COMPOSITION AND DYNAMICS OF STRESS GRANULES

by

TYLER MATHENY

B.S., University of Tulsa, 2013

B.A., University of Tulsa, 2013

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Biochemistry

2018

This thesis entitled:
Composition and Dynamics of Stress Granules
written by Tyler Matheny
has been approved for the Department of Biochemistry

Roy Parker

Dylan Taatjes

Date:

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Tyler Matheny (Ph.D., Biochemistry)
Composition and Dynamics of Stress Granules
Thesis directed by Professor Roy Parker

ABSTRACT

Stress granules are non-membranous assemblies of RNA and protein that form during conditions in which translation initiation is limited. Stress granules are of interest due to their potential role in the regulation of gene expression and their relation to other cytoplasmic RNP granules, such as p-bodies, germ granules, and neuronal granules. Additionally, stress granules are related to pathological aggregates observed in neurodegenerative diseases since (1) they share many protein components with these aggregates and (2) mutations that promote stress granule assembly or inhibit their clearance are causative in neurodegenerative disease.

The goals of this thesis are (1) to understand the dynamics of stress granule assembly and disassembly, (2) to elucidate the RNA composition of stress granules, and (3) to understand how RNAs are targeted to stress granules. Herein we make several important observations about the composition and dynamics of stress granules. i) Stress granules assemble through a multistep process in which stable substructures, termed cores, form first and coalesce into larger stress granules. ii) Stress granules are a composite of many mRNAs and lncRNAs, and transcripts vary in their degree of recruitment from <1% to >95% enriched in stress granules. iii) RNAs partition to stress

granules on the basis of their length and translation status, with long poorly translated transcripts showing enhanced partitioning to stress granules. iv) The RNA composition of stress granules is largely conserved from yeast to humans, between different stresses, and between different protein cores. v) RNA can partition to stress granules independently of individual RBPs, suggesting that RNA partitioning to stress granules is either based on RNA-RNA interactions, or through the summation of the effects of many different RBPs.

Our findings also pave the way for testing stress granule function by providing a list of stress granule enriched RNAs which will presumably be most affected by stress granules. Finally, our results also have implications outside of stress granule biology and provide insights into how RNAs might partition into other RNP granules.

ACKNOWLEDGEMENTS

I believe that the act of thinking basically boils down to having a conversation with yourself about a given subject. The more perspectives one draws upon for this internal conversation and the higher the quality of these perspectives, the more robust the thinking and, therefore, the better the ideas generated. The perspectives we draw upon are usually voiced by the people with whom we surround ourselves. Thus, my thoughts are not strictly my own, but rather a summation of all the perspectives of the company I keep.

Therefore, I owe Roy and the members of the Parker Lab a great debt of gratitude for helping me become the scientist and person I am today. I am deeply grateful to have worked in such an amazing lab with such talented scientists and all-around good people. I will miss each and every one of you and thank you for all of your help throughout the years.

I would also like to thank my parents, Rick and Angela, for always encouraging me to chase my dreams. They have been very supportive throughout my time in grad school. I am fortunate to have such amazing parents.

Last, but certainly not least, I would like to thank my wife, Sarah, for her love and support throughout graduate school. If the only thing I had accomplished in graduate school was meeting you and marrying you, I would have considered my graduate school experience a success. The rest of this thesis is just icing on the cake.

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

PRINCIPLES AND SPECIFICITY OF RNA ACCUMULATION IN RNP GRANULES

SUMMARY

Cells contain numerous non-membrane-bound organelles that contain both RNA and protein, termed ribonucleoprotein (RNP) granules. Stress granules, P-bodies, neuronal RNA granules, and germ granules are various examples of cytosolic RNP granules, whereas paraspeckles, nuclear speckles, the nucleolus, and Cajal bodies represent examples of nuclear RNP granules. Recently, the transcriptome of neuronal RNA granules (Fatimy et al., 2016), P-bodies (Hubstenberger et al., 2017), and stress granules (Khong et al., 2017; Namkoong et al. 2018) have been elucidated through various granule purification strategies and RNA sequencing. Herein we review these three classes of cytoplasmic granules, as well as other RNP granules, and compare and contrast the physical basis for RNA partitioning to these various granules. From this comparison we find that translation, length, nucleotide composition, RNA-RNA interactions, and association with RNA binding proteins (RBPs) are parameters that can act to define which RNAs partition to granules. We also examine the relative contribution of RNA-RNA interactions and RBP interactions and suggest that these two types of interactions work synergistically to define the RNA and protein composition of stress granules. Finally, we explore which factors may explain differential partitioning of transcripts between P-bodies and stress granules. We suggest that the underlying

principles for partitioning to these assemblies may be the same and that translation may, in part, dictate specificity by defining the available substrate pool.

INTRODUCTION

Ribonucleoprotein (RNP) granules are non-membranous assemblies of RNA and protein. There are many different types of RNA granules including stress granules, P-bodies, germ granules, and neuronal RNA granules in the cytoplasm and paraspeckles, nuclear speckles, the nucleolus, and Cajal bodies in the nucleus **(summarized in Table 1)**. Nuclear RNP granules contain various types of RNAs (such as lncRNAs, snRNAs, snoRNAs, mRNAs, and rRNAs), and function in snRNP assembly, splicing, transcriptional regulation, ribosome assembly, and nuclear export of RNA (Neugebauer, 2018; Fox et al., 2018; Galganski et al., 2017; Lam and Trinkle-Mulcahy, 2015). Cytosolic RNP granules, such as P-bodies, stress granules, and neuronal RNA granules, contain translationally silent mRNAs (although they can contain lncRNAs) and are thought to serve as RNA storage hubs, or as transport granules (in the case of neuronal granules, which transport translationally silenced mRNP complexes to the synapse where they are translated).

The RNA composition of RNP granules is generally related to granule function. For example, ribosomal RNAs localize to the nucleolus, where they are assembled into ribosomal subunits (Savino et al., 2001). Similarly, snRNAs localized to paraspeckles, where they are organized into snRNPs (Fox and Lamond, 2010). Additionally, individual RNAs can be critical structural components of RNP assemblies, acting as scaffolds. For example, the lncRNA NEAT1 serves as a scaffold for assembly of paraspeckles (Bond and Fox, 2009).

Until recently, the transcripts that comprise different types of RNA granules have remained an enigma due to the lack of granule purification techniques. In the last few years, a variety of purification strategies have been devised to sequence the RNAs that localize to various cytosolic RNP granules based on differential centrifugation alone (Namkoong et al., 2018), differential centrifugation followed by immunoprecipitation or particle sorting (Jain et al., 2016, Khong et al., 2017, Hubstenberger et al., 2017), or differential migration through sucrose gradients (Fatimy et al., 2016). These techniques have allowed researchers to characterize the RNA composition of stress granules, P-bodies and neuronal granules as well as elucidate properties of RNA that may play a role in their partitioning to granules.

Herein we review the RNA composition of various granules and discuss the general principles that govern RNA recruitment to each of these assemblies. We compare these principles in order to describe general rules for RNA partitioning. Finally, we propose why P-bodies and stress granules harbor similar, yet distinct, transcriptomes and propose a new model in which translation equilibrium may serve as a major determinant in RNA partitioning.

TABLE 1.1 Summary of RNP granules and their associated RNAs

assembly	*localization*	*function(s)*	*associated RNAs*
stress granules	Cytoplasmic localization. Often in proximity of, or docked, with p-bodies	Function is still debated. Possibilities include storage and/or protection of transcripts from degradation.	Long translationally repressed mRNAs and long lincRNAs.
P-bodies	Cytoplasmic localization. Increase in numbers and size and dock with SGs during stress	Thought to play a role in storage and/or degradation.	Long AU-rich mRNAs and some lincRNAs
neuronal RNA granules	Throughout the neuronal cytoplasm, especially concentrated near the synapse	Transport of translationally repressed mRNAs to the synapse.	Long transcripts and/or transcripts with long 3' UTRs
germ granules	Cytoplasm of germ cells. Localize to posterior of embryo and establish asymmetry	Regulation of translation, mRNA localization, and mRNA stability. Important for totipotent state of germline	Maternal mRNAs.
Cajal bodies	Nuclear localization. Interchromatin space. Localize adjacent to nucleoli	snRNP assembly, regulation of splicing	Small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs)
paraspeckles	Nuclear localization. Interchromatin space. Form adjacent to paraspeckles	Retention of RNAs double stranded structures in the nucleus. Possible role in transcription	*NEAT1 and Ctn1* RNAs with double stranded regions
Nuclear speckles	Nuclear localization Interchromatin space. Form adjacent to paraspeckles	Regulation of splicing, transcription, and RNA export	*MALAT1/NEAT2* Poly-A positive RNAs (particularly those containing introns), snRNAs, 7SK RNA
nucleolus	Nuclear localization Interchromatin space.	Ribosome assembly	rRNA

RNP Background and Purification strategy

In the following section we give a brief background of stress granules, p-bodies, and neuronal granules, as well as the purification strategies utilized to isolate these various RNP granules. We also give a brief overview of the RNA composition of each of these assemblies.

Stress Granules

Stress granules were first observed in the cytoplasm of tomato plant cells after heat shock (Nover et al., 1983). In an elegantly designed experiment, the *in vitro* translation of RNA from heat shock granules (later termed stress granules) showed that RNAs encoding heat shock proteins do not localize to heat shock granules, suggesting that these assemblies contain a specific subset of mRNAs (Nover et al., 1989). Stress granules were subsequently shown to also form in mammalian cells, as well as in yeast (Kedersha et al., 1999, Buchan et al., 2008). Stress granules are thought to contain translationally silenced mRNPs since (1) stress granules contain poly-adenylated mRNAs (Kedersha et al., 1999), (2) stress granules form under conditions in which translation initiation is limited (Kedersha et al., 1999), (3) stress granules lack large ribosomal subunits (Kedersha et al., 2002, Kimball et al. 2003), (4) trapping ribosomes on transcripts with cycloheximide limits stress granule formation (Kedersha et al., 2000), (5) mRNAs with long ORFs are recruited to stress granules later than transcripts with short ORFs (Khong and Parker, 2018), and (6) reporters that localize to granules are always in a translationally repressed state (Moon et al., 2018). The prevailing

hypothesis in the field is that stress granules act as sites of RNA storage and, upon removal of stress, the transcripts within stress granules can either re-enter the translational pool, or undergo degradation (Anderson and Kedersha, 2002, Buchan and Parker 2009). However, a clear function for stress granules, and whether they are consequential to altering gene expression remains unclear.

The RNA composition of stress granules was recently characterized through stress granule core purification via differential centrifugation and immunoprecipitation followed by RNA sequencing (Khong et al., 2017). The purification strategy was confirmed by immunofluorescence for proteins identified via mass spectrometry (Jain et al. 2016) as well as single molecule fluorescent in situ hybridization (smFISH) for RNAs identified via RNA-seq (Khong et al., 2017). Additionally, others have pelleted RNA granules (without an immunoprecipitation step) and found a high degree of correlation with the stress granule transcriptome (Namkoong et al., 2018).

Analysis of the stress granule transcriptome revealed many new interesting findings. Only 9-13% of bulk RNA was found to localize to stress granules, based on quantitative estimates of the stress granule transcriptome and oligo(dT) FISH. Stress granules were also shown to contain both mRNAs, and lncRNAs; however, mRNAs are more abundant within stress granules, presumably because mRNAs are longer than lncRNAs on average (see below). Stress granules contain most RNAs, but different RNAs vary in the extent of their recruitment. Interestingly, the most abundant RNAs in stress granules were not necessarily enriched transcripts. For example, GAPDH was

identified as a depleted mRNA as only 6% of its transcripts partition to stress granules. However, since there are over 1000 copies of GAPDH per cell it is still one of the most abundant transcripts within stress granules (Khong et al., 2017). Taken together, these results suggest that stress granules are a composite of many different RNAs.

Processing bodies

Processing bodies (P-bodies) were first discovered when cells stained for XRN1 showed an unusual punctate staining pattern (Bashkirov et al., 1997). P-bodies are constitutive and contain many other components of the RNA degradation machinery (Sheth and Parker, 2003). While P-bodies are constitutively present in the cell, the number of P-bodies can increase under conditions in which translation initiation is limited or RNA degradation is inhibited (Kedersha et al., 2005, Teixiera et al., 2005). Transcripts that associate with P-bodies can either be degraded or return to translation (Brengues et al., 2007, Parker and Sheth, 2007). mRNPs that localize to P-bodies are thought to be translationally repressed as they lack the translational machinery necessary for protein synthesis (reviewed in Decker and Parker 2012, Teixiera et al., 2005). Whether P-bodies play a role in actively degrading transcripts has been a topic of debate in the field (Parker and Sheth, 2007, Decker and Parker, 2012, Hubstenberger et al 2017). On one hand, the presence of decay factors and the observation that decay intermediates localize to P-bodies, suggest that degradation can occur within P-bodies. However, RNA decay has also been shown to occur outside of P-bodies (Eulalio et al.,2007; Stadler and Muhlman, 2009; Decker et al., 2007; Decker

and Parker, 2012), and P-bodies have been proposed to be sites of mRNA storage, rather than decay (Hubstenberger et al., 2017).

Recently, P-body enriched RNAs were purified and sequenced (Hubstenberger et al., 2017). In brief, a fluorescently tagged P-body marker (GFP-LSM14A) was expressed in a HEK293T cell line. P-bodies were then purified through differential centrifugation followed by fluorescence-activated particle sorting (Hubstenberger et al. 2017). In this manner, fluorescently tagged p-bodies were separated from unlabeled heavy complexes. Proteins and RNAs that were enriched by this particle sorting methodology were then detected via mass spectrometry and RNA-seq. Novel P-body proteins and transcripts were confirmed to localize to P-bodies *in vivo* by either immunostaining or single molecule inexpensive fluorescent in situ hybridization (smiFISH) (Tsanov et al., 2016).

Over 6,000 different transcripts were found to be significantly enriched in p-bodies (Hubstenberger et al., 2017). mRNAs were found to be the most abundant type of RNA in stress granules, although some lncRNAs and other noncoding RNAs were also identified (Hubstenberger et al., 2017). Interestingly, less RNA seems to partition to p-bodies than to stress granules, which is consistent with p-bodies being much smaller than stress granules (Hubstenberger et al., 2017). For example, SPEN, a highly enriched transcript within P-bodies via RNA-seq, only shows ~15% colocalization with P-bodies *in vivo*. This amount of colocalization is much smaller than that observed in stress granules (with some stress granule-enriched transcripts > 80% colocalization)

suggesting that p-bodies (1) contain less RNA than stress granules and (2) that P-body enriched transcripts show a more limited partitioning to p-bodies compared to stress granule enriched transcripts partitioning to stress granules. Nevertheless, similar to stress granules, P-bodies are also a composite of many different RNAs.

Neuronal RNA granules

Neuronal RNA granules have long been known to transport RNA in neurons (Ainger et al., 1993). Neuronal granules harbor translationally-repressed RNAs and transport these RNAs to the synapse where they are needed for local protein synthesis (reviewed in Kiebler and Bassell, 2006). Neuronal RNA granules are thought to be storage sites for these transcripts, thereby poising these transcripts for translation (Liu-Yesucevitz et al., 2011). Protein synthesis is required to retain synaptic plasticity, thus neuronal RNA granules have major implications in memory formation.

Neuronal granules have been purified based on differential sedimentation rates of RNA granules and polysomes (Fatimy et al., 2016). In brief, RNA granules were purified from mouse brain homogenates via linear sucrose gradients. L7 (a polyribosome marker) and FMRP (a granule marker) were tracked throughout these gradients via western blot to distinguish between polyribosomes and neuronal RNA granules. In this manner, neuronal granules were separated from polysomes. Similar to stress granules and P-bodies, neuronal granules were found to contain a composite of many different mRNAs.

General Principles of RNA partitioning into RNP granules

Rather than cytoplasmic RNPs harboring a highly specific subset of transcripts, RNA partitioning to RNP granules seems to act as a gradient. Most transcripts seem to be able to localize to granules, but individual transcripts can vary in the degree to which they are recruited. This phenomena is most striking for stress granules, in which transcripts varied in their degree of recruitment, ranging from < 1% to > 95% (Khong et al. 2017). Similarly, P-bodies and neuronal granules are also a composite of many RNAs that show varying degrees of enrichment. Since RNAs can vary in their degree of recruitment to RNP granules, this led to examination of the physical basis of recruitment. Herein we review the general principles that seem to govern RNA partitioning to RNP granules **(see Table 2)**.

Length: Transcript length is the predominant metric that correlates with RNA partitioning to stress granules (Khong et al., 2017). Indeed, the average stress granule-enriched RNA was 7.1 kilobases, while the average stress granule-depleted transcript was 2.5 kilobases. Similar to stress granules, neuronal RNA granules and P-bodies are biased towards harboring long transcripts (Fatimy et al., 2016; Hubstenberger et al., 2017; Courel et al. 2018). Indeed, P-body enriched transcripts had an average length of ~ 4.4 kb and neuronal granule enriched RNAs were ~ 6 kb on average.

