimeric proteins and overexpression of PrLD-GFP fusions.

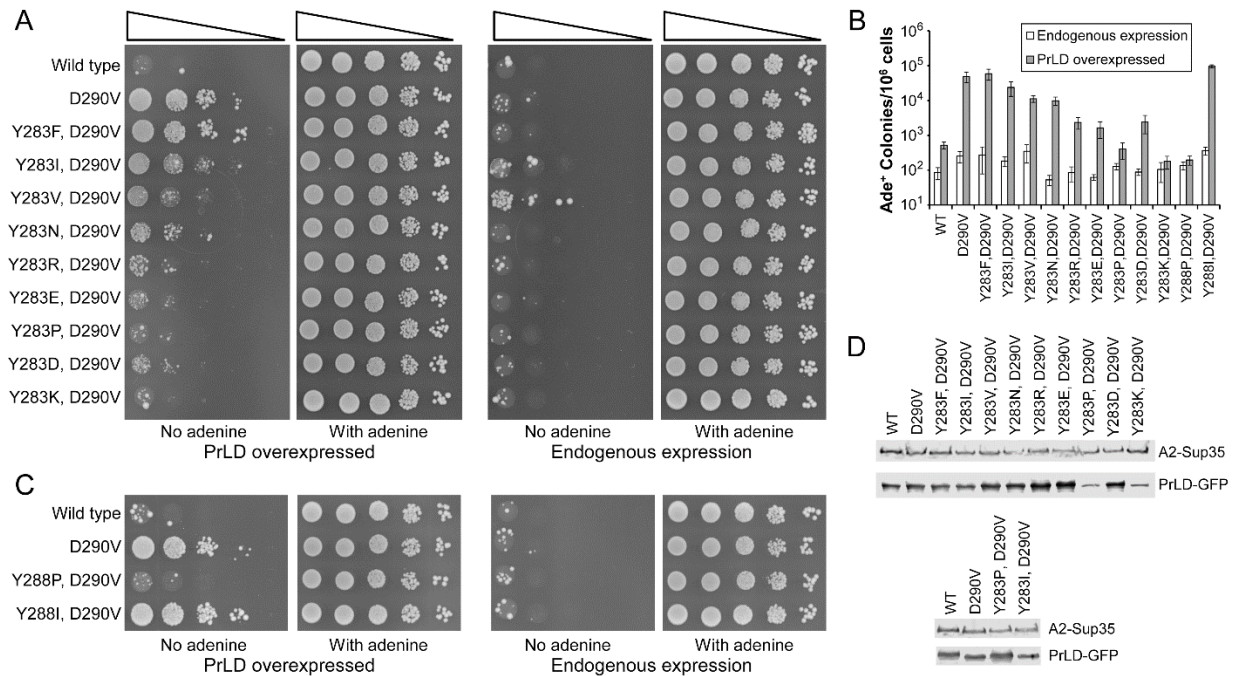


Figure A.4: Compensatory mutations. (A) Prion-inhibiting mutations effectively offset the effects of the D290V mutation. Y283 in the hnRNPA2-Sup35 (D290V) fusion was replaced with either other prion promoting amino acids (F, I, V), a neutral amino acid (N), or prion-inhibiting amino acids (R, E, P, D, K). Each of the predicted prion-inhibiting amino acids partially or completely reversed the effects of the D290V mutation. (B) Quantification of Ade⁺ colony formation. Data represent mean \pm s.d.; $n \geq 3$. (C) Y288 in the hnRNPA2-Sup35 (D290V) fusion was replaced with either a prion-inhibiting proline or a prion-promoting isoleucine. (D) Western blot analysis of endogenous expression of full-length hnRNPA2-Sup35 chimeric proteins and overexpression of PrLD-GFP fusions.

tyrosine with other prion-promoting amino acids had little effect (Figure A.4A,B). Similar results were seen with a more limited panel of mutants at a second position (Y288; Figure A.4C). For each of the mutants in which Y was replaced with a prion-promoting amino acid, the majority of Ade⁺ colonies formed upon PrLD overexpression were curable by guanidine HCl, consistent with prion formation (data not shown). Y283 and Y288 mutants showed only modest differences in expression of the full-length hnRNPA2-Sup35 chimeras, although two of the Y283 mutants (Y283P and Y283K) showed lower levels of PrLD-GFP overexpression (Figure A.4D), potentially explaining why these two mutations showed the strongest aggregation-inhibiting effects.

Zipper Segments are Neither Necessary nor Sufficient for Prion Aggregation

Each of the disease-associated mutations in hnRNPA2B1 and A1 are predicted by ZipperDB to create strong steric zipper segments (Figure A.5A,B; (18)). Each of the prion-promoting residues tested in Figure A.2 are likewise predicted to create strong zipper segments, so it is unclear whether the mutations enhance prion formation solely because of compositional effects, or due to creation of a steric zipper. The presence of a strong zipper segment is clearly not sufficient for prion formation, as the compensatory mutations in Figure A.4A prevent prion formation without disrupting the predicted zipper segment (Figure A.5C and data not shown). We designed additional mutations to test whether zipper segments are necessary for prion formation by the hnRNPA2-Sup35 chimera. Because aspartic acid is predicted by PAPA to be strongly prion-inhibiting, deletion of aspartic acid is predicted to enhance prion activity. Indeed, deletion of one or both aspartic acids in the core A2 PrLDs strongly enhanced Ade⁺ colony formation by the fusion proteins (Figure A.5D,E), and the majority of Ade⁺ colonies formed upon PrLD overexpression for these mutants were curable by guanidine HCl. This effect is not due to differences in expression level; the Δ D290 mutant showed similar levels of expression to

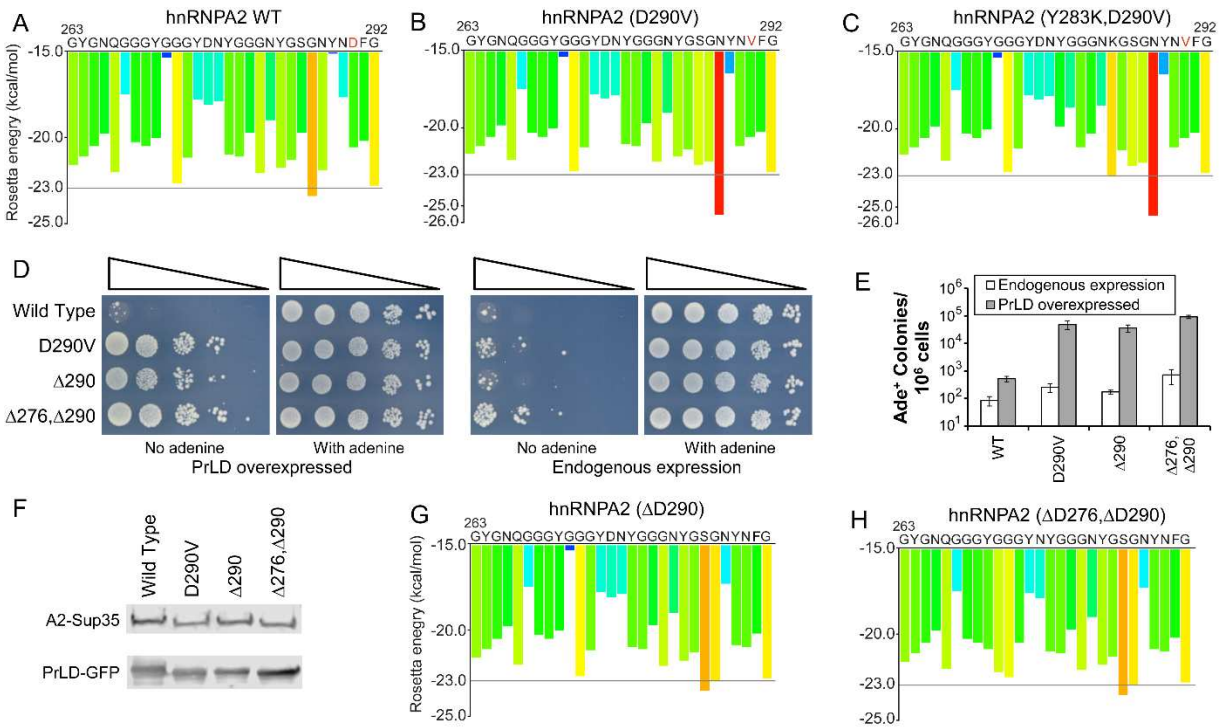


Figure A.5: Predicted strong steric zipper segments are neither necessary nor sufficient for prion activity. (A, B) ZipperDB analysis (24) of the core PrLDs of hnRNPA2 wild-type and D290V. Segments with a Rosetta energy below -23.0 kcal/mol are predicted to form steric zippers. The D290V mutation creates a strong predicted steric zipper segment from amino acids 287-292. The remainder of the core PrLD is not scored by ZipperDB due to the presence of prolines at positions 262 and 298. Adapted from Kim et al. (18). (C) The Y283K mutation blocks prion formation by the hnRNPA2-Sup35 (D290V) mutant (Figure A.4), but does not affect the predicted strong steric zipper segment. (D) Both Δ D290 and a Δ D276/ Δ D290 double mutant substantially increase Ade⁺ colony formation by the hnRNPA2-Sup35 fusion. (E) Quantification of Ade⁺ colony formation. Data represent mean \pm s.d.; $n \geq 3$. (F) Western blot analysis of endogenous expression of full-length hnRNPA2-Sup35 chimeric proteins and overexpression of PrLD-GFP fusions. (G, H) Neither the Δ D290 nor Δ D276/ Δ D290 mutations are predicted to create a strong steric zipper.

wild-type for both the full-length hnRNPA2-Sup35 chimera and the PrLD-GFP fusion (Figure A.5F), and although the double deletion showed modestly higher PrLD-GFP overexpression, this difference would be unlikely to explain the multiple orders of magnitude increase in Ade⁺ colony formation relative to the wild-type protein (Figure A.5F). However, although they substantially increased prion formation, neither of these mutations is predicted to create a strong steric zipper segment (Figure A.5G,H), indicating that strong zipper segments are neither necessary (Figure A.5D-H) or sufficient (Figure A.5C) for prion-like aggregation.