Longer transcripts have an increased valency and can therefore form many types of interactions. Presumably, the increased valency of long transcripts helps to form both specific interactions and non-specific interactions that may contribute to RNP

sequestration. Indeed, live cell imaging techniques reveal that long transcripts tend to form stable interactions within stress granules, while shorter transcripts tend to form more transient interactions (Moon et al., 2018). Therefore, these results suggest that the length of an mRNA is an RNP granule targeting metric.

Translation: Poorly translated transcripts also accumulate in stress granules. Three observations have shown that stress granule-localized RNAs tend to be poorly translated. (1) Comparison of the stress granule transcriptome to various data sets measuring the translation efficiency (ribosome protected reads/ total RNA reads) of mRNAs showed that mRNAs that have poor basal rates of translation tend to enrich in stress granules during stress (Khong et al., 2017; Matheny et al., 2018 (see chapter4); Sidrauski et al., 2015; Subtelny et al 2014). (2) Comparison of the stress granule transcriptome to a dataset which examined newly translated proteins during arsenite stress revealed that translated mRNAs evade sequestration into stress granules, presumably due to their association with ribosomes (Baron et al. 2018). (3) Codon composition shows a strong correlation with stress granule enrichment in yeast, with transcripts containing a large percentage of non-optimal codons (which indicates poor translation rate) enriching in stress granules (reviewed in Hanson and Coller, 2018; Khong et al., 2017).

Like stress granules, P-bodies were found to be biased towards poorly translated transcripts. Since p-bodies form during normal growth conditions, one would anticipate that p-bodies would be made up of only the poorest translators. Consistent with this

hypothesis, P-body enrichment shows a much stronger negative correlation with translation than stress granules. This observation was based (1) on comparing P-body enriched transcripts to a quantitative HEK293T proteome (Hubsenberger et al., 2017; Geiger et al 2012) and (2) ribosome profiling data obtained in HEK293T cells (Hubsenberger et al., 2017; Subtelny et al., 2014). In both cases a negative correlation was observed between P-body enrichment and translation. This observation is consistent with the hypothesis that P-bodies contain translationally silenced mRNPs and suggests that translational silencing is required for targeting to P-bodies. Moreover, P-body enriched mRNAs showed less correlation between mRNA levels and protein levels than P-body depleted transcripts, suggesting P-bodies decouple RNA and protein abundance and may play an active role in repressing translation (Hubstenberger et al., 2017). Additionally, depletion of DDX6, which suppresses P-body formation, leads to increased translation of P-body enriched RNAs (Hubstenberger et al., 2017).

Taken together these results suggest poorly translated transcripts partition to both P-bodies and stress granules and that translation probably defines the available substrate pool during RNA partitioning to RNP granules.

Nucleotide composition: The strongest determinant for RNA partitioning to P-bodies is nucleotide content. Indeed, P-body enrichment and AU content show a strong positive correlation (Courel et al., 2018). Consistent with P-bodies containing AU-rich transcripts, p-body-enriched transcripts contain a disproportionately high frequency of rare codons, which are biased to containing either an adenine or a uridine at the third

base position (Courel et al., 2018). Moreover, the absolute number of rare codons per transcript was sufficient to partition the enriched transcripts from depleted transcripts (Courel et al., 2018). Thus, AU composition and rare codon frequency are major determinants for RNA partitioning to p-bodies, which is consistent with P-body enriched RNAs being poorly translated.

AU-rich transcripts also tend to accumulate in stress granules, albeit to a lesser extent than p-bodies. Indeed, stress granule enrichment has been shown to show a negative correlation with GC content in both yeast and human cells, suggesting nucleotide composition may play a role in localization to stress granules, however we question this role later in this review (see “Differential sorting of RNA into P-bodies and stress granules”) (Khong et al., 2017).

Consistent with P-body and stress granule-enriched RNAs showing an AU bias, an RNA granule dataset has detected AU rich element (ARE) motifs are enriched in RNA granules by **D**iscriminitive **R**egular **E**xpression **M**otif **E**licitation (DREME) analysis (Namkoong et al., 2018). While AREs are more prevalent on longer transcripts by chance alone, the authors note that AREs show a better correlation than transcript length during ER stress. This may be due to this dataset containing a mixed population of stress granules and P-bodies in the preparation, as P-bodies show a stronger AU-bias than stress granules.

Taken together, these results suggest a clear role for AU composition driving RNA partitioning to P-bodies.

***Transcription*:** More recently, it was shown that newly transcribed RNAs tend to avoid sequestration to stress granules (Matheny and Parker, 2018; see chapter 4). This may be due to the newly transcribed RNAs being translated to mount a functional stress response. Alternatively, newly transcribed RNAs may be modified in some yet-to-be-determined manner in order to evade stress granule sequestration. Additionally, stress granules have been shown to contain short poly-A tails (Woo et al., 2018). During ER stress, the length of poly-A tails increases, however the poly-A tails of transcripts targeted to stress induced RNA granules are shorter (Woo et al., 2018). Since newly synthesized mRNAs have longer poly(A) tails, this work suggests a mechanism by which newly transcribed transcripts may evade stress granule sequestration.

Isoforms: Different isoforms of a transcript can show differential partitioning into stress granules (Namkoong et al., 2018). In most cases, stress granules seem to favor the longer isoform of a transcript, consistent with length playing a critical role in defining the RNA composition of stress granules (Namkoong et al., 2018). This finding was reminiscent of NEAT1 partitioning to paraspeckles, in which only the long isoform of NEAT1 is critical for paraspeckle assembly (Jiang et al., 2017). Moreover, different isoforms of RNA have also been shown to localize to different regions of the neuron (Taliaferro et al., 2016), which may be due to their differential partitioning to neuronal granules. Taken together, different isoforms of a transcript can partition differently to RNP granules, either through their altered length, or possibility due to the elements they contain.

RBP association: Finally, RBP association correlates with RNA recruitment to stress granules and P-bodies. By comparing the stress granule transcriptome to eCLIP datasets, a bias was observed for stress granule-enriched RBPs showing a preference for stress granule-enriched transcripts (Khong et al., 2017). Similarly, proteins involved in ALS and proteins containing prion-like domains show preferential binding of stress granule-enriched transcripts. Similar to stress granules, by examining CLIP datasets, P-body enriched RBPs showed preference for P-body-enriched transcripts (Hubstenberger et al., 2017, Yang et al., 2015). Additionally, novel proteins were identified that localize to P-bodies based on CLIP data alone, suggesting that RBP CLIP sites correlate with localization to P-bodies. We delve further into the role of RBPs in stress granule formation in the next section.

Table 1.2 Properties of RNA that play a role in partitioning to RNP granules

parameter	*rationale*	**observed in which RNP granules?**
Transcript length	Longer transcripts are capable of forming more interactions and therefore increase chance of sequestration	**Stress granules** show strongest length bias. **Neuronal granules** show a bias too and have long 3' UTRs **P-bodies** show a weaker, but still significant length bias
Translation status	Poorly translated transcripts have a greater propensity to form RNA-RBP interaction and RNA-RNA interactions since ribosomes aren't continually disrupting interactions	**P-bodies** show the strongest translation bias. **Stress granules** show a weaker bias.
AU content	AU rich transcripts often have higher instances of non-optimal codons and therefore show poor translation. AU rich transcripts may also show a bias to bind certain RBPs	AU-bias is very strong in **p-bodies** and is the best predictor of p-body enrichment. **Stress granules** show a weak AU-bias, but this is more likely due to long RNAs having higher AU bias
RNA-RNA interactions	RNA-RNA interactions are thought to play a role in RNP granule formation because RNA assemblies formed in vitro from total yeast RNA show a remarkable overlap with the SG transcriptome	RNA-RNA interactions likely play a role in most RNP assemblies. However, the correlation observed between RNA droplet transcriptomes and **stress granules**.
Alternate isoforms	Different isoforms can vary drastically in length and/or can associate with different RBPs and form different RNA-RNA interactions	Long isoforms generally partition to **stress granules**. However, there are instances where a small exon can influence localization. In **paraspeckles**, the long isoform of NEAT1 plays a major role in scaffolding

The role of protein and RNA interactions in stress granule formation

The transcriptome of cytoplasmic RNP granules seems to be dominated by long transcripts. This observed length effect begs the question “what aspect of transcript length is causative in RNA partitioning into RNP granules?”. Long transcripts are capable of forming many more interactions than shorter RNAs. Long RNAs also have a higher propensity to form both RNA-RBP interactions as well as RNA-RNA interactions. Herein, we will review what is known about the relative contributions of each of these types of interactions.

RNA binding proteins have been shown to be crucial for stress granule formation. An important experiment has shown deletion of the stress granule RNA-binding protein G3BP inhibits the ability of cells to form stress granules during certain stresses (such as arsenite, thapsigargin, and Pateamine A), while in other stresses (such as NaCl, sorbitol, and heat shock) stress granules can still form (Kedersha et al., 2016). Similarly, deletion of other stress granule RBPs, such as FMRP, limit stress granule formation (Didiot et al., 2009). Moreover, individual stress granule proteins are capable of undergoing phase separation *in vitro*, suggesting that proteins have the ability to assemble in the absence of RNA. Finally, overexpression of G3BP or TIA-1 has been shown to induce stress granule formation in the absence of stress (Gilks et al., 2004, Tourrierre et al., 2003). Taken together, these results demonstrate a clear role for RBPs in stress granule formation.

While stress granule-enriched RNAs show a preference for stress granule proteins, the idea that individual RBPs play a significant role in defining the RNA composition of stress granules has been challenged. Recent work has shown that Pumilio1/2 knockout cells fail to alter the localization of NORAD, which is a stress granule enriched RNA that contains numerous Pumilio binding sites, to stress granules (Namkoong et al., 2018). Further, G3BP knockout cells have an incredibly similar stress granule transcriptome to wild type stress granules during sorbitol induced stress (Matheny and Parker, 2018). Since G3BP is one of the most abundant RBPs in stress granules, the fact that its deletion does not impact RNA partitioning suggests that RBPs may play a limited role in targeting individual transcripts to stress granules. Thus, stress granule enriched RNAs seem to have more binding sites for stress granule RBPs, but it seems that stress granule-enriched RNAs can partition to stress granules independently of the effects of individual RBPs. One possibility that is consistent with these findings is that stress granule targeting is a function of the summation of numerous RBPs. Thus, deletion of any given RBP will only minutely affect localization. However, there remains the possibility that some unknown RBP will have a major effect in altering the localization of transcripts.

Taken together, these results argue that while individual RNA binding proteins can play a role in stress granule formation, they may play a general role by either providing non-specific interactions or providing protein-protein interactions needed to form granules, rather than driving specific RNAs into granules. Consistent with RBPs

playing a more general role in stress granule formation, in cell lines lacking G3BP, TIA1 overexpression can rescue stress granule formation (Kedersha et al., 2016). Nevertheless, it remains a possibility that certain RBPs are capable of conferring specificity and more experiments are needed in order to rule out the effects of individual RBPs.

If individual RBPs play a general role in stress granule formation, how is specificity achieved? Recent studies have proposed RNA-RNA interactions may play a major role in defining the stress granule transcriptome. Indeed, two observations have argued that RNA-RNA interactions may play a critical role in stress granule formation and composition. (1) RNA can form large assemblies *in vitro* through RNA-RNA interactions (Jain and Vail, 2017; Van Treeck et al., 2018). (2) Assemblies made *in vitro* from total yeast RNA have a transcriptome that shows remarkable overlap with the yeast stress granule transcriptome, suggesting that RNA-RNA interactions alone may be sufficient to define the RNA composition of stress granules (VanTreeck et al., 2018). Taken together, these results are consistent with RNA-RNA interactions playing a role in defining stress granule composition and potentially other RBP assemblies.

An intriguing possibility is that RNAs may act to localize RBPs to stress granules. Since stress granules contain regions of increased RBP concentration, termed cores (Jain et al., 2016; Wheeler et al., 2016; Niewedok et al., 2018), might these stable cores represent sites at which RBPs are stably associated with RNAs? Consistent with this model, RNAs show incredibly long residence times inside of stress granules (Moon et

al., 2018). Since RBPs show slower exchange within cores (Niewedok et al., 2018), cores may simply represent sites at which RBPs stably associate with RNA.

As proof of principle of RNAs driving RBPs to stress granules, in an artificial system, MS2-GFP can be localized to stress granules when β-galactosidase RNA tagged with MS2 stem loops was transfected (Mollet et al., 2008). However, while RNAs might help to localize RBPs to stress granules, it is unlikely that deletion or knockdown of any single RNA will lead to a significant change in the RBPs associated with stress granules. Since there are numerous binding sites on many RNAs for a given RBP, it seems unlikely that knockdown of a single RNA will greatly alter the localization of RBPs to stress granules. Consistent with this model, knockdown of the lincRNA NORAD does not alter the localization of Pumilio to stress granules (Namkoong et al., 2018).

Do RNAs localize the RBPs or RBPs localize the RNAs? On the surface this question seems analogous to asking the more straight-forward question “Does the dog wag the tail, or the tail wag the dog?”. In reality, this question is more akin to asking “Does the left side of the dog, or the right side of the dog, localize the dog?”. In many ways, it is a totally nonsensical question. RNPs are likely localized through the synergistic actions of RNA and RBPs. There are many interactions within a stress granule that can lead to sequestration. Thus, stress granule formation is a robust process and it will likely be difficult to perturb the localization of a given transcript to stress granules. Deleting any individual RBP likely will only have a minute effect on

RNA localization, as RNAs probably have a multitude of different interactions that work synergistically to keep the transcript in the granule. Similarly, deletion of any single transcript will likely have minimal effect on localization of any given RBP, since other stress granule RNAs will likely have compensatory RBP sites.

Differential sorting of RNA into P-bodies and stress granules

Stress granules and P-bodies harbor many of the same transcripts. Indeed, P-body and stress granule enrichment scores show a positive correlation (R = .33, **Figure 1A**). Consistent with this correlation, both P-body and stress granules are both biased towards harboring long, poorly-translated, AU-rich transcripts. While P-bodies and stress granule enrichment scores show a significant correlation, the correlation is far from perfect. There are numerous examples in which a transcript is modestly enriched in P-bodies but highly enriched in stress granules. This phenomenon implies there is some underlying specificity in RNA partitioning to P-bodies vs. stress granules. How then is any specificity achieved in targeting transcripts to these assemblies? We suggest that the answer may lie in the relative impact of translation and length.

While length is the predominant metric that correlates with RNA enrichment in stress granules, AU composition is the predominant metric that corelates with RNA enrichment in P-bodies. Stress granules have been shown to also have an AU bias, but the correlation is much less striking than for P-bodies. One possibility that may explain the bias in AU content for stress granule enriched RNAs is that length and nucleotide

composition are correlated **(Figure 1B)**. Moreover, when stress granules are binned by length, the AU-bias no longer seems to play a role in determining stress granule enrichment **(Figure 1C)**. This is in contrast to P-bodies, which show a strong AU bias, even when transcripts are binned by length **(Figure 1D)**. Thus, the AU-bias is a predominant metric for p-body targeting, but is probably not a bona-fide metric for stress granule targeting. Rather, AU composition is a confounding metric for stress granule partitioning due to long transcripts being biased towards AU-rich transcripts.

To take the analysis further, we plotted P-body enrichment scores, vs. stress granule enrichment scores and color-coded the plot based on AU composition **(Figure 1E)**. A striking bias was observed, suggesting that AU composition may play a role in discriminating p-body targeted RNAs from stress granule RNAs. Further, plotting AU composition vs. the ratio of p-body/stress granule enrichment yielded a strong correlation **(Figure 1F)**. Taken together, these results indicate AU composition may play a role in differentially partitioning transcripts into P-bodies or stress granules.

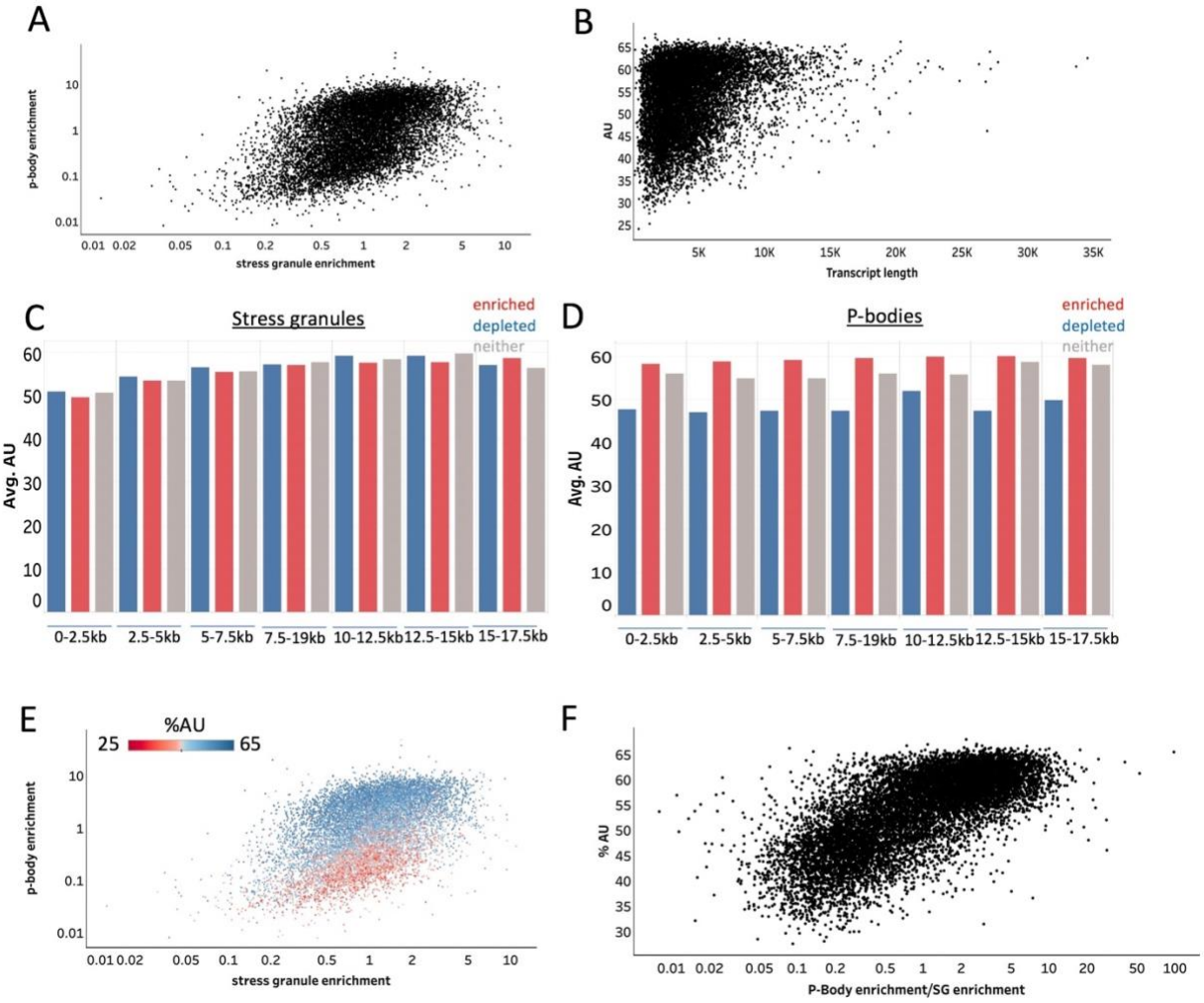

A
p-body enrichment
stress granule enrichment
B
AU
Transcript length
C
Stress granules
enriched
depleted
neither
Avg. AU
0-2.5kb
2.5-5kb
5-7.5kb
7.5-19kb
10-12.5kb
12.5-15kb
15-17.5kb
D
P-bodies
enriched
depleted
neither
Avg. AU
E
%AU
25
65
p-body enrichment
stress granule enrichment
F
% AU
P-Body enrichment/SG enrichment

Figure 1.1 AU composition correlates with differential p-body and stress granule RNA partitioning

(A) Scatterplot showing the correlation between P-body RNA enrichment scores and stress granule RNA enrichment scores. (B) Scatterplot of %AU composition vs. transcript length. (C) Bar graph of average AU content for stress granule enriched and depleted transcripts, binned by length. (D) Same as C, but for P-bodies. (E) Same as A, but color coded for %AU composition. (F) Scatterplot of %AU composition vs P-body:stress granule enrichment score ratio.

What is the underlying mechanism of this phenomenon? We suggest that translation may be responsible. AU-rich transcripts have high instances of rare codons (Courel et al., 2017). Rare codons have been shown to markedly reduce the translation of transcripts (reviewed in Hanson and Coller. 2018). Consistent with this observation, P-bodies are biased towards poorly translated transcripts, and show a larger translation bias.

We propose that substrate availability may explain the specificity discrepancy. During normal growth conditions, translation is robust. During these conditions, only the poorest translators would be available for sequestration into RNP granules as translation would compete to keep the majority of transcripts out of P-bodies. Thus, the observed translation biases, AU-rich biases, and non-optimal codon biases for p-body sequestration are expected, while length would play a more limited role if a given long transcript were still efficiently translated **(Figure 2A)**.

Under conditions of stress, bulk translation is greatly impaired. Thus, the translational equilibrium is greatly shifted. Most transcripts are now translationally impaired, except for short stress response genes, and may explain why AU content and translation efficiency show lowered correlation in stress granules compared to P-bodies. Since most transcripts are translationally repressed, length would now serve as the dominant metric for partitioning to stress granules and/or P-bodies since these transcripts are capable of forming many more interactions than short transcripts **(Figure**

2B).

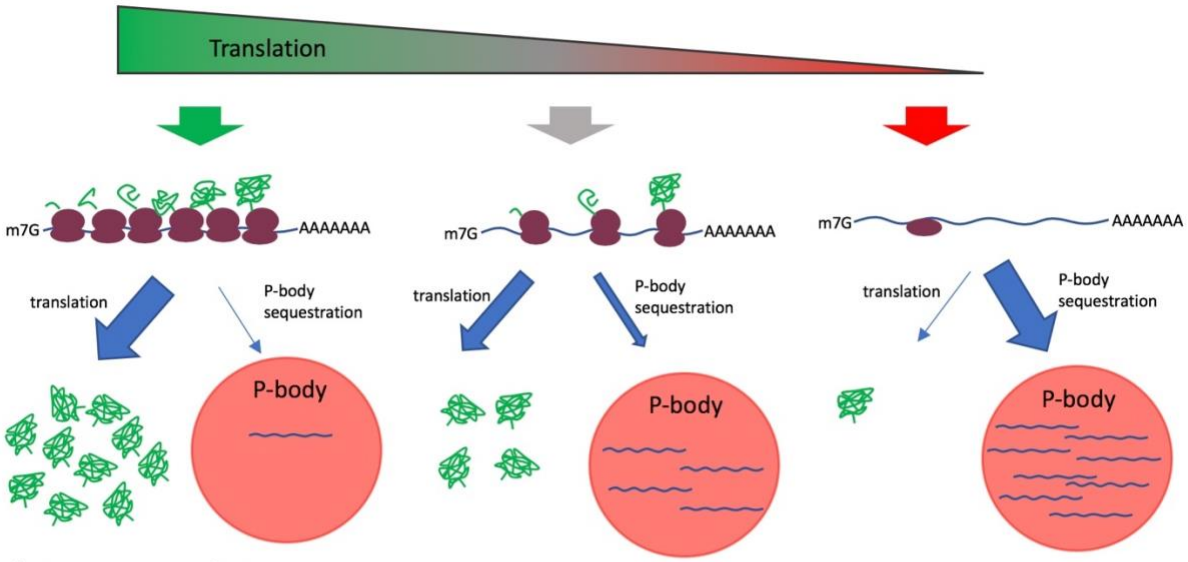
This translation-dependent model of RNA partitioning to RNA granules makes a couple of interesting predictions. First, it suggests that specificity is determined, in part, by the availability of RNA, which, in turn, is governed by translation (as well as transcription). Thus, RNA partitioning into other granules may be dominated by (1) the available pool of non-translating transcripts and (2) the propensity of the available transcripts to form assemblies.

Secondly, it suggests that the transcriptome of P-bodies may change under conditions of stress where translation is inhibited. Since stress causes global inhibition of translation, multiple RNAs are made available for sequestration into RNP granules. If the metrics that govern RNA localization to P-bodies and stress granules are similar, one would anticipate that the RNA composition of P-bodies would shift during stress and possibly more closely resemble the stress granule transcriptome. Since the repression of translation initiation increases the number of P-bodies per cell in addition to forming stress granules, an intriguing, albeit challenging, experiment would be to examine whether the transcriptome of newly formed P-bodies differs from the transcriptome of pre-stress P-bodies.

While AU-rich transcripts are poorly translated, AU composition may also lead to P-body sequestration due to the interactions these transcripts form. For example, AU-rich transcripts may have drastically different association with RBPs and may be able to

form unique RNA-RNA interactions. This may, in part, lead to their differential sorting into P-bodies. It is important to note that AU composition defining the available substrate pool and AU composition altering the types of interactions RNAs form are not mutually exclusive models, and both models likely play a role in defining differential partitioning to stress granules and P-bodies.

(A) Normal growth conditions



(B) Stress conditions

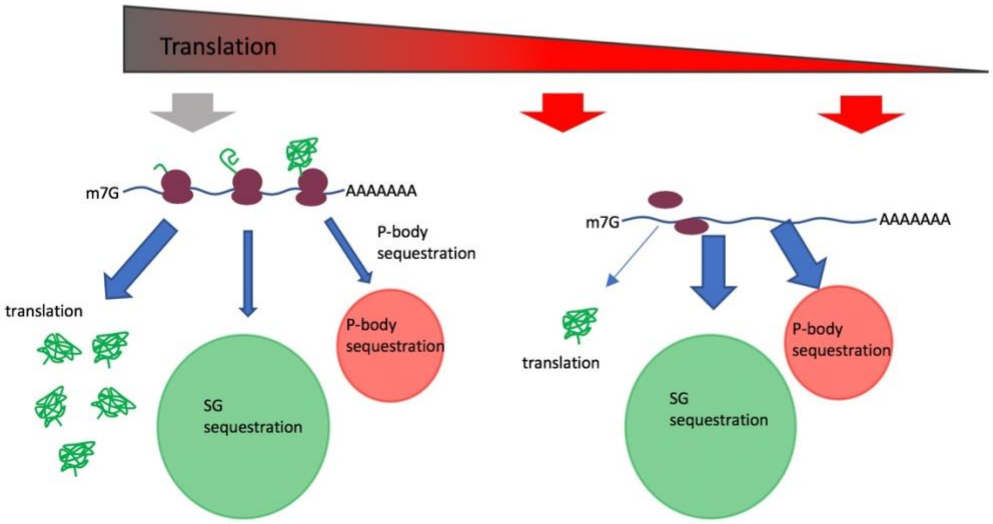


Figure 1.2 Translation equilibrium model for differential RNA partitioning

(A) During normal growth conditions, translation is robust. Therefore, only the poorest translators are available for P-body sequestration. (B) During stress conditions, translation is greatly reduced, making many more new substrates available for RNP granules sequestration. (P-bodies represented as red circle, stress granules represented as green cirlcles)

CONCLUSION

In this review we summarize the predominant metrics that correlate with RNA localization to RNP assemblies. Based on the transcriptomes of stress granules, P-bodies, and neuronal granules, we compile a list of metrics that correlate with enrichment in RNP granules. Length, AU content, translation, poly-A tail length, RNA-RNA interactions and association with RBPs are all key metrics that correlate with RNA partitioning although the relative contribution of these metrics depend on the RNP assembly (i.e. stress granules correlate strongly with length while p-bodies correlate strongly with AU content).

Moving forward a major avenue of research will be to parse out which other metrics play a role in RNP targeting. As transcript length increases, the amount of potential RNA-RBP and RNA-RNA interactions also increases. Additionally, long transcripts contain more sites for RNA modifications than shorter RNAs. Undoubtedly, parsing out which targeting elements are buried in this length parameter will yield fruitful insights into RNP granule assembly.

However, we would like to point out that transcript length is an inherent regulatory mechanism for an RNA. The probabilistic tendency of long transcripts to harbor motifs and/or modifications is itself a mechanism for regulation. For example, in neurons, a variety of genes are needed to be targeted to the synapse. One evolutionary strategy for targeting shorter transcripts to the synapse would be to evolve

highly specific targeting sequences. However, this does not seem to be the strategy neurons have utilized. Neurons have taken a more simplistic approach, and evolved long transcripts to localize to the synapse. Such a strategy is beautiful in its simplicity, because it takes advantage of a probabilistic certainty.

Finally, the unveiling of these RNP granule transcriptomes will hopefully allow for clear functions to be elucidated for these assemblies. The function of stress granules and P-bodies has remained elusive, largely due to a lack of information of which transcripts are most highly targeted to these assemblies. Whether these assemblies are consequential to gene expression remains a mystery, but with the characterization of the stress granule and P-body transcriptomes, function will soon be able to be tested on candidate transcripts.

CHAPTER 2

DISTINCT STAGES IN STRESS GRANULE ASSEMBLY AND DISASSEMBLY[1]

SUMMARY

Stress granules are non-membrane bound RNA-protein (RNP) assemblies that form when translation initiation is limited and contain a biphasic structure with stable core structures surrounded by a less concentrated shell. The order of assembly and disassembly of these two structures remains unknown. Time course analysis of granule assembly suggests that core formation is an early event in granule assembly. Stress granule disassembly is also a stepwise process with shell dissipation followed by core clearance. Perturbations that alter liquid-liquid phase separations (LLPS) driven by intrinsically disordered protein regions (IDR) of RNA binding proteins in vitro have the opposite effect on stress granule assembly in vivo. Taken together, these observations argue that stress granules assemble through a multistep process initiated by stable assembly of untranslated mRNPs into core structures, which could provide sufficient high local concentrations to allow for a localized LLPS driven by IDRs on RNA binding proteins.

---

[1] This chapter has been modified from the following publication: Wheeler JR.*, Matheny T.*, Jain S.*, Abrisch R., and Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. Elife 2016 Sep 7, 5. *These authors contributed equally.

INTRODUCTION

Stress granules are non-membranous assemblies of mRNA and protein (mRNP) that form when translation initiation is limiting, which occurs during many stress responses. Stress granules are thought to influence mRNA function, localization, and to affect signaling pathways (Buchan, 2014; Buchan and Parker, 2009; Kedersha et al., 2013). Normally, stress granule formation is a dynamic, reversible process. However, pathological mutations in proteins that either increase the formation, or decrease the clearance of stress granules, can lead to abnormal accumulation of aggregates that share components with stress granules (Buchan et al., 2013; Dormann et al., 2010; Izumi et al., 2015; Kim et al., 2013). Abnormal accumulation of stress granule-like aggregates is associated with neurodegenerative disease (Li et al., 2013; Ramaswami et al., 2013). The molecular interactions and mechanisms that regulate stress granule assembly and how stress granules may become altered in disease remain unknown.

Stress granules are members of an emerging class of non-membrane bound organelles and are thought to represent multicomponent viscous liquid droplets that from spontaneously by liquid-liquid phase separation (LLPS) (Brangwynne, 2013; Brangwynne et al., 2009, 2015; Elbaum-Garfinkle et al., 2015; Hyman and Simons, 2012; Hyman et al., 2014). In vitro LLPS droplets display many properties of in vivo stress granules including fusion, wetting, shearing, and dynamicity (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015; Zhang et al., 2015). LLPS in vitro can be driven by high local concentrations of proteins containing intrinsically

disordered regions (IDRs) (Lin et al., 2015; Molliex et al., 2015; Nott et al., 2015; Wright and Dyson, 2015). Since IDR-containing proteins are enriched in stress granules and other mRNP granules (Decker et al., 2007; Gilks et al., 2004; Jain et al., 2016; Kato et al., 2012; King et al., 2012; Reijns et al., 2008), this has led to the suggestion that IDRs on mRNA binding proteins form hetero- and homotypic interactions that drive initial formation of stress granules and P-bodies (Gilks et al., 2004; Decker et al., 2007; Lin et al., 2015; Molliex et al., 2015; Patel et al., 2015; Toretsky and Wright, 2014; Weber and Brangwynne, 2012; Zhang et al., 2015). This model is supported by observations in vitro wherein high concentrations of IDRs are sufficient to spontaneously form LLPS possibly through the interplay of weak electrostatic and hydrophobic homo- and heterotypic protein-protein interactions (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Nott et al., 2015; Pak et al., 2016; Patel et al., 2015).

The interactions amongst IDRs driving phase-separation are enhanced by altering the molecular milieu either through lower salt concentration, molecular crowding, addition of RNA, or lower temperature (Lin et al., 2015; Molliex et al., 2015; Nott et al., 2015; Patel et al., 2015). In contrast, small aliphatic molecules, such as 1,6-Hexanediol, are hypothesized to perturb weak hydrophobic interactions, and thus disassemble assemblies that exhibit liquid-like properties in vitro (Patel et al., 2007; Ribbeck and Görlich, 2002). These in vitro observations suggest that manipulation of these properties might be useful for discerning the molecular interactions and mechanisms that drive assembly and regulate stress granules in vivo. Indeed, since mammalian stress granules have been reported to disassemble when cells are treated

with 1,6-Hexanediol, they have been inferred to be LLPSs (Kroschwald et al., 2015).