Effects of Mutations in Drosophila

For each of the mutations tested in yeast, prion activity closely correlated with PAPA predictions. Because hnRNPA2(D290V) primarily causes myopathy in humans (18), we were interested in whether our yeast results could accurately predict myopathy in a multi-cellular organism. Expression of aggregation-prone prion or prion-like proteins in muscle tissue of various model systems can cause muscle disorganization (18, 50, 51). We previously showed that when expressed in *Drosophila*, wild-type hnRNPA2 localizes to the nucleus and is predominantly detergent soluble, whereas hnRNPA2(D290V) forms cytoplasmic inclusions, is largely detergent insoluble, and leads to muscle degeneration (Figure A.6; (18)). Similar results were observed with two other antibodies: DP3B3 with untagged hnRNPA2, and anti-Flag antibody with Flag-tagged hnRNPA2 (data not shown). The cytoplasmic inclusions formed by hnRNPA2(D290V) are RNA granule assemblies, containing various RNA binding proteins (52).

To test whether other predicted prion-promoting mutations at the 290 position would also increase insolubility and promote formation of cytoplasmic foci, we expressed the hnRNPA2(D290F) in *Drosophila*. While there was less RIPA-insoluble (urea-soluble) protein for hnRNPA2(D290F) than for hnRNPA2(D290V) (Figure A.6B), hnRNPA2(D290F) was nevertheless largely RIPA-insoluble (Figure A.6C); interestingly it predominantly formed

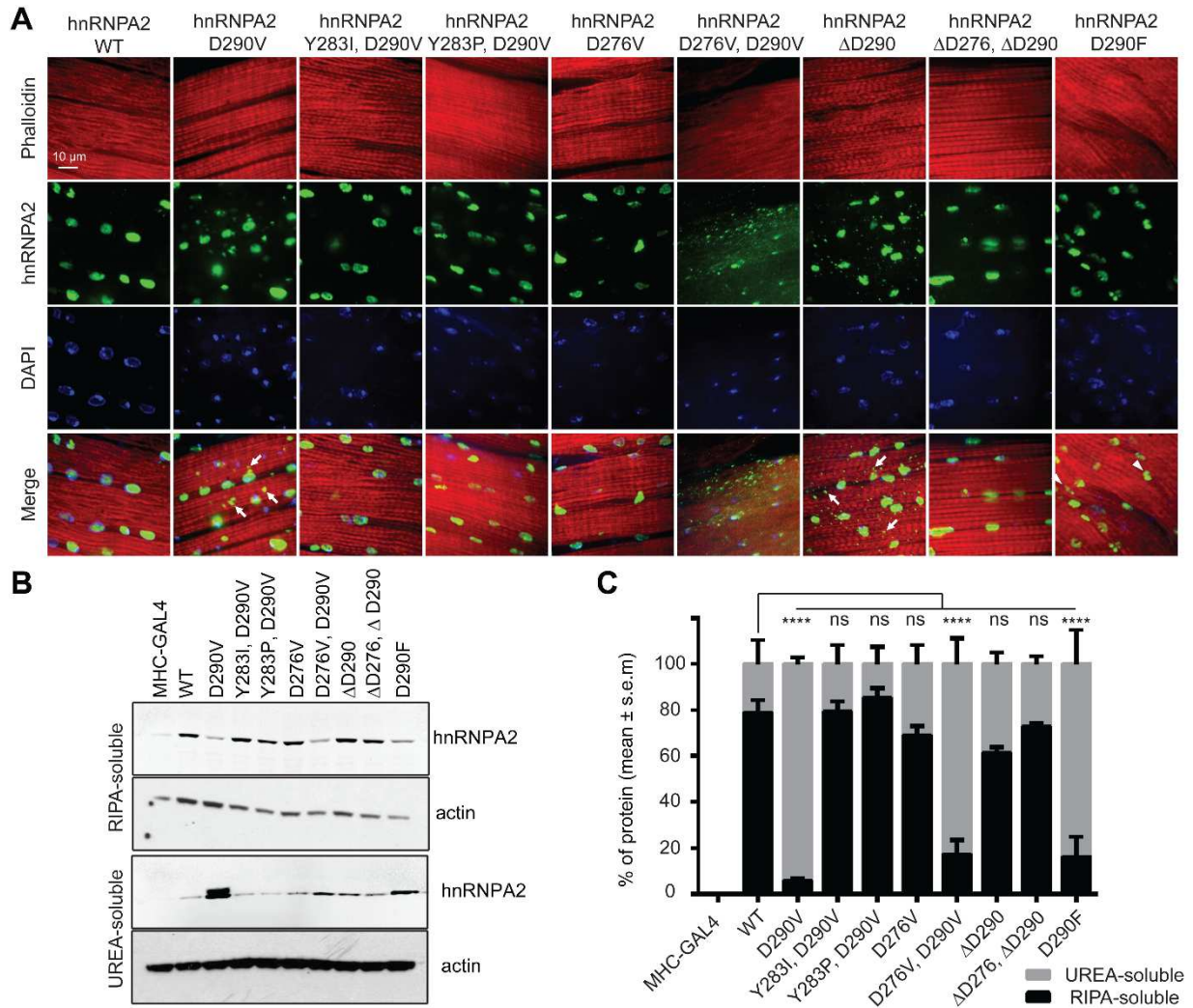


Figure A.6: Effects of mutations in *Drosophila*. (A) Adult fly thoraces were stained with anti-hnRNPA2B1 (green), Texas Red-X phalloidin (red) and DAPI (blue). Wild type hnRNPA2 localizes exclusively to the nuclei, whereas the D290V mutant also forms cytoplasmic foci. The other mutants show a range of localization patterns, including much more substantial cytoplasmic foci for the D276V, D290V double mutant. Examples of cytoplasmic and nuclear foci are indicated with arrows and arrow heads, respectively. (B) Thoraces of adult flies were dissected and sequential extractions were performed to examine the solubility profile of hnRNPA2. (C) Quantification of the blot shown in (B). Data represent mean \pm s.e.m.; $n = 3$; **** $P < 0.0001$; two-way ANOVA test with Bonferroni's post hoc test.