Analysis of mammalian stress granules has revealed they are comprised of two phases: a less concentrated shell, which disassembles upon cell lysis and thus behaves like a LLPS formed by weak interactions; and more stable, internal “core” structures (Jain et al., 2016). Interestingly, formation of in vitro liquid droplets by IDRs have been observed to mature into a second, less dynamic phase comprised in part of stable amyloid-like assemblies (Han et al., 2012; Kato et al., 2012; Kwon et al., 2013; Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015). A parsimonious model therefore is that stress granules may initially assemble to first form a LLPS comprised of multivalent, weak and dynamic interactions, and over time a less dynamic, more viscous second phase matures from the continued supersaturation of IDRs, thereby generating stress granule cores (Jain et al., 2016). Here, we set out to test several predictions of this model: (i) that lower temperature should enhance granule assembly, as LLPS driven by IDRs on RNA binding proteins have been shown to be enhanced at lower temperatures (Molliex et al., 2015; Nott et al., 2015); (ii) that stress granules are sensitive to drugs, such as 1,6-Hexanediol, that are predicted to disrupt weak hydrophobic interactions; (iii) that stress granule dynamics decrease and core size should increase over time as stress granule structure matures.

In contrast to the expected results, we observed that stress granule assembly is inhibited at lower temperatures, despite the extent of translation repression being temperature independent. Moreover, we observe that 1,6-Hexanediol triggers stress granule formation in both yeast and mammalian cells, as well as altering many cellular

structures including the cytoskeleton and the nuclear pore complex. Finally, we observe stress granule dynamics and core size is unchanged during prolonged stress. Given these observations we suggest an alternative model where untranslating mRNPs initially oligomerize into stable cores, thereby nucleating an initial shell layer and providing a platform for the growth of a more dynamic shell around these cores, followed by merger of these individual core/shell assemblies into large mature stress granules. Stress granule disassembly is also a stepwise process where shell dissipation occurs first followed by clearance of cores.

Two models of granule assembly can explain a biphasic stress granule architecture

We have previously described that mammalian stress granules formed after 60 minutes of sodium arsenite ($NaAsO_2$) treatment contain substructures (referred to as cores) that are stable in lysates, as well as a less concentrated surrounding structure, referred to as the shell (Jain et al., 2016). The shell region of stress granules does not persist in our lysis conditions, which suggests it is sensitive to dilution, one of the properties of an LLPS.

The biphasic stress granules seen after 60 minutes in mammalian cells could form by one of two models (**Figure 2.1 A**). First, the increased pool of untranslated mRNAs bound by proteins containing IDRs, could lead to the formation of a LLPS based on IDR-IDR interactions that would be very fluid and dynamic. In this view, cores would assemble within this initial LLPS in a second step, possibly due to the increased local concentration of core components (Jain et al., 2016). In support of this view, LLPS derived from RNA binding protein IDRs *in vitro* mature to include more stable, possibly amyloid-like, substructure (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015).

In a second model, stress granule formation could be initiated by the oligomerization of mRNPs into stable core structures. Because these stable core

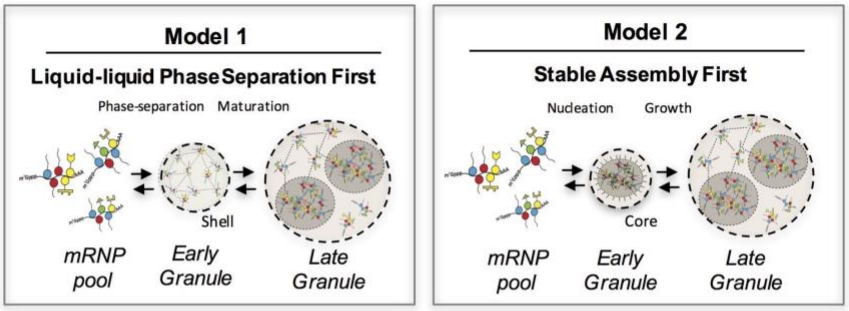
structures would concentrate mRNA binding proteins with IDRs, stable cores could function as nucleation sites for LLPS by creating a high local concentration of IDRs to drive LLPS formation. Over time, cores may then form into larger LLPS assemblies with multiple core structures within each stress granule. In this latter model the formation of the "shell" structure of the stress granule would be analogous to the formation of the nuclear pore, wherein a high local concentration of FG-repeat containing proteins is imposed by the structure of the nuclear pore, which then facilitates the formation of phase transition in the pore region (Frey and Görlich, 2007; Hülsmann et al., 2012; Schmidt and Görlich, 2016).

To distinguish between these two models for stress granule assembly we examined the time course of stress granule assembly, as well as the biochemical, ultrastructural, and dynamic properties of stress granules at different stages of their assembly.

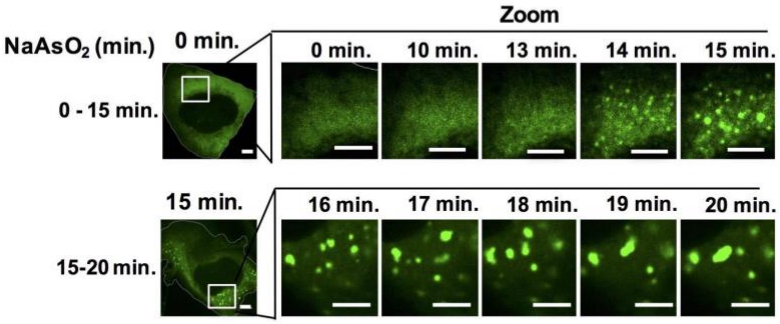
To examine the time course of stress granule assembly, we exposed human osteosarcoma cells (U-2 OS) expressing a GFP-G3BP1 fusion protein (Kedersha et al., 2008; Ohn et al., 2008) to $NaAsO_2$ and collected images over time using a spinning disc confocal microscope. We observed that stress granules began to form 13-15 minutes after stress induction, which correlated with an increase in eIF2 phosphorylation (**Figure 2.1B, 2.1C**). The initial stress granules then grew in size both by accretion of material and by fusion with other stress granules (**Figure 2.1B**). The area of GFP-G3BP1 stress granules continued to increase with the majority of stress granule assembly completed

by 40 minutes. Stress granules induced with thapsigargin or osmotic stress showed a similar time course of stress granule growth albeit the rates of granule induction were slightly different (**Figure 2.1C**). Our data is consistent with the idea that assembly of a mature stress granule occurs via a multistep process: an initial nucleation event that may be influenced by the stress stimulus and a consistent growth phase that requires an addition of substrate and fusion of the small initial stress granules into larger assemblies.

A)


Model 1
Liquid-liquid Phase Separation First
Phase-separation
Maturation
Shell
mRNP pool
Early Granule
Late Granule
Model 2
Stable Assembly First
Nucleation
Growth
Core
mRNP pool
Early Granule
Late Granule

B)


Zoom
NaAsO2 (min.)
0 min.
0 - 15 min.
0 min.
10 min.
13 min.
14 min.
15 min.
15 min.
15-20 min.
16 min.
17 min.
18 min.
19 min.
20 min.

C)

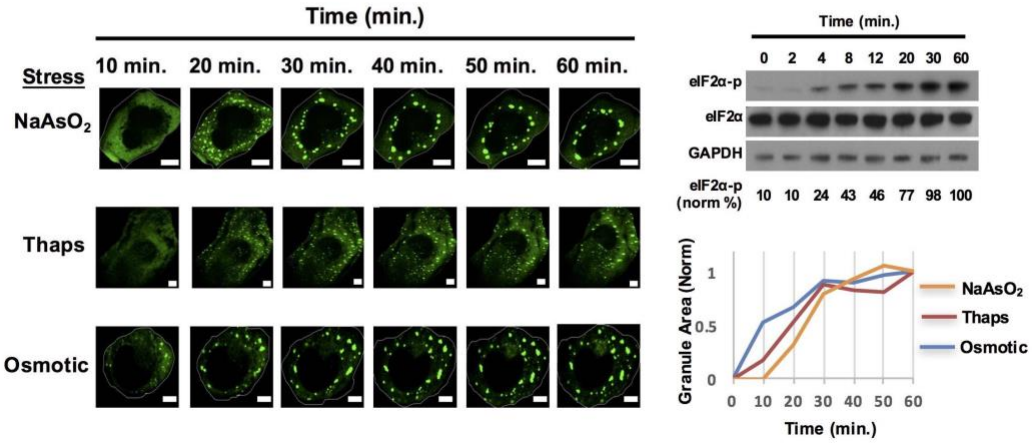

Time (min.)
Stress
10 min.
20 min.
30 min.
40 min.
50 min.
60 min.
NaAsO2
Thaps
Osmotic
Time (min.)
0 2 4 8 12 20 30 60
eIF2α-p
eIF2α
GAPDH
eIF2α-p (norm %)
10 10 24 43 46 77 98 100
Granule Area (Norm)
1
0.5
0
0 10 20 30 40 50 60
Time (min.)
NaAsO2
Thaps
Osmotic

Figure 2.1: Time course of stress granule assembly

(A) Two models for stress granule assembly. In Model 1, an increased pool of untranslated mRNAs with bound proteins containing IDRs, could lead to the formation of a LLPS based on weak, dynamic IDR-IDR interactions and over time core formation occurs as a second phase due to the supersaturation of local concentration of core components. In Model 2, untranslating mRNAs with bound proteins containing IDRs could initially oligomerize into stable cores that provide a platform for LLPS and eventual coalescence of multiple cores results in formation of a larger LLPS assembly. (B) Time-lapse imaging stress granule assembly and early dynamics in U-2 OS cells expressing GFP-G3BP1 during $NaAsO_2$ stress using 100X objective. (C) Time-lapse imaging of stress granule assembly under $NaAsO_2$ (0.5mM), thapsigarin (Thaps, 100nM), or osmotic stress (375M sorbitol). Western blot for eiF2 phosphorylation status following exposure to $NaAsO_2$ stress over time. Percent eiF2 phosphorylation was normalized to total eiF2. eiF2 and GAPDH serve as loading controls. Graph shows average granule area plotted for each stress condition normalized to maximal granule area at 60 minutes. All scale bars are 5 m.

Core formation is an early step in stress granule formation

If a stress granule shell forms first, then stress granules formed at early time points should be comprised of only a shell, no core structures, and would not be stable in lysates. To test this, we examined if the earliest detectable stress granules contain compartments that are stable in lysates (Jain et al., 2016).

Lysis of cells at 5 minutes, 10 minutes, and 15 minutes after $NaAsO_2$ induction revealed that by 15 minutes, when stress granules are first visible in cells, we observed an increase in stable stress granule cores detected in lysates (**Figure 2.2A**). We identify these GFP-G3BP1 foci in lysates as stress granule cores since they contain PABP1, poly(A)+ RNA, and eIF4G (**Figure 2.2B**). Similarly, yeast stress granules are stable in lysates as soon as they are microscopically detectible *in vivo* (data not shown). We interpret this observation to indicate that as soon as stress granules are observed in cells by microscopy they contain core assemblies that are stable in lysates.

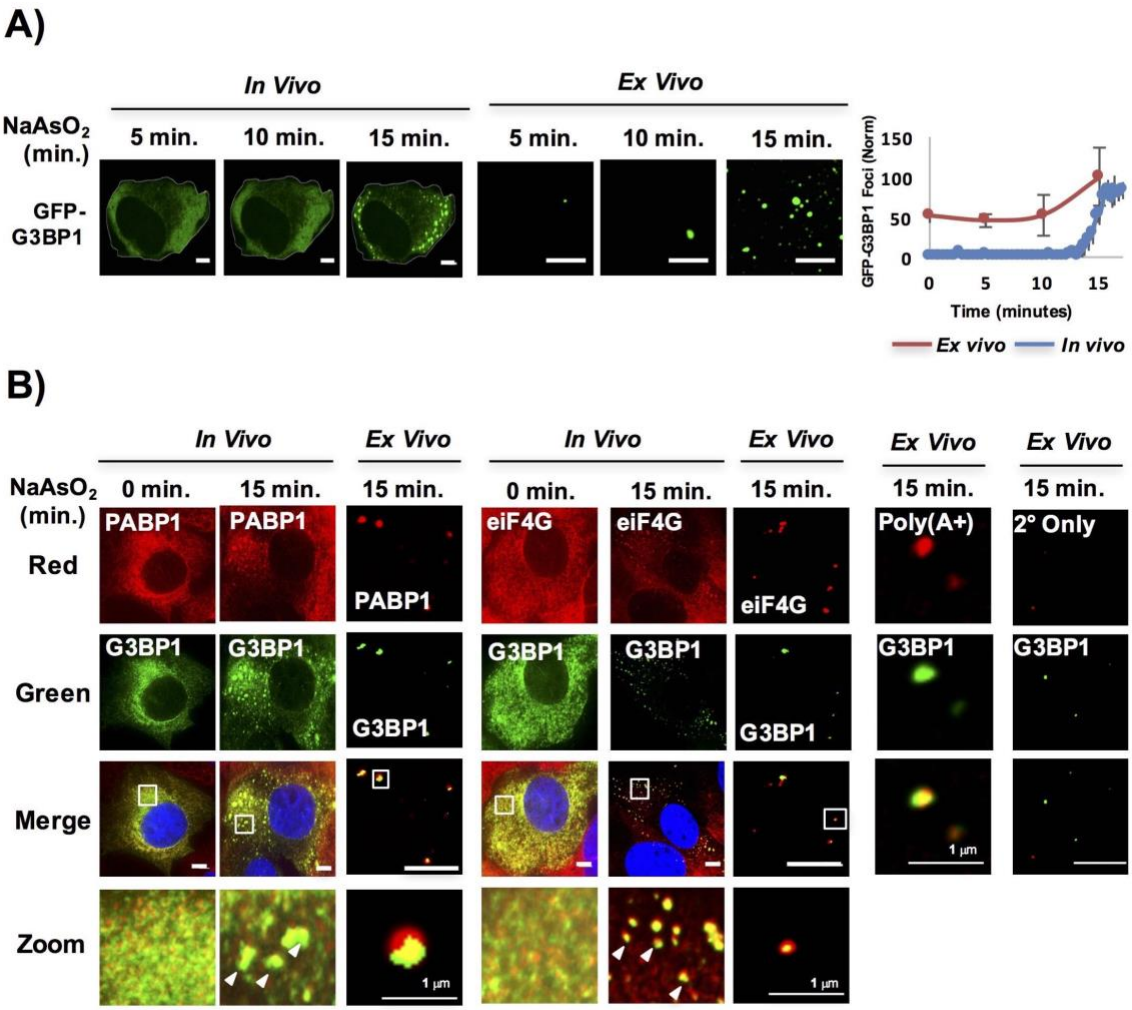

A)
In Vivo
Ex Vivo
NaAsO2 (min.)
5 min.
10 min.
15 min.
5 min.
10 min.
15 min.
GFP-G3BP1
GFP-G3BP1 Foci (Norm)
150
100
50
0
0
5
10
15
Time (minutes)
Ex vivo
In vivo
B)
In Vivo
Ex Vivo
In Vivo
Ex Vivo
Ex Vivo
Ex Vivo
NaAsO2 (min.)
0 min.
15 min.
15 min.
0 min.
15 min.
15 min.
15 min.
15 min.
Red
PABP1
PABP1
PABP1
eiF4G
eiF4G
eiF4G
Poly(A+)
2° Only
Green
G3BP1
G3BP1
G3BP1
G3BP1
G3BP1
G3BP1
G3BP1
G3BP1
Merge
1 µm
Zoom
1 µm
1 µm

Figure 2.2: Stress granule cores are stable in lysates early in granule assembly

(A) GFP-G3BP1 U-2 OS cells and cell lysates following treatment with $NaAsO_2$ for 15 minutes. Graph shows percent GFP-G3BP1 foci for cells (*in vivo*) and cell lysates (*in vitro*) and normalized to maximal number of foci detected at 15 minutes (*in vivo*) or number of foci detected at time 0 minute (*in vitro*). (B) GFP-G3BP1 U-2 OS cells and cell lysates probed for the stress granule markers, PABP1 or eIF4G following treatment with $NaAsO_2$ for 15 minutes. GFP-G3BP1 foci in lysates probed for poly (A+) RNA by oligo-dT or secondary-only antibody from cells $NaAsO_2$ stressed for 15 minutes. GFP-G3BP1 foci in lysates were probed with secondary antibody only (Alexa-647) at same concentrations used for primary (PABP1 or eIF4G) antibody detection. Zoom represents magnified inset. Unless otherwise noted, all scale bars are 5 m.

Early and late stress granules are non-uniform by super resolution microscopy

As an additional test to see if stress granules formed at early times contain core substructures, we examined the ultrastructure of stress granules by super-resolution microscopy (Structured Illumination Microscopy [SIM]) at different times. Using SIM for G3BP1, as assessed by mapping pixel intensity for individual granules, we observed that stress granules contained substructure as soon as they were large enough to be examined by SIM (20 minutes) (**Figure 2.3**).

Based on these two observations, we conclude that core assembly represents an early event in stress granule formation.