nuclear foci, rather than the cytoplasmic foci seen for hnRNPA2(D290V) (Figure A.6A). To test whether mutations at other sites would mimic the D290V mutation, we expressed the hnRNPA2(D276V) and the hnRNPA2(D276V,D290V) mutants. As in yeast, the D276V mutation had a smaller effect than the D290V mutation. D276V slightly increased the fraction of detergent insoluble protein compared to wild-type hnRNPA2, although this increase was not statistically significant (Figure A.6C). As in yeast, the double mutant had a strongly additive effect. The fraction of insoluble protein was actually slightly lower than for the hnRNPA2(D290V), likely because hnRNPA2(D290V) had higher protein levels, and was already almost entirely insoluble (Figure A.6C); however, the double mutant had a much more dramatic immunohistological phenotype, with many small foci throughout the cytoplasm and nucleus (Figure A.6A). Additionally, it showed clear disruption of muscle fibers, seen as a loss of the regular striations normally observed with phalloidin staining of healthy muscle (Figure A.6A).

Two of the mutations showed different behavior in yeast and *Drosophila*. As expected, the predicted prion-inhibiting Y283P mutation largely offset the effect of the D290V mutation, restoring solubility and nuclear localization (Figure A.6). However, the control Y283I mutation, which had little effect in yeast (Figure A.4A), also offset the effect of the D290V mutation in *Drosophila* (Figure A.6). As in yeast, the Δ D290 mutation decreased solubility of the protein in *Drosophila*, albeit not statistically significantly (Figure A.6C), and caused formation of cytoplasmic inclusions (Figure A.6A); however, the hnRNPA2(Δ D276, Δ D290) double mutant actually appeared to be more soluble.

One other striking difference was observed between hnRNPA2(D290V) and all other mutants tested: only hnRNPA2(D290V) showed two bands on the western blot, likely reflecting an uncharacterized post-translational modification. The significance of this second band is

unclear; given that it was not observed in the hnRNPA2(D276V,D290V) double mutant, clearly it is not required for insolubility, mislocalization, or muscle pathology.

In Vitro Analysis of Mutants

Prediction algorithms like PAPA are generally designed to predict the intrinsic aggregation propensity of peptides or proteins. However, mutations can influence aggregation by affecting activities other than intrinsic aggregation propensity, including: altering interactions with other cellular factors or with other parts of the protein; changing expression levels or protein stability; or altering localization. For the mutants that showed divergent behavior in yeast and *Drosophila*, we hypothesized that this divergent behavior likely reflected effects of the mutation beyond intrinsic aggregation propensity. To test this hypothesis, we utilized an *in vitro* aggregation assay to examine the intrinsic aggregation propensity of these mutants in the absence of other cellular factors.

We generated 35-amino acid peptides from the core PrLDs. We tested each for amyloid aggregation using thioflavin T, a dye that fluoresces upon interaction with amyloid fibrils, but not soluble proteins or amorphous aggregates (53). Most amyloid-forming proteins show sigmoidal aggregation kinetics, with a lag time, followed by a growth phase in which there is a rapid increase in aggregation, and then a plateau, as soluble material is exhausted. For each protein where the yeast and *Drosophila* results diverged, the *in vitro* aggregation kinetics mimicked the yeast results and PAPA predictions; higher frequencies of prion formation in yeast correlated with shorter lag times and a steeper growth phase. Specifically, the wild-type protein showed very slow aggregation kinetics, with a lag time of approximately 25 h (Figure A.7A). The D290V mutation substantially accelerated aggregation, shortening the lag phase to about 4 h (Figure A.7B). The Δ D290 likewise showed accelerated aggregation, which was further enhanced in the Δ D276/ Δ D290 double mutant (Figure 2.7A). The Y283P mutant was largely

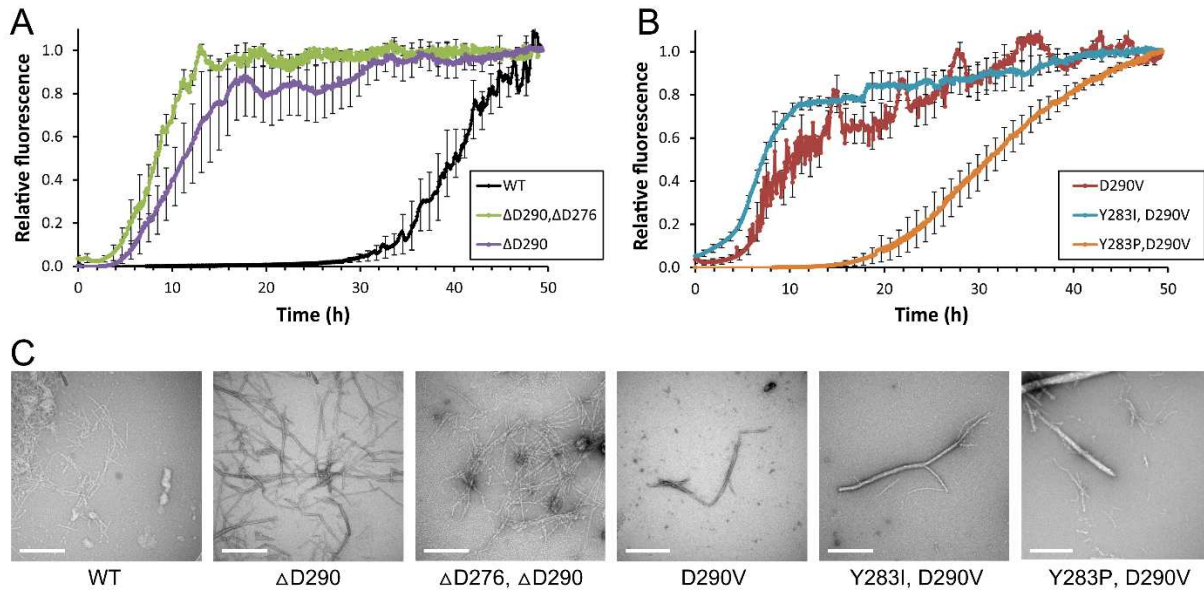


Figure A.7: In vitro amyloid formation by hnRNPA2 mutants. (A, B) Synthetic 35-amino-acid peptides from the hnRNPA2 core PrLD were generated with the indicated mutations. Peptides were resuspended under denaturing conditions, and then diluted to initiate amyloid formation. Reactions were incubated at room temperature with intermittent shaking. Amyloid formation was monitored by thioflavin T fluorescence. Data represent the mean \pm s.e.m., with error bars shown for every 10th data point; $n = 3$. (C) Electron micrographs of amyloid formation assays after 48 h. Scale bars, 500 nm

able to offset the aggregation promoting effect of D290V, while the more conservative Y283I mutation had little effect (Figure A.7B). In all cases, the increase in thioflavin T signal was associated with fiber formation (Figure A.7C).

Collectively, these results indicate that both the PAPA prediction algorithm and yeast fusion system can accurately predict the effects of mutations on the intrinsic aggregation propensity of peptides, but that intrinsic aggregation propensity is an imperfect predictor of *in vivo* aggregation of the peptides in the context of their respective full-length proteins.

Discussion

Although prion formation in humans is generally thought of as pathogenic, an emerging theory suggests that many prion-like domains may be evolved to form weak or transient interactions that mediate the formation of membrane-less organelles (9). For example, P bodies and stress granules are two types of RNA-protein assemblies that regulate translation and mRNA turnover. The prion-like domain of TIA-1 helps mediate the formation of stress granules (12), and in yeast, the prion-like domain of Lsm4 is involved in P body formation (11). This suggests the intriguing hypothesis that mutations may cause disease by disrupting the dynamics of these assemblies.

However, the exact relationship between PrLD aggregation propensity and disease is unclear. The cytoplasmic inclusions seen for disease-associated hnRNPA2B1 and A1 mutations are not just simple aggregates of these proteins, but instead are RNA granules (52). Therefore, these complex structures are normally under regulatory control, so it is unclear whether simple increases in aggregation propensity are sufficient to cause disease. Examining this question is challenging, as our understanding of these diseases is currently based on a limited set of mutations. For example, only one disease-associated mutation has been characterized in hnRNPA2B1. Furthermore, targeted mutations to investigate the role of PrLD aggregation in protein function and pathology have often involved dramatic changes to protein sequence, such

as deletion or replacement of the entire PrLD, so these mutations likely have effects beyond just changing aggregation propensity.

By contrast, we were able to cause profound changes in the aggregation propensity of hnRNPA2 with just single or double point mutations. This ability to rationally design more subtle mutations to alter aggregation propensity will provide a powerful tool to explore the role of PrLDs in functional and pathological aggregation. While not all mutations behaved as expected in *Drosophila* in the context of full-length hnRNPA2, most did, suggesting that it is relatively simple to design mutations to modulate aggregation propensity. This will facilitate experiments both to explore the normal role of functional aggregation and to test whether increasing aggregation propensity disrupts the dynamics of these aggregates and leads to disease.

Our success rate in predicting prion aggregation in yeast was surprising. We designed 13 mutations expected to increase prion formation relative to the wild-type hnRNPA2-Sup35 fusion (D290F, I, Y, and N; D290V paired with Y283F, I, V, or N; D290V paired with Y288I; D276V; D276V,D290V; Δ D290; and Δ D276, Δ D290) and 10 mutations expected to decrease prion activity relative to the D290V mutant (D290R, E, P, and K; D290V paired with Y283R, E, P, D, K; and D290V paired with Y288P). Numerous factors can affect Ade⁺ colony formation in our assay: our mutants showed subtle differences in protein expression, which will influence prion formation; for wild-type Sup35, many prions formed are toxic to cells (54), so the fraction of toxic prions for a given mutant will affect Ade⁺ colony formation; the assay may not detect weak, poorly propagating prions; and interactions with other cellular proteins may influence prion activity. Despite all of these potentially confounding factors, we accurately predicted the direction of the effect relative to the wild-type or D290V reference for all 23 mutations tested. Although the strength of the effect of each mutation was not perfectly predictable, this

remarkable prediction success despite the limitations of the assay demonstrates that even single point mutations can have a profound effect on aggregation propensity.