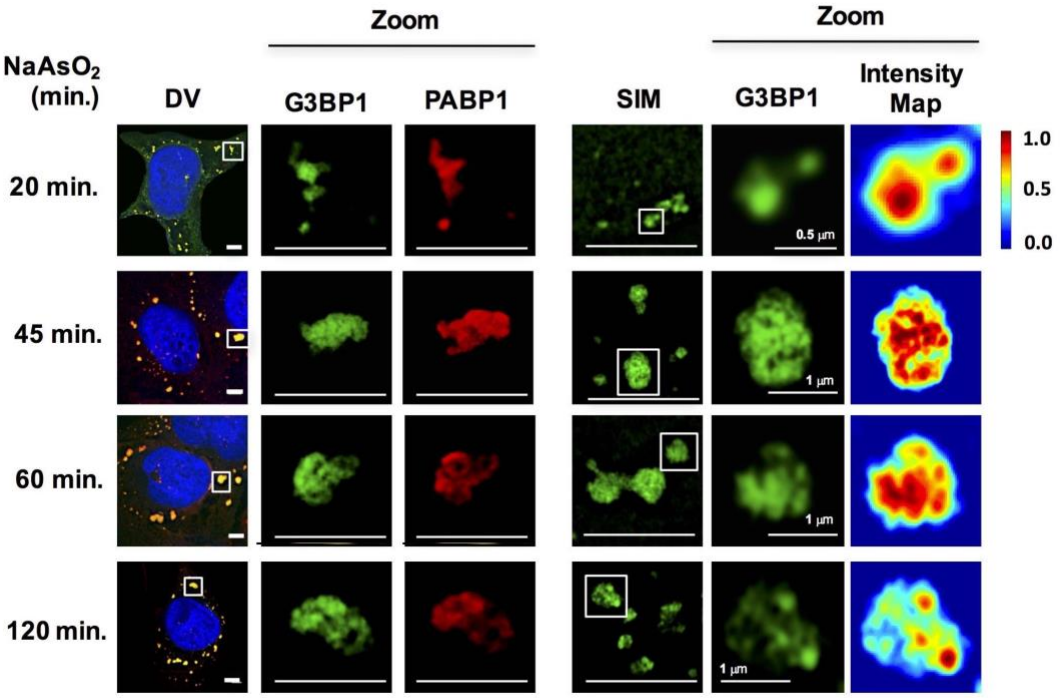


Figure 2.3: Early and late stress granules contain a non-uniform substructure
GFP-G3BP1 stress granules were stained with the stress granule marker PABP1 and imaged by deconvolution microscopy (DV) following $NaAsO_2$ stress at different time points. GFP-G3BP1 was assessed in the same cells by structured illumination microscopy (SIM). Intensity map represents relative gray scale intensity. Zoom represents magnified inset. Unless otherwise noted, all scale bars are 5 m.

Stress granule core size and dynamics do not change over time

Previous *in vitro* experiments demonstrate that IDR-containing stress granule proteins can undergo a maturation event leading to the formation of more stable complexes and possibly amyloid-like assemblies (Kato et al., 2012; Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Nott et al., 2015; Patel et al., 2015; Zhang et al., 2015). To test if stress granules show a similar “maturation” in cells, we examined both the dynamics and core dimensions at different times during stress granule formation.

We examined the properties of the stable components of stress granules in lysates to see if there was an increase in size of the core structures over time, which would imply a continued growth of these stable core structures and might reveal maturation of the core structure within a larger stress granule. By Nanosight Nanoparticle Tracking analysis of lysates from U-2 OS cells at different times after stress, we observed that GFP-G3BP1 granule core size does not significantly change between 15 minutes (median size, 259 74 nm), 30 minutes (median size, 258 24nm), 60 minutes (median size, 273 12nm), and 120 minutes (median size, 247 4nm) of $NaAsO_2$ stress (one-way ANOVA, p-value = 0.94) (**Figure 2.4**). We interpret these results to demonstrate that GFP-G3BP1 granule core size does not appreciably mature into larger stable assemblies within the time periods assessed.

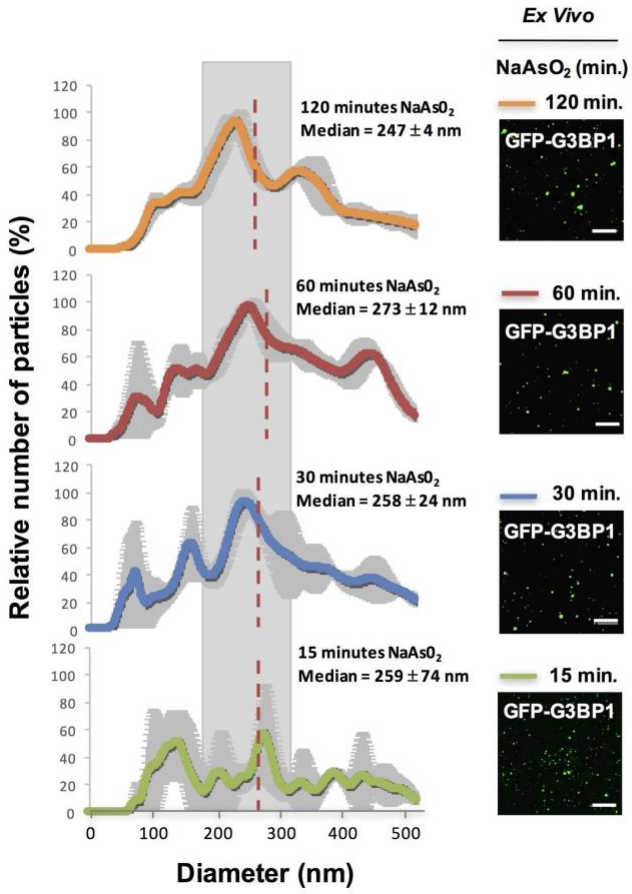


Figure 2.4: Similar GFP-G3BP1 stress granule core size during early and prolonged stress

Mean of 3 Nanosight experiments ± standard deviation for GFP-G3BP1 mammalian stress granule cores isolated from U-2 OS cells stressed with $NaAsO_2$ for 15, 30, 60, or 120 minutes. Relative number of particles for each time point are represented as a percentage of maximal number of particles per unit size. Mean median from these 3 experiments is highlighted. Representative images from GFP-G3BP1 cores in lysates at respective time points are shown. All scale bars are 5 m.

In a second experiment, we examined if the dynamics of stress granules change over time as assessed by FRAP (Fluorescence Recovery After Photobleaching) analysis of GFP-G3BP1. This experiment is analogous to experiments done *in vitro* where the recovery rates of LLPS driven by IDRs of RNA binding components decrease over time revealing a maturation process within these reconstituted assemblies (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015; Zhang et al., 2015). Thus, we examined if the dynamics of normal stress granules change over time, which might reveal an analogous maturation process in cells. We observed that the recovery rates of GFP-G3BP1 were essentially identical after 30 minutes, 60 minutes, or 120 minutes of exposure to $NaAsO_2$ (**Figure 2.5**), although the half-time of recovery was slightly faster at 30 minutes (18.5 seconds) as compared to 60 minutes (35 seconds) and 120 minutes (31 seconds). This slightly faster recovery rate is likely to simply reflect that at 30 minutes stress granules are still undergoing growth, which will increase the apparent rate of recovery. We interpret this observation to argue there is not a global transition of stress granule components to a more stable state within the time periods assessed. However, we cannot rule out that specific components of stress granules become less dynamic over time.

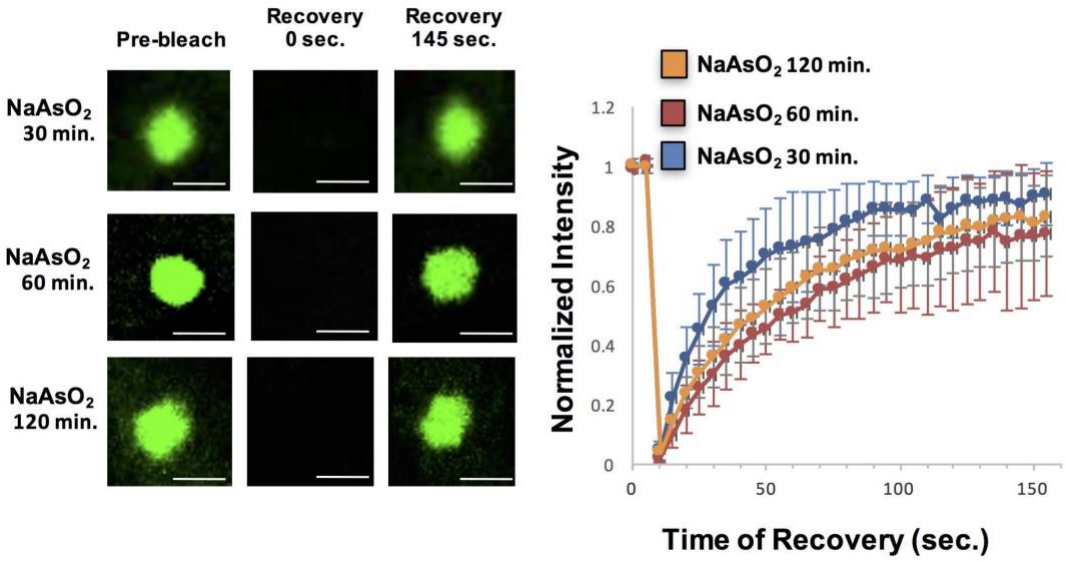


Figure 2.5: Similar GFP-G3BP1 stress granule dynamics during early and prolonged stress

Granules shown prior to photobleaching, at 0 seconds, and at 145 seconds after photobleaching. Cells were treated for either 30, 60, 120 minutes with $NaAsO_2$. Graph shows recovery curves as an average of 6 granules ± standard deviation at each respective time point. Scale bars are 1.5 µm.

Mammalian stress granule assembly is inhibited at lower temperatures

Our observations suggest that core formation is an early event in stress granule assembly; however, we cannot distinguish by light microscopy if cores precede shell formation or not. To distinguish between these early events, we first tested whether stress granule assembly is enhanced at lower temperatures. The logic of this experiment is based on multiple experiments *in vitro* showing that IDRs from RNA binding proteins can form either LLPS (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Nott et al., 2015; Zhang et al., 2015) or hydrogels, which in this context are a meshwork assembly of protein filaments (Han et al., 2012; Kato et al., 2012). In these experiments, the LLPS or hydrogel formation is enhanced at lower temperatures, which would be expected to lower the critical concentration required for assembly (Molliex et al., 2015; Murakami et al., 2015; Nott et al., 2015; Kato et al., 2012). If these types of interactions drive stress granule assembly *in vivo*, then stress granule assembly should be more efficient at lower temperatures. To test this possibility, we examined stress granule formation at different temperatures in GFP-G3BP1 U-2 OS cells, in response to $NaAsO_2$ treatment.

We observed that stress granules formed less efficiently at lower temperatures with a continuing decline in the rate, overall number and size of stress granule formed as the temperature declined from 37°C to 27°C (**Figure 2.6A, 2.6B**). Two observations suggest that lower temperatures do not perturb translational inhibition. First, we observe

a stress-specific induction of eiF2-phosphorylation to similar levels amongst all the temperatures assessed (**Figure 2.6C**). Second, via polysome profiling, we observed a similar extent of translation repression after 20 minutes of $NaAsO_2$ stress at 30C as compared to 37C. (**Figure 2.6E**). Taken together, we conclude that despite a sufficient extent of translation repression and an environment shown to be conducive for phase separation *in vitro* (i.e. lower temperatures), we observe a delay in stress granule formation at lower temperatures.

In principle, stress granule formation could be dependent on higher temperatures either because the interactions holding stress granules together are disrupted at lower temperatures, or because the assembly of stress granules is an ordered process with some step in the process dependent on higher temperatures. If the interactions holding stress granules together are dependent on higher temperatures, then stress granules formed at higher temperatures should disassemble at lower temperatures. Conversely, if stress granules are assemblies that have a temperature dependent step in their assembly, then once formed they should persist after a drop in the temperature. To test these possibilities, we induced stress granule formation at 37C for 45 minutes then imaged the cells at a lower temperature (30C) for 15 minutes (**Figure 2.6D**). We observed that once formed, stress granules persist at a lower temperature and do not appreciably decrease in size. It should be noted that shifting cells to 30C (without $NaAsO_2$) does not induce stress granule formation (**Figure 2.6D**). This is inconsistent with stress granules being held together by interactions that are enhanced at higher temperatures, and instead argues stress granules have a step in their assembly that is

inhibited at low temperatures, but once formed stress granules are stable at both high and low temperatures. This suggests that the primary driving forces of stress granule assembly are not weak interactions between IDR domains that are enhanced at low temperatures.

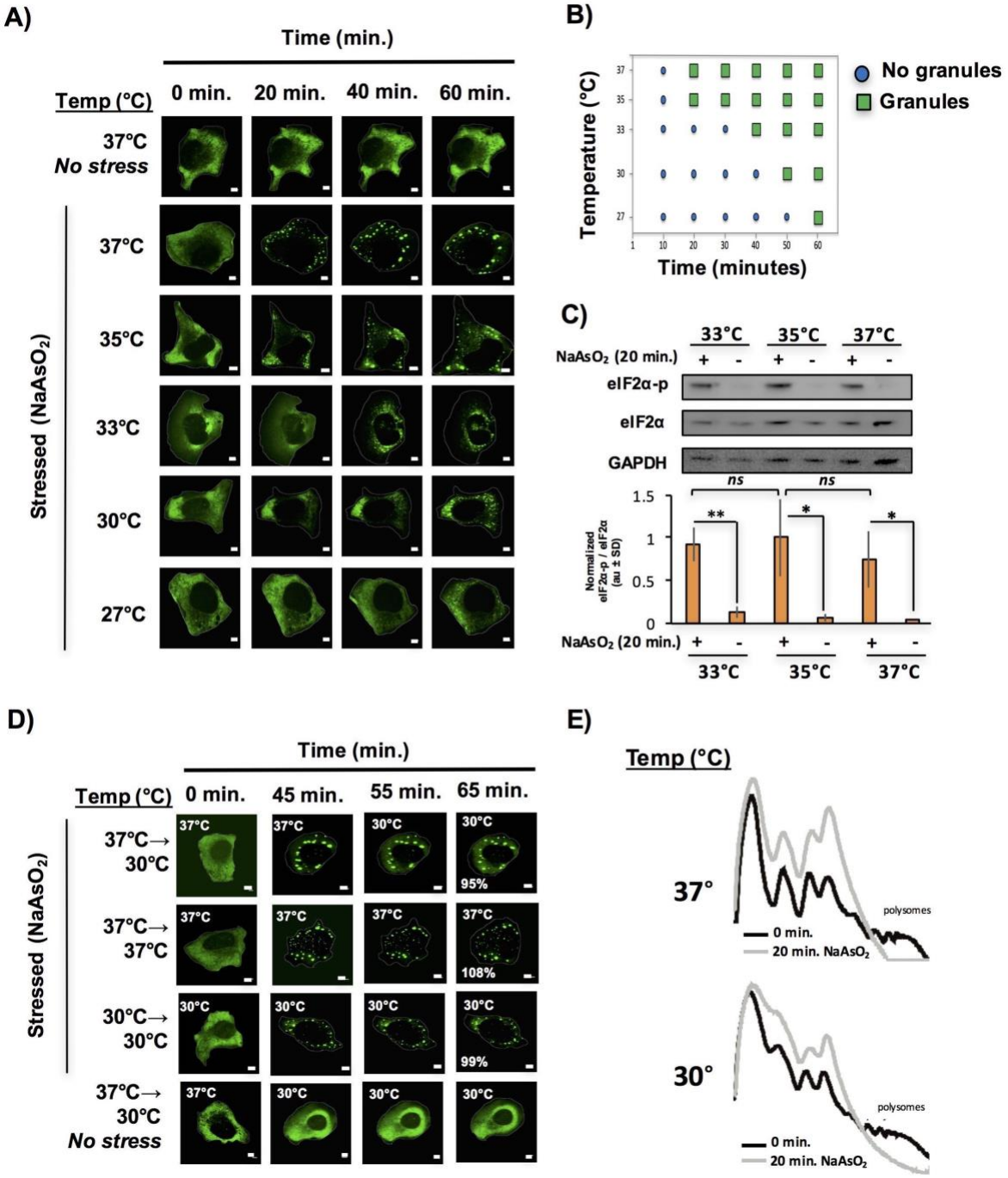


Figure 2.6: Temperature inversely affects kinetics of stress granule formation

(A) Time-lapse microscopy of U-2 OS cells expressing GFP-G3BP1. U-2 OS cells imaged at respective temperatures in the presence of $NaAsO_2$ stress for 1hr using 100X

objective. (B) Graphical representation of temperature on stress granule formation in U-2 OS cells. (C) Western blot for eiF2α phosphorylation status following +/- exposure to 20 minutes of $NaAsO_2$ stress at respective temperatures. eiF2α and GAPDH serve as loading controls. Normalized to total eiF2α (N=3, au = arbitary units, SD = standard deviation). (p-value: *<0.05; **<0.01; ns = not significant). (D) Time-lapse microscopy of U-2 OS cells expressing GFP-G3BP1 at respective temperatures following 45 minutes $NaAsO_2$ stress using 100X objective. Temperatures at which cells were imaged are indicated in top left of each panel. Percentages listed in final panel represent average area of granules normalized to the area of granules at the start of image acquisition. (E) Polysome analysis of U-2 OS cells (grown either at 30°C or 37°C) during steady state or following 20 minutes $NaAsO_2$ stress. All scale bars are 5 m.