Nevertheless, our experiments also highlight remaining challenges for predicting the effects of mutations. In the past decade, mutations in numerous PrLD-containing RNA binding proteins, including FUS (55, 56), TDP-43 (57), TAF15 (58), EWSR1 (59), hnRNPD (60), hnRNPA1 (18), and hnRNPA2 (18) have been linked to various degenerative diseases. While some algorithms have proven successful at identifying the PrLDs in these proteins (7), predicting the effects of mutations has proven more difficult (6, 7). One issue is that while mutations in these proteins appear to cause disease by disrupting of RNA homeostasis, increasing the aggregation propensity of these RNA binding proteins is just one of many mechanisms by which RNA homeostasis could be disrupted. For example, for TDP-43, a subset of disease-associated mutations do not cause a detectable increase in aggregation propensity (61). Likewise, for FUS and hnRNPA1, some of the disease-associated mutations are found in a predicted nuclear localization signal (62, 63), so may lead to the formation of cytoplasmic inclusions by disrupting nuclear localization rather than directly increasing aggregation propensity.

Our results suggest an additional challenge in predicting the effects of mutations: while we clearly have made substantial strides in predicting the effects of mutations on the aggregation propensity of isolated PrLDs, this was an imperfect predictor of foci formation by the full-length protein in *Drosophila*. The reasons for this discrepancy are unclear. For natively folded proteins, native state stability can be a critical determinant of aggregation propensity (64), so the intrinsic aggregation propensity of a protein (i.e., the propensity of the protein to form aggregates from a denatured state) is an imperfect predictor of the aggregation propensity of the native protein. However, the hnRNPA2 PrLD is predicted to be intrinsically disordered, so native state stability should be less of an issue. For a disordered protein, other factors, including changes in

localization, interactions with other proteins or nucleic acids, post-translational modifications, or expression levels could indirectly affect aggregation propensity. Because RNA granule formation is a highly regulated process, interactions of a mutation with the normal regulatory machinery may influence the effect of a mutation, potentially resulting in the observed disconnect between intrinsic aggregation propensity and observed foci formation and insolubility in cells. Examining these outliers may ultimately provide a more complete understanding of the factors that affect pathological protein aggregation. Furthermore, it remains to be determined whether intrinsic aggregation propensity (in yeast and *in vitro*) or foci formation in *Drosophila* is more predictive of disease.

These experiments also highlight the risk of using small isolated peptides to examine aggregation propensity. This was seen at two levels: as discussed above, the peptides tested *in vitro* and in yeast were imperfect predictors of the behavior of full-length proteins; and ZipperDB, which was developed from analysis of 6-amino-acid peptides, was an imperfect predictor of longer ~35-amino-acid peptides. ZipperDB utilizes structure-based prediction, threading sequences into the crystal structure of a 6-amino-acid peptide in an amyloid-like conformation (24). There are two main types of interactions that stabilize the structure: in-register parallel β -sheet interactions that run the length of the amyloid fibril, and steric zipper packing interactions between β -sheets. Because the structure is based on a 6-amino-acid peptide, the predicted steric zipper interactions are intermolecular, between two identical peptides. Therefore, when ZipperDB predicts the ability to form steric zippers, it is essentially predicting self-complementarity of a peptide. However, in the context of longer peptides, steric zipper interactions may be intramolecular, between different segments of a single peptide (65). Such intramolecular zippers are seen in a recent high-resolution structure of amyloid- β fibers (66, 67). These intramolecular zippers are currently not predicted by ZipperDB, potentially explaining

why it inaccurately predicts some of the peptides here. However, it should be noted that while strong predicted zipper segments were clearly not sufficient for efficient aggregation in yeast (Figure A.4, 5), in *Drosophila* (Figure A.6), or *in vitro* (Figure A.7), whether they are necessary in *Drosophila* is less clear. The high aggregation propensity of the $\Delta D276/\Delta D290$ double mutant *in vitro* (Figure A.7) and in yeast (Figure A.5) clearly shows that a strong predicted steric zipper is not necessary in these contexts; however, this double mutant showed low aggregation propensity in *Drosophila* (Figure A.6), so we cannot rule out the possibility that a strong zipper segment is important in the context of the full-length protein.