1,6-Hexanediol disrupts cellular structures in yeast and alters cellular morphology in HeLa cells

Our observations suggest that promoting LLPS by lower temperature is not sufficient to enhance granule formation. To test if an LLPS shell precedes and is necessary for granule core, we decided to test the effect of inhibiting or disrupting LLPS formation by the addition of 1,6-Hexanediol.

1,6-Hexanediol has previously been used to disrupt various structures that are expected to represent LLPS, *in vitro* and *in vivo* (Kroschwald et al., 2015; Molliex et al., 2015; Patel et al., 2007; Ribbeck and Görlich, 2002; Updike et al., 2011). As the nuclear pore is proposed to be a LLPS (Patel et al., 2007; Ribbeck and Görlich, 2002; Updike et al., 2011), a prediction is that 1,6-Hexanediol should disrupt this structure *in vivo.* Consistent with this prediction, the addition of 10% 1,6-Hexanediol to yeast cells (Kroschwald et al., 2015), rapidly disrupted the nuclear pore, as adjudged by the loss of punctate, peri-nuclear localization of the nuclear pore protein, Nsp1 (**Figure 2.7A**). However, we noticed that the effect of 1,6-Hexanediol was not limited to cellular structures that are known to represent a LLPS. We observed that 1,6-Hexanediol (but not 1, 4, 6-Hexanetriol) caused a significant disruption of actin and tubulin organization, as adjudged by the localization of Sac6-GFP (an actin binding protein) and Tub1-GFP (Alpha-Tubulin) respectively (**Figure 2.7A and data not shown**). This disruption of cytoskeleton organization occurred within two minutes of 1,6-Hexanediol treatment.

Interestingly, Sac6 changed localization twice – once rapidly to go from punctate to diffuse; and later from diffuse to into rod-like structures in the cytoplasm (**Figure 2.7A**). Membrane bound cellular structures such as the Endoplasmic Reticulum (ER, observed using Get1-GFP) and mitochondria (observed using Atp1-GFP) were not affected by 1,6-Hexanediol. These observations suggest that 1,6-Hexanediol will disrupt a variety of cellular structures.

We also examined how 1,6-Hexanediol affected mammalian cells. Two observations suggest 1,6-Hexanediol affects cellular viability and gross cellular morphology of mammalian cells. First, HeLa cells demonstrated a 40% reduction in cellular viability as assessed by RealTime-Glo MT Cell Viability Assay following exposure to 3.5% 1,6-Hexanediol as compared to untreated cells. Second, HeLa cells exposed to 1,6-Hexanediol demonstrated rapid widespread morphological changes within seconds including membrane blebbing (**Figure 2.7B**). Thus, owing to the adverse effects of 1,6-Hexanediol on cell viability and morphology, it's use in mammalian cells in culture may be complicated.

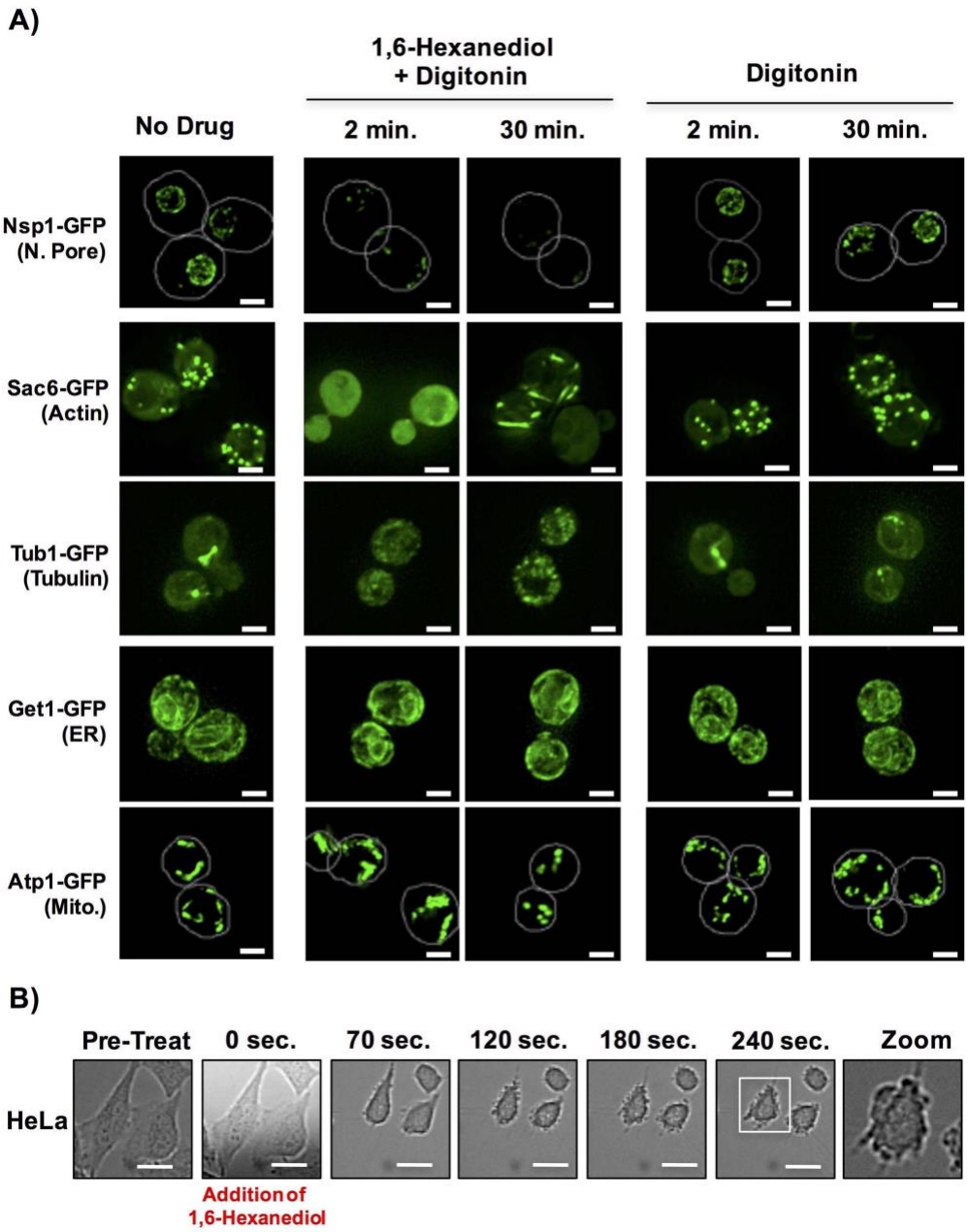

A)
1,6-Hexanediol
+ Digitonin
Digitonin
No Drug
2 min.
30 min.
2 min.
30 min.
Nsp1-GFP
(N. Pore)
Sac6-GFP
(Actin)
Tub1-GFP
(Tubulin)
Get1-GFP
(ER)
Atp1-GFP
(Mito.)
B)
Pre-Treat
0 sec.
70 sec.
120 sec.
180 sec.
240 sec.
Zoom
HeLa
Addition of
1,6-Hexanediol

Figure 2.7: 1,6-Hexanediol disrupts many cellular structures in yeast and alters cellular morphology in HeLa cells

(A) The localization of various GFP-tagged proteins is shown either without treatment (No Drug); in the presence of 10% 1,6-Hexanediol and 10 µg/mL Digitonin for 2 minutes or 30 minutes; or just 10 µg/mL Digitonin for 2 minutes or 30 minutes. For Nsp1-GFP and Atp1-GFP, white lines depict cell boundaries. N. Pore = nuclear pore. ER = Endoplasmic Reticulum. Mito. = Mitochondria. Get1, Guided Entry of Tail-anchored proteins (YGL020C). Atp1, ATP synthase (YBL099W). Scale bar is 2 µm. (B) Time-lapse microscopy of HeLa cells treated with 3.5% 1,6-Hexanediol reveals alterations in cellular morphology. HeLa cells were exposed to 3.5% 1,6-Hexanediol (time zero) and imaged for 5 minutes using a 20X objective. Scale bars are 20 µm.

1,6-Hexanediol induces stress granule formation in yeast

To test the effect of 1,6-Hexanediol on already assembled stress granules and P-bodies in yeast, we glucose starved yeast cells for 15 minutes, followed by incubation with 10% 1,6-Hexanediol as previously described (Kroschwald et al. 2015). Markers of stress granules and P-bodies changed localization twice (similar to Sac6-GFP above). First, both assemblies quickly disassembled within two minutes, however, after ten minutes, they reappeared, often completely co-localized (**Figure 2.8A**). This is in contrast to a previous report that suggested that yeast P-bodies, but not stress granules, are sensitive to 1,6-Hexanediol treatment under glucose starvation conditions (Kroschwald et al. 2015). We observed that P-bodies formed upon 1,6-Hexanediol treatment are less intense than the P-bodies formed under glucose starvation conditions and this may have limited their earlier detection.

By using strains with tagged components of stress granules and P-bodies, we observed that in the absence of any other stress, 1,6-Hexanediol (and not a Digitonin control) also induced robust Pab1-GFP (a stress granule marker) granule formation within 10 minutes of treatment (**Figure 2.8B**). Interestingly, these Pab1-GFP foci also overlapped with Edc3-mCherry – a P-body protein (**Figure 2.8B**). Thus, exposure of yeast cells to 1,6-Hexanediol is an effective inducer of stress granule-like assemblies.

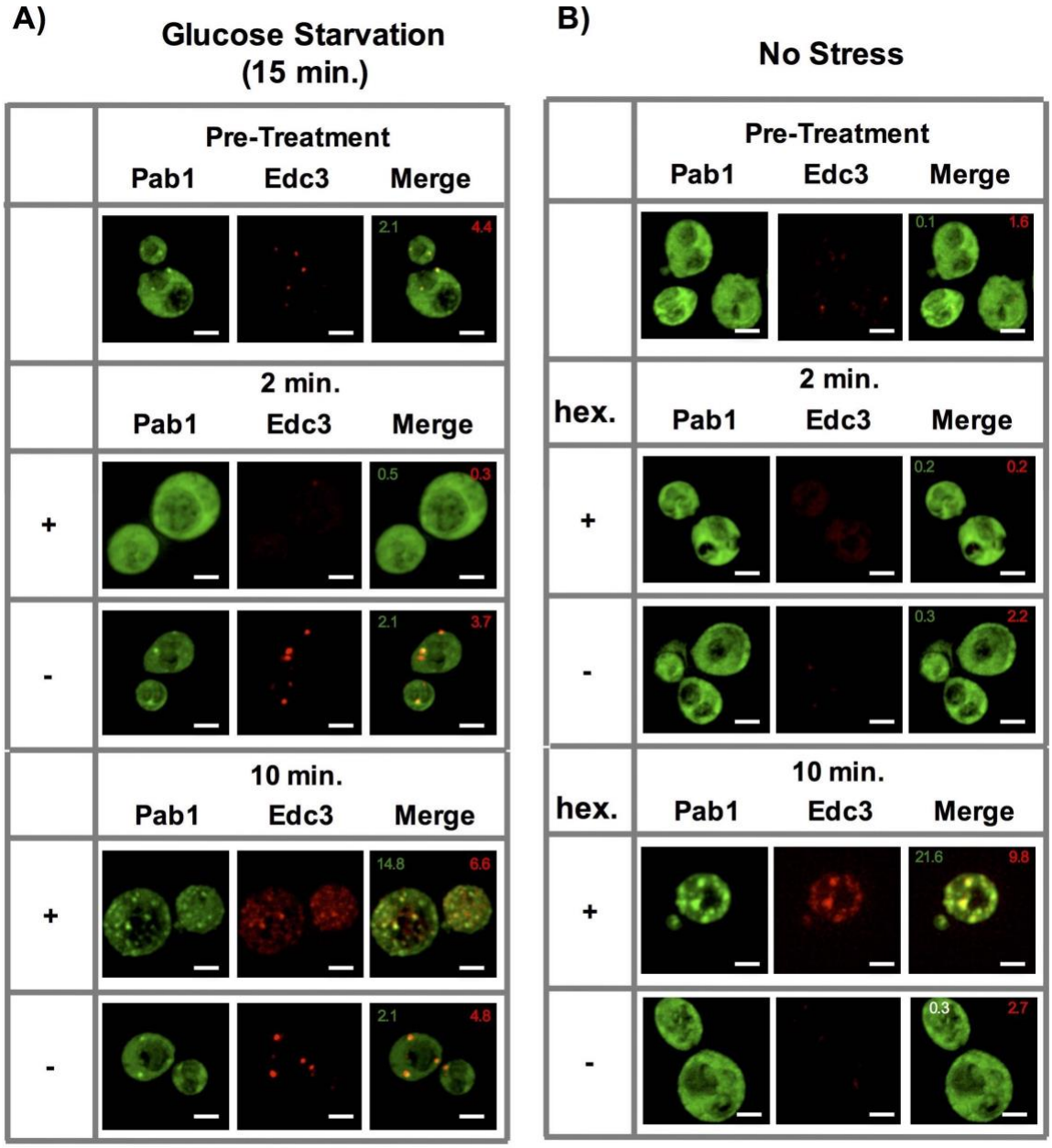

A)
Glucose Starvation
(15 min.)
Pre-Treatment
Pab1
Edc3
Merge
2.1
4.4
2 min.
Pab1
Edc3
Merge
+
0.5
0.3
-
2.1
3.7
10 min.
Pab1
Edc3
Merge
+
14.8
6.6
-
2.1
4.8
B)
No Stress
Pre-Treatment
Pab1
Edc3
Merge
0.1
1.6
hex.
2 min.
Pab1
Edc3
Merge
+
0.2
0.2
-
0.3
2.2
hex.
10 min.
Pab1
Edc3
Merge
+
21.6
9.8
-
0.3
2.7

Figure 2.8 Effect of 1,6-Hexanediol on stress granules and P-bodies in yeast

(A) Localization of Pab1-GFP and Edc3-mCherry prior to treatment with (Pre-treat) or after 2 minutes or 10 minutes of 10 µg/mL Digitonin ± 10% 1,6-Hexanediol following 15 minutes of glucose starvation. All cells continue to be in media lacking glucose throughout the course of the experiment. (B) Same as (A), after 10 µg/mL Digitonin ± 10% 1,6-Hexanediol treatment only. Numbers in green (or white) indicate average number of stress granules/ cell. Numbers in red indicate average number of P-bodies/ cell. All scale bars are 2 µm.

1,6-Hexanediol induces stress granule formation in HeLa cells

To determine how pre-existing mammalian stress granules react to 1,6-Hexanediol, we induced stress granules with $NaAsO_2$ in HeLa cells and then treated cells with 1,6-Hexanediol. Similar to yeast stress granules, we observed an initial reduction in total stress granule numbers, which coincided with alterations in cellular morphology; however, after 10 minutes, stress granules re-appeared (**Figure 2.9A**). These effects of 1,6-Hexanediol were also observed in U-2 OS cells, suggesting this is not a cell specific phenotype (data not shown).

In the absence of any other stress, 1,6-Hexanediol treatment resulted in the formation of stress granule-like assemblies. These assemblies stained positive for stress granule markers G3BP1 and PABP1 (**Figure 2.9B**). Thus, similar to yeast cells, 1,6-Hexanediol treatment of mammalian cells in culture can induce stress granule-like assemblies.

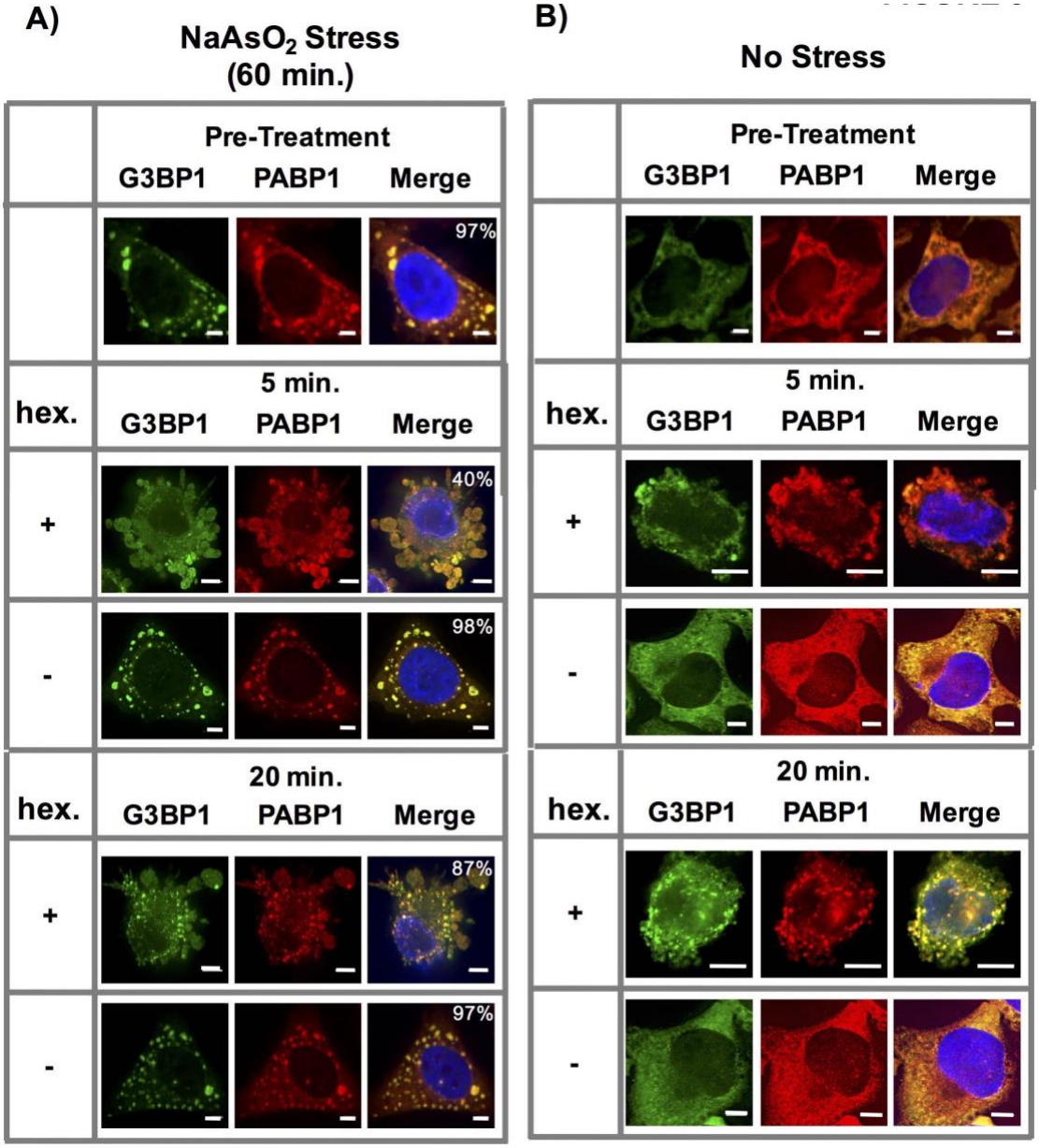
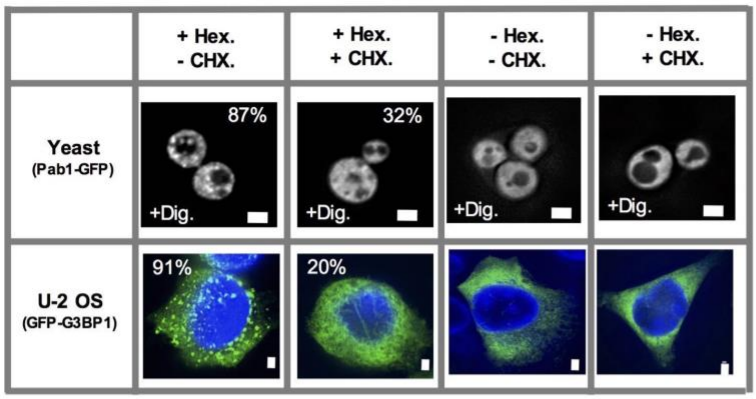

A)
NaAsO2 Stress
(60 min.)
B)
No Stress
Pre-Treatment
G3BP1
PABP1
Merge
97%
hex.
5 min.
+
40%
-
98%
20 min.
+
87%
-
97%

Figure 2.9: 1,6-Hexanediol induces formation of stress granule-like assemblies in HeLa cells

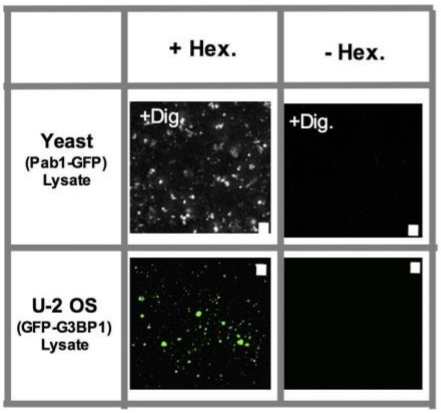
(A) Affects of 1,6-Hexanediol on $NaAsO_2$-induced stress granules in HeLa cells. HeLa cells were stressed with $NaAsO_2$ for 60 minutes. Addition of 3.5% 1,6-Hexanediol in the presence of $NaAsO_2$ was added at indicated time points. For all experiments, HeLa cells were fixed and co-stained for G3BP1 (FITC labeled secondary) and PABP1 (Alexa-647 labeled secondary). Percentages refer to cells with stress granules defined as G3BP1 and PABP1 double positive foci (N=3 per condition, 50 cells/ replicate). (B) 1,6-Hexanediol induces stress granule formation in HeLa cells. HeLa cells were exposed to media containing 3.5% 1,6-Hexanediol for indicated time periods. All scale bars are 5 μm.

Several observations suggest the assemblies induced by 1,6-Hexanediol are bona fide stress granules. First, 1,6-Hexanediol induces the formation of assemblies that contain known stress granule proteins in both yeast and mammalian cells (HeLa and U-2 OS cells) (**Figure 1.10A**). Second, 1,6-Hexanediol induced assemblies are sensitive to cycloheximide treatment, a drug known to block stress granule assembly, in both yeast and mammalian cells (**Figure 2.10A**). Third, 1,6-Hexanediol-induced assemblies are stable in lysates suggesting that the material properties of these assemblies contain a relatively stable set of interactions similar to normal stress granules (**Figure 2.10B**). Fourth, eiF2 phosphorylation is induced to a similar extent by 1,6-Hexanediol compared to $NaAsO_2$ (**Figure 2.10C**). Finally, FRAP analysis of 1,6-Hexanediol-induced granules in mammalian cells indicates that 1,6-Hexanediol induced stress granules behave similar to normal stress granules (**Figure 2.10D**).

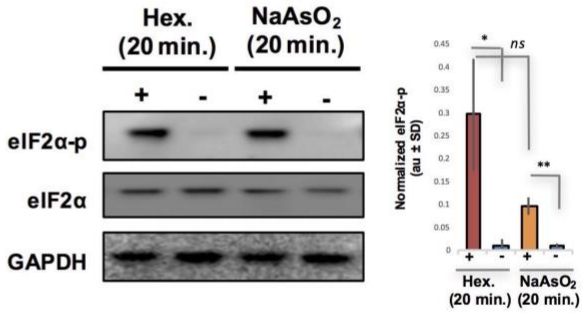
A)


+ Hex.
- CHX.
+ Hex.
+ CHX.
- Hex.
- CHX.
- Hex.
+ CHX.
Yeast
(Pab1-GFP)
87%
+Dig.
32%
+Dig.
+Dig.
+Dig.
U-2 OS
(GFP-G3BP1)
91%
20%

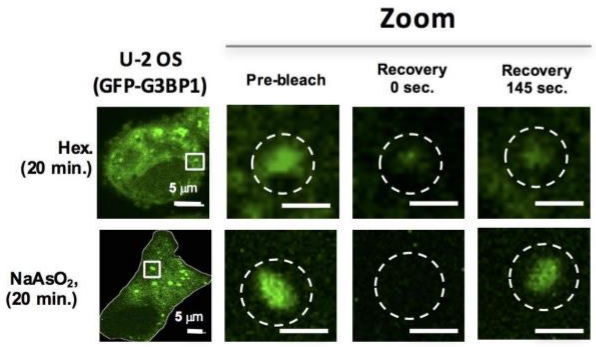
B)


+ Hex.
- Hex.
Yeast
(Pab1-GFP)
Lysate
+Dig.
+Dig.
U-2 OS
(GFP-G3BP1)
Lysate

C)


Hex.
(20 min.)
NaAsO2
(20 min.)
+
-
+
-
eIF2α-p
eIF2α
GAPDH
Normalized eIF2α-p
(au ± SD)
*
ns
**
Hex.
(20 min.)
NaAsO2
(20 min.)

D)


Zoom
U-2 OS
(GFP-G3BP1)
Pre-bleach
Recovery
0 sec.
Recovery
145 sec.
Hex.
(20 min.)
NaAsO2,
(20 min.)
5 µm
5 µm

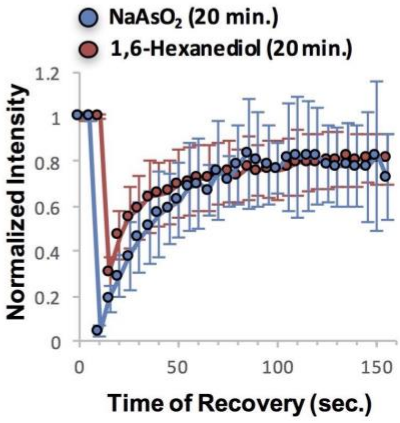

NaAsO2 (20 min.)
1,6-Hexanediol (20 min.)
Normalized Intensity
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
Time of Recovery (sec.)

Figure 2.10: Stress granules induced by 1,6-Hexanediol are cyclohexamide sensitive, stable in lysates, and dynamic

(A) 1,6-Hexanediol induced stress granules are cyclohexamide sensitive in yeast and U-2 OS cells. Yeast were grown in the presence of 10% 1,6-Hexanediol and 10 µg/mL Digitonin for 10 minutes or just 10 µg/mL Digitonin ± cyclohexamide (100 µg/mL). U-2 OS cells were exposed to media containing 3.5% 1,6-Hexanediol for 20 minutes ± cyclohexamide (10 µg/mL). Percentages refer to cells with stress granules (N=3 bioloigcal replicate, 50 cells per replicate). (B) Yeast and U-2 OS 1,6-Hexanediol induced granules are stable in lysates. Yeast were grown in presence of 1,6-Hexanediol and 10 µg/mL Digitonin for 10 minutes or just 10 µg/mL Digitonin prior to lysis. Mammalian cells were grown in the presence or absence of 3.5% 1,6-Hexanediol containing media for 15 minutes prior to lysis. (C) Western blot for eiF2α phosphorylation status following +/- exposure to 1,6-Hexanediol (3.5%) or $NaAsO_2$ (0.5mM) stress for 20 minutes. Normalized to eiF2α and GAPDH (N=3, au = arbitary units, SD = standard deviation). (p-value: *<0.01; **<0.001; ns = not significant). (D) Granules shown prior to photobleaching, at 0 seconds, and at 145 seconds after photobleaching. Cells were treated with either 1,6-Hexanediol or $NaAsO_2$ for 20 minutes. Graph shows recovery curves as an average of 6 granules ± standard deviation for each respective condition. Abbreviations: Hex., 1,6-Hexanediol; Dig., Digitonin; CHX, cyclohexamide. All scale bars are 2 µm unless otherwise noted.

Stress granules disassemble through multiple discrete steps

Our data suggests a model for stress granule assembly wherein mRNPs oligomerize to form stable cores and over time individual cores then dock with one another through a more dynamic shell. We reasoned stress granule disassembly may occur in a reverse process where a less stable shell may dissipate initially followed by core disassembly or clearance. To test this prediction, we examined stress granule disassembly using live cell imaging. Several observations suggest that, like stress granule assembly, stress granule disassembly occurs in a multi-step process.

First, live cell imaging of stress granule disassembly shows that stress granules start out as large complexes and disassemble within a narrow time window. Within the limitations of our ability to detect stress granule disassembly, we observe for large stress granules, disassembly appears to occur in two steps: first, larger stress granules break into smaller foci followed by disassembly and/or clearance of these smaller foci. During disassembly, we observe the appearance of filamentous structures emerging at the edges of the disassembling stress granule; however, we should point out that due to the resolution of our microscope we are unable to characterize the nature of these structures (**Figure 2.11B, 2.11C**). Interestingly, once stress granules disassemble into smaller assemblies, these small assemblies become microscopically undetectable within a similar time scale as required for their initial detection during assembly (**Figure 2.11A, 2.11B**).

Second, examination of stress granule disassembly reveals non-uniform distribution within stress granules undergoing disassembly wherein areas of relative concentration appear to persist once the surrounding shell structures have dissolved (**Figure 2.11B**). Finally, we observe stress granules are stable in lysates at the same time points that disassembly is occurring (**Figure 2.11C**).

Taken together, stress granule disassembly may occur through multiple steps wherein RNA is titrated out of stress granules and into translation, leading to structural instability and subsequent disassembly of a larger stress granule complex into smaller core structures that are then disassembled or cleared by autophagy.

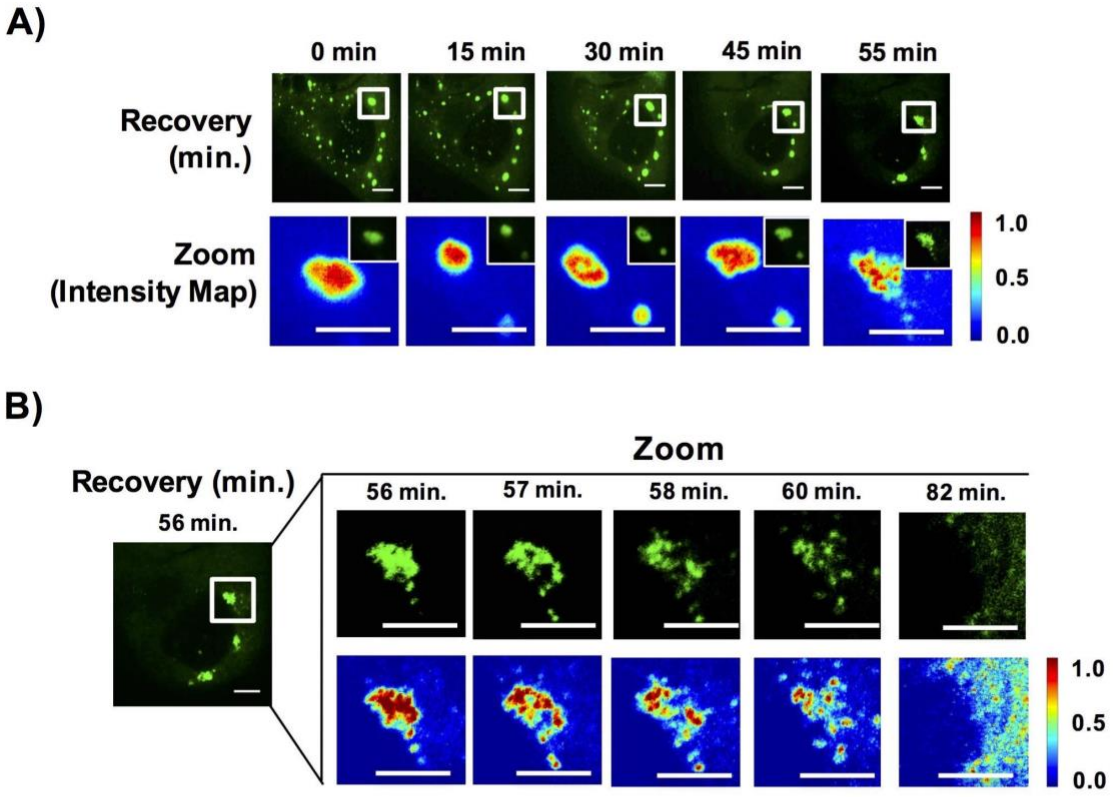

A)
0 min
15 min
30 min
45 min
55 min
Recovery (min.)
Zoom (Intensity Map)
1.0
0.5
0.0
B)
Zoom
Recovery (min.)
56 min.
56 min.
57 min.
58 min.
60 min.
82 min.
1.0
0.5
0.0
C)
Zoom, In vivo
Ex vivo
Recovery (min.)
90 min.
DV
G3BP1
PABP1
G3BP1

FIGURE 2.11: Stress granules disassemble through multiple discrete steps

(A) Time-lapse imaging of stress granule disassembly in normal media following 60 minutes of $NaAsO_2$ stress using a 100X objective. Intensity map represents relative gray scale intensity of zoomed inset. (B) Same as (A) imaged during stress granule fracturing. (C) *In vivo* GFP-G3BP1 stress granules were stained with PABP1 (Alexa-647 labeled secondary) and *ex vivo* GFP-G3BP1 stress granules cores imaged by deconvolution microscopy (DV) following 90 minutes of recovery from $NaAsO_2$ stress. All scale bars are 5 μm.