REFERENCES

1. Kisilevsky R, Fraser PE (1997) A beta amyloidogenesis: unique, or variation on a systemic theme? *Crit Rev Biochem Mol Biol* 32(5):361-404.
2. Sipe JD, Cohen AS (2000) Review: history of the amyloid fibril. *J Struct Biol* 130(2-3):88-98.
3. Liebman SW, Chernoff YO (2012) Prions in yeast. *Genetics* 191(4):1041-72.
4. Wickner RB, Shewmaker FP, Bateman DA, Edskes HK, Gorkovskiy A, Dayani Y, et al (2015) Yeast prions: structure, biology, and prion-handling systems. *Microbiol Mol Biol Rev* 79(1):1-17.
5. Du Z (2011) The complexity and implications of yeast prion domains. *Prion* 5(4):311-6.
6. Cascarina SM, Ross ED (2014) Yeast prions and human prion-like proteins: sequence features and prediction methods. *Cell Mol Life Sci* 71(11):2047-2063.
7. King OD, Gitler AD, Shorter J (2012) The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res* 1462:62-80.
8. Oldfield CJ, Dunker AK (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu Rev Biochem* 83:553-84.
9. Ramaswami M, Taylor JP, Parker R (2013) Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders. *Cell* 154(4):727-36.
10. Wolozin B (2012) Regulated protein aggregation: stress granules and neurodegeneration. *Mol Neurodegener* 7:56.
11. Decker CJ, Teixeira D, Parker R (2007) Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *J Cell Biol* 179(3):437-49.
12. Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM, et al (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15(12):5383-98.
13. Reijns MAM, Alexander RD, Spiller MP, Beggs JD (2008) A role for Q/N-rich aggregation-prone regions in P-body localization. *J Cell Sci* 121(15):2463-72.
14. Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, et al (2012) Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149(4):753-67.
15. Lin Y, Protter DS, Rosen MK, Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60(2):208-19.
16. Molliex A, Temirov J, Lee J, Coughlin M, Kanagaraj AP, Kim HJ, et al (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163(1):123-33.
17. Patel A, Lee HO, Jawerth L, Maharana S, Jahnelt M, Hein MY, et al (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* 162(5):1066-77.
18. Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, et al (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495(7442):467-73.
19. March ZM, King OD, Shorter J (2016) Prion-like domains as epigenetic regulators, scaffolds for subcellular organization, and drivers of neurodegenerative disease. *Brain Res* 1647:9-18.

20. Osherovich LZ, Cox BS, Tuite MF, Weissman JS (2004) Dissection and design of yeast prions. *PLoS Biol* 2(4):E86.
21. Toombs JA, Petri M, Paul KR, Kan GY, Ben-Hur A, Ross ED (2012) De novo design of synthetic prion domains. *Proc Natl Acad Sci USA* 109(17):6519-24.
22. Toombs JA, McCarty BR, Ross ED (2010) Compositional determinants of prion formation in yeast. *Mol Cell Biol* 30(1):319-32.
23. Ross ED, Maclea KS, Anderson C, Ben-Hur A (2013) A bioinformatics method for identifying Q/N-rich prion-like domains in proteins. *Methods Mol Biol* 1017:219-28.
24. Goldschmidt L, Teng PK, Riek R, Eisenberg D (2010) Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci USA* 107(8):3487-92.
25. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riek C, Grothe R, et al (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435(7043):773-8.
26. Sherman F (1991) Getting started with yeast. *Methods Enzymol.* 194:3-21.
27. MacLea KS, Paul KR, Ben-Musa Z, Waechter A, Shattuck JE, Gruca M, et al (2015) Distinct amino acid compositional requirements for formation and maintenance of the [PSI(+)] prion in yeast. *Mol Cell Biol* 35(5):899-911.
28. Gonzalez Nelson AC, Paul KR, Petri M, Flores N, Rogge RA, Cascarina SM, et al (2014) Increasing prion propensity by hydrophobic insertion. *PloS One* 9(2):e89286.
29. Paul KR, Hendrich CG, Waechter A, Harman MR, Ross ED (2015) Generating new prions by targeted mutation or segment duplication. *Proc Natl Acad Sci USA* 112(28):8584-9.
30. Ross ED, Edskes HK, Terry MJ, Wickner RB (2005) Primary sequence independence for prion formation. *Proc Natl Acad Sci USA* 102(36):12825-30.
31. Fredrickson EK, Gallagher PS, Clowes Candadai SV, Gardner RG (2013) Substrate recognition in nuclear protein quality control degradation is governed by exposed hydrophobicity that correlates with aggregation and insolubility. *J Biol Chem* 288(9):6130-9.
32. Bagriantsev SN, Kushnirov VV, Liebman SW (2006) Analysis of amyloid aggregates using agarose gel electrophoresis. *Methods Enzymol* 412:33-48.
33. Chernoff YO, Lindquist SL, Ono B, Inge-Vechtomov SG, Liebman SW (1995) Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi+]. *Science* 268(5212):880-4.
34. Wickner RB (1994) [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* 264(5158):566-9.
35. Liu JJ, Sondheimer N, Lindquist SL (2002) Changes in the middle region of Sup35 profoundly alter the nature of epigenetic inheritance for the yeast prion [PSI+]. *Proc Natl Acad Sci USA* 99 Suppl 4:16446-53.
36. Ter-Avanesyan MD, Dagkesamanskaya AR, Kushnirov VV, Smirnov VN (1994) The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi+] in the yeast *Saccharomyces cerevisiae*. *Genetics* 137(3):671-6.
37. Ter-Avanesyan MD, Kushnirov VV, Dagkesamanskaya AR, Didichenko SA, Chernoff YO, Inge-Vechtomov SG, et al (1997) Deletion analysis of the SUP35 gene of the yeast *Saccharomyces cerevisiae* reveals two non-overlapping functional regions in the encoded protein. *Mol Microbiol* (5):683-92.
38. Li L, Lindquist S (2000) Creating a protein-based element of inheritance. *Science* 287(5453):661-4.