DISCUSSION

In this chapter, we set out to distinguish between two different models to explain the processes of stress granule formation, maturation, and disassembly. In the first model, stress granules form by a loose condensation of IDR containing proteins to form an initial phase separated, dynamic structure and over time stable cores mature within this dynamic structure. This model is suggested by recent findings *in vitro*, demonstrating LLPS can be driven by IDRs and these LLPS subsequently mature to form a second, less dynamic phase (Lin et al., 2015; Molliex et al., 2015; Murakami et al., 2015; Patel et al., 2015; Zhang et al., 2015). In the second model, we considered that stress granules initially condense to form stable cores, possibly with a nascent shell layer, and these cores merge into larger structures through interactions between the shell structures.

Several lines of evidence argue that stress granule assembly involves the early formation of stable core structures that then assemble into larger stress granules, each of which can contain multiple cores, surrounded by a less dense shell layer (**Figure 2.12**). First, as soon as stress granules are observed in cells, they are stable in lysates (**Figure 2.2**). Second, the size distribution of the stable components of stress granules in lysates is similar from 15 to 120 minutes of stress, arguing that stress granule core formation is not a late step in stress granule assembly (**Figure 2.4**). Third, as soon as technically possible to assess, the FRAP behavior, as observed by the GFP-G3BP1 mobile fraction, of stress granules remains unchanged from 30 minutes to 120 minutes

of stress (**Figure 2.5**). This also argues that stress granules are not undergoing a dramatic change in biochemical state, and is inconsistent with models where stress granules initially consist of a uniform and dynamic LLPS, that then matures into a mature biphasic stress granule. Finally, when examined using super resolution microscopy, we observe granules exhibit a heterogeneous distribution of protein even at early time points (**Figure 2.3**). Taken together, these observations are consistent with granules assembling from the oligomerization of individual mRNPs into stable core complexes. It should be noted we cannot rule out the formal caveat that stress granules initially condense into a weak dynamic assembly, that immediately transitions to a stable assembly, or that stress granule shells and cores form independently as two immiscible liquid phases which later fuse as has been suggested to explain the assembly of nucleolus (Feric et al., 2016). We hypothesize that these stable core assemblies could provide a structural platform, due to the high concentration of IDRs on stress granule components to then rapidly phase separate a dynamic shell that grows as a result of fusion with other small granules and surface exchange with an increasing pool of untranslating mRNPs.

The assembly of stress granule cores as an initial step in assembly solves a conundrum. The issue is that IDRs are thought to interact with each other to drive LLPS often by weak dynamic interactions, proposed to involve Arg-Aromatic interactions in some cases (Nott et al., 2015; Pak et al., 2016). Since these interactions are limited to few amino acids, one anticipates that any given interaction would be expected to have low specificity, and as such would be expected to interact with many proteins when

dispersed in the cytosol. Thus, for such weak and promiscuous interactions to drive *a priori* the assembly of a stress granule would be difficult since the interactions between stress granule components would always be in competition with other cytosolic factors. When the initial step in stress granule assembly is formation of a stable core, which by definition involves relatively stable interactions, the core structure could then provide a high local concentration of IDRs on stress granule components, which could then be envisioned to trigger a local phase separation, possibly what we refer to as the shell phase of a stress granule. This general principle is also seen in the formation of the nuclear pore, wherein the creation of a high local concentration of FG repeat proteins by the structure of the nuclear pore, allows for a phase transition within the nuclear pore itself, even if the interactions between the FG repeats are relatively weak and non-specific (Frey and Görlich, 2007; Hülsmann et al., 2012; Schmidt and Görlich, 2016).

This work provides two observations that suggest a difference between the process of *in vitro* LLPS driven by IDRs on RNA binding proteins and the formation of stress granules in cells. First, although low temperatures promote LLPS driven by IDRs of RNA binding proteins (Molliex et al., 2015; Nott et al., 2015), we observed that stress granules formed less efficiently at lower temperatures, despite efficient polysome disassembly (**Figure 2.6**). Second, although 1,6-Hexanediol can prevent LLPS by IDRs from RNA binding proteins (Molliex et al, 2015), we observed that 1,6-Hexanediol effectively induces stress granules in both yeast and mammals (**Figures 2.8, 2.9, and 2.10**). These observations suggest that the process of stress granule assembly in cells is not primarily driven by weak, dynamic homo- and heterotypic interactions between

IDRs on RNA binding proteins analogous to the LLPS driven by these protein domains *in vitro*. However, we hypothesize that weak, dynamic interactions between IDRs could be important in forming the shell structure of stress granules, once the core is nucleated by more stable interactions.

In contrast to our results, 1,6-Hexanediol has been reported to dissolve yeast P-bodies and mammalian stress granules (Kroschwald et al., 2015). We do not understand the basis for the difference between our results and those published earlier, although it could be due to differences in cell handling, growth conditions, or analysis. In any case, the sensitivity of a cellular structure to 1,6-Hexanediol does not appear sufficient to conclude the structure forms by LLPS. We suggest this cautionary note for three reasons. First, we observed that some, but not all, cellular structures were sensitive to 1,6-Hexanediol, including well described cytoskeleton elements (**Figure 2.7**). Second, one should anticipate that, depending on the chemical nature of interactions driving assembly, many different types of assemblies could be sensitive to 1, 6-Hexanediol. Finally, one should also expect LLPS assemblies can form by many types of promiscuous or specific interactions, some of which would be expected to be resistant to 1,6-Hexanediol.

Previous work has suggested that small stress granules that form early in a stress response require microtubules and their associated motors to form larger stress granules (Chernov et al., 2009; Fujimura et al., 2009; Ivanov et al., 2003; Loschi et al., 2009). In those experiments, treatment of cells with nocadazole, which depolymerizes

microtubules, prevents the fusion of the small initial stress granules into larger assemblies. Based on these results, a reasonable model is that cores are brought together to form larger assemblies by movement on microtubules. Disassembly may involve a reverse process where the microtubule network may facilitate retrograde transport of mRNPs away from a disassembling stress granule.

We observe stress granule disassembly occurs in a reverse process where a less stable shell dissolves initially followed by core disassembly and or clearance by autophagy. This observation has several corollaries. First, upon re-establishment of translation, mRNA is thought to be in rapid equilibrium between the cytosol and stress granules. This exchange appears to influence stress granule structural integrity and may account for titration of select RNAs into translation. The lag in the clearance of granule cores may reflect the requirement of a myriad of ATP-dependent remodeling complexes (e.g. heat shock 70 or p97/VCP AAA-ATPase complexes (Buchan et al., 2013; Walters et al., 2015)) and could serve as a cytoprotective mechanism to acutely re-nucleate stress granules if the cell re-encounters stress.

In mammalian cells, a biphasic stress granule architecture provides multiple layers of functionality within the larger tunable stress-specific RNA transcriptome. First, a shell in mammalian cells provides a scaffold for dynamic exchange of select RNAs out of the translational pool. This could represent a thermostatic mechanism where select RNAs (possibly important housekeeping genes) could be still translated under stress without overtly outcompeting the translation of stress-specific messages essential for

survival under stress. These shell-specific RNAs would be expected to be primed for reentry into translation during recovery and may also represent a metastable structural scaffold that fractures during stress granule disassembly. Meanwhile, further compartmentalization into less dynamic cores could provide an additional layer of sequestration of select messages. This could possibly serve as a biochemical sorting mechanism where “fit” mRNPs could be remodeled under stress by concentrating similar RNPs, while “unfit” mRNPs could be sorted for decay or not remodeled to reenter the translational pool.

In degenerative disease, this multistep process might become altered leading to aberrant formation of granules. Hyper-nucleation or inefficient clearance of stress granule cores may facilitate a pathological transition by lowering the free energy landscape to seed an -amyloid aggregate by providing a continued platform for nucleation

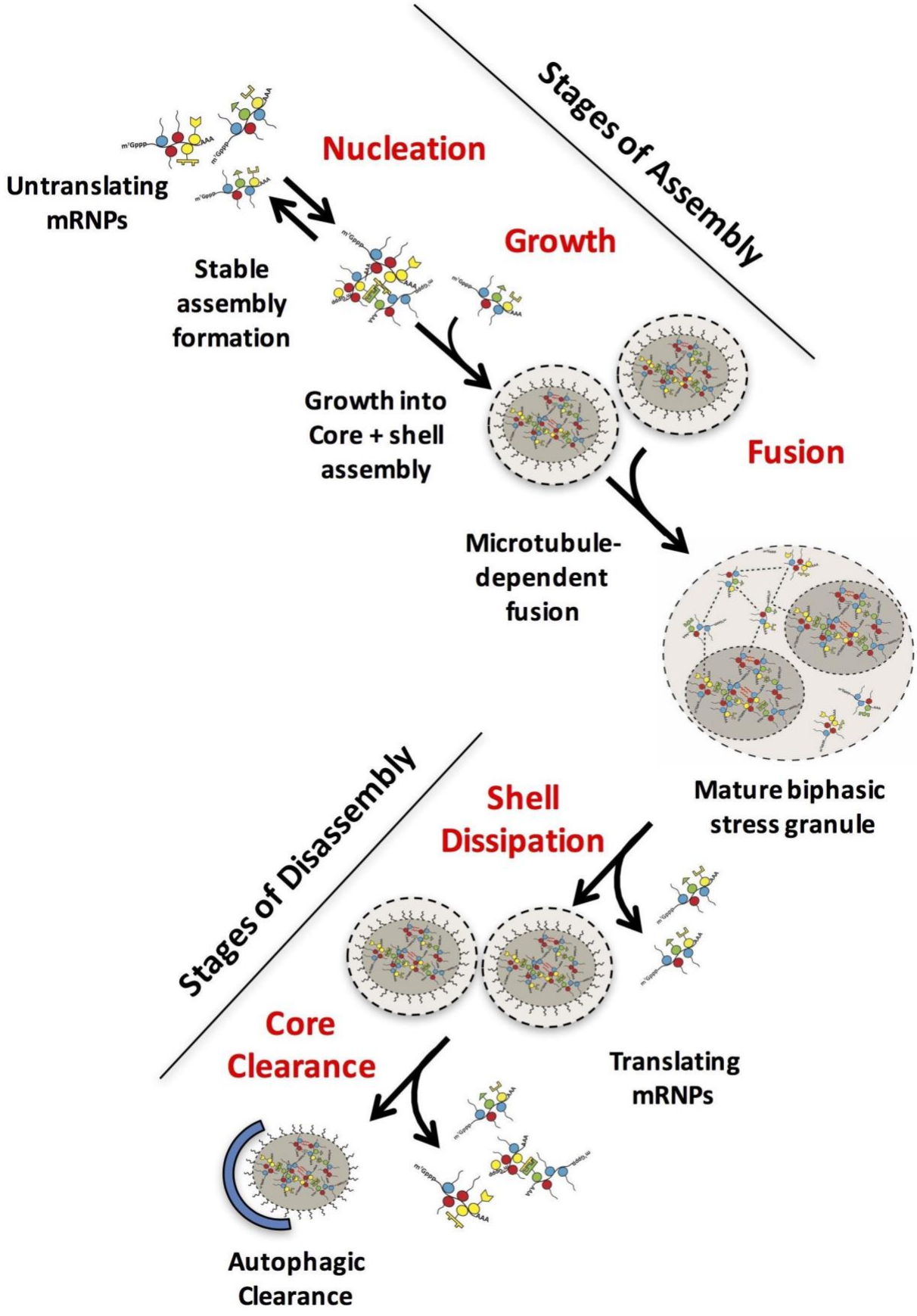

Stages of Assembly
Nucleation
Untranslating mRNPs
Growth
Stable assembly formation
Growth into Core + shell assembly
Fusion
Microtubule-dependent fusion
Mature biphasic stress granule
Stages of Disassembly
Shell Dissipation
Translating mRNPs
Core Clearance
Autophagic Clearance

FIGURE 2.12: Model for distinct stages of stress granule assembly and disassembly

Possible steps in granule assembly and disassembly are shown. Untranslating mRNPs nucleate an early stable mRNP (core) complex and grow to rapidly include an early core/ shell (biphasic) assembly through an ATP-dependent, microtubule-independent process. Fusion of biphasic stress granules forms a larger higher order mature assembly in part by a microtubule-dependent process. Disassembly is likely to occur through shell dissipation by exchange of weakly associated granule (shell) mRNPs into a recovering translational mRNP pool. More stable core assemblies may then be disassembled by ATP-requiring remodeling complexes or autophagy. Dashed lines between mRNPs represent weak physical interactions in phase-separated shell (light gray). Red wavy lines represent strong interactions between IDRs in stable cores (darker gray).

MATERIALS AND METHODS

Yeast and mammalian cell growth conditions

All GFP-tagged strains shown in Figure 7 (Sac6-GFP, Tub1-GFP, Nsp1-GFP, Get1-GFP and Atp1-GFP) were taken from the yeast GFP collection. These strains were grown in minimal media supplemented with a complete set of amino acids and 2% Dextrose at 30°C. For the experiments presented in Figure 8, BY4741 yeast was transformed with a single plasmid expressing Pab1-GFP and Edc3-mCherry (pRP1657). These strains were grown in –Ura media with 2% Dextrose at 30°C. Experiments in Figure 10A were done with 4741 transformed with a plasmid expressing Pab1-GFP (pRP 1363). Human osteosarcoma U-2 OS (expressing GFP-G3BP1, mRFP-DCP1a) (Kedersha et al., 2008; Ohn et al., 2008) and HeLa cells, maintained in DMEM, High Glucose, GlutaMAX with 10% fetal bovine serum, 1% penicillin/streptomycin, and 1mM sodium pyruvate at 37°C/5% CO_2 were used for all biochemical experiments, stress granule purification, and imaging experiments.

Stressing cells

Yeast: NaN_3 stress: Cells were treated with 0.5% NaN_3 for 30 minutes at 30°C (unless otherwise stated). Mammalian: $NaAsO_2$ stress: Cells were treated with 0.5mM $NaAsO_2$ at 37°C/5% CO_2 (unless otherwise stated). Thapsigargin stress: Cells were treated with

100nM Thapsigargin in DMSO for 1hr at 37°C/5% CO_2. Osmotic stress: Cells were treated with 375M Sorbitol at 37°C/5% CO_2 for 1hr.

Analysis of yeast and mammalian stress granule cores

A 'granule-enriched' fraction was prepared from yeast and mammalian cells as previously described (Jain et al., 2016). For yeast granule isolation, 50mL of yeast were grown. For mammalian granule isolation, stressed U-2 OS cells were collected from 15cm plates, pelleted and snap frozen in liquid nitrogen. Cell pellets (re-suspended in lysis buffer) were used to prepare granule enriched fractions. In lysate immunofluorescence and oligo-(dT) staining of stress granule cores was performed as previously described (Jain et al., 2016). 1,6-Hexanediol granules were isolated from U-2 OS cells grown in the presence of 3.5% 1,6-Hexanediol for 15 minutes.

Nanosight Particle Size Analysis

Stress granule core enriched fractions were collected at respective time points and diluted in lysis buffer (1:2). Particles were analyzed using the NanoSight Nanoparticle Tracking Analysis system NS300 (Malvern) with syringe pump, a 488nm laser, and a sCMOS camera. Five videos of 60 seconds were collected for each sample using the 488nm laser and analyzed by NTA 3.0 software. Maximal concentration of particles (particles/ nm) was used for normalization.

FRAP analysis of U-2 OS GFP-G3BP1 cells

U-2 OS cells were cultured on glass bottom 35mm dishes. FRAP experiments were performed at indicated times as previously described (Jain et al., 2016).

Digitonin, Cyclohexamide and 1,6-Hexanediol Treatment in Yeast

Cells were grown to an OD_{600} of 0.4-0.7 in minimal media supplemented with the appropriate amino acids, at 30°C. These cells were then imaged for the 'No Drug' or 'Pre-treat' controls. These cultures were then pelleted at 3220Xg for 1 minute, washed once with medium containing 10 µg/ ml Digitonin ± 10% 1,6-Hexanediol (similar to the concentrations used in Kroschwald et al., 2015), then resuspended in the same medium and returned to a shaker at 30°C. Cells were then imaged after 2 minutes, 10 minutes, 20 minutes and 30 minutes. For glucose starvation experiments, cells at log phase were pelleted at 3220Xg for 1 minute, washed once in media without Dextrose, then resuspended in media without Dextrose and returned to a shaker at 30°C. After 15 minutes, these cells were then pelleted again at 3220Xg for 1 minute and washed once and then resuspended in dextrose free media containing 10 µg/ ml Digitonin ± 10% 1,6-Hexanediol. These cells were then returned to a shaker at 30°C, and imaged after 2, 10, 20 and 30 minutes. All cyclohexamide treatments were performed at a final concentration of 100 µg/mL added to media at the same time as Digitonin ± 1,6-Hexanediol.

1,6-Hexanediol Treatment