39. Alberti S, Halfmann R, King O, Kapila A, Lindquist S (2009) A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins. *Cell* 137(1):146-58.
40. Osherovich LZ, Weissman JS (2001) Multiple Gln/Asn-rich prion domains confer susceptibility to induction of the yeast [PSI(+)] prion. *Cell* 106(2):183-94.
41. Sondheimer N, Lindquist S (2000) Rnq1: an epigenetic modifier of protein function in yeast. *Mol Cell* 5(1):163-72.
42. DePace AH, Santoso A, Hillner P, Weissman JS (1998) A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 93(7):1241-52.
43. Toombs JA, Liss NM, Cobble KR, Ben-Musa Z, Ross ED (2011) [PSI+] maintenance is dependent on the composition, not primary sequence, of the oligopeptide repeat domain. *PLoS One* 6(7):e21953.
44. Ross ED, Toombs JA (2010) The effects of amino acid composition on yeast prion formation and prion domain interactions. *Prion* 4(2):60-5.
45. Cox BS (1965) PSI, a cytoplasmic suppressor of super-suppressor in yeast. *Heredity* 26:211-32.
46. Tuite MF, Mundy CR, Cox BS (1981) Agents that cause a high frequency of genetic change from [psi+] to [psi-] in *Saccharomyces cerevisiae*. *Genetics* 98(4):691-711.
47. Ferreira PC, Ness F, Edwards SR, Cox BS, Tuite MF (2001) The elimination of the yeast [PSI+] prion by guanidine hydrochloride is the result of Hsp104 inactivation. *Mol Microbiol* 40(6):1357-69.
48. Jung G, Masison DC (2001) Guanidine hydrochloride inhibits Hsp104 activity in vivo: a possible explanation for its effect in curing yeast prions. *Curr Microbiol* 43(1):7-10.
49. Patino MM, Liu JJ, Glover JR, Lindquist S (1996) Support for the prion hypothesis for inheritance of a phenotypic trait in yeast. *Science* 273(5275):622-6.
50. Nussbaum-Krammer CI, Park KW, Li L, Melki R, Morimoto RI (2013) Spreading of a prion domain from cell-to-cell by vesicular transport in *Caenorhabditis elegans*. *PLoS Genet* 9(3):e1003351.
51. Park KW, Li L (2008) Cytoplasmic expression of mouse prion protein causes severe toxicity in *Caenorhabditis elegans*. *Biochem Biophys Res Commun* 372(4):697-702.
52. Li S, Zhang P, Freibaum BD, Kim NC, Kolaitis RM, Molliex A, et al (2016) Genetic interaction of hnRNPA2B1 and DNAJB6 in a *Drosophila* model of multisystem proteinopathy. *Hum Mol Gen* 25(5):936-50.
53. LeVine H (1999) Quantification of beta-sheet amyloid fibril structures with thioflavin T. *Methods Enzymol* 309:274-84.
54. McGlinchey RP, Kryndushkin D, Wickner RB (2011) Suicidal [PSI+] is a lethal yeast prion. *Proc Natl Acad Sci USA* 108(13):5337-41.
55. Kwiatkowski TJ, Jr., Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, Russ C, et al (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323(5918):1205-8.
56. Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323(5918):1208-11.
57. Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, et al (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314(5796):130-3.

58. Couthouis J, Hart MP, Shorter J, Dejesus-Hernandez M, Erion R, Oristano R, et al (2011) A yeast functional screen predicts new candidate ALS disease genes. *Proc Natl Acad Sci USA* 108:20881-90.
59. Couthouis J, Hart MP, Erion R, King OD, Diaz Z, Nakaya T, et al (2012) Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Hum Mol Gen* 21(13):2899-911.
60. Vieira NM, Naslavsky MS, Licinio L, Kok F, Schlesinger D, Vainzof M, et al (2014) A defect in the RNA-processing protein HNRPDL causes limb-girdle muscular dystrophy 1G (LGMD1G). *Hum Mol Gen* 23(15):4103-10.
61. Johnson BS, Snead D, Lee JJ, McCaffery JM, Shorter J, Gitler AD (2009) TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J Biol Chem* 284(30):20329-39.
62. Da Cruz S, Cleveland DW (2011) Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr Opin Neurobiol* 21(6):904-19.
63. Liu Q, Shu S, Wang RR, Liu F, Cui B, Guo XN, et al (2016) Whole-exome sequencing identifies a missense mutation in hnRNPA1 in a family with flail arm ALS. *Neurology* 87(17):1763-9.
64. Kelly JW (1998) The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol.* 8(1):101-6.
65. Paul KR, Ross ED (2015) Controlling the prion propensity of glutamine/asparagine-rich proteins. *Prion* 9(5):347-54.
66. Eisenberg DS, Sawaya MR (2016) Implications for Alzheimer's disease of an atomic resolution structure of amyloid-beta(1-42) fibrils. *Proc Natl Acad Sci USA* 113(34):9398-400.
67. Walti MA, Ravotti F, Arai H, Glabe CG, Wall JS, Bockmann A, et al (2016) Atomic-resolution structure of a disease-relevant A beta(1-42) amyloid fibril. *Proc Natl Acad Sci USA* 113(34):E4976-E84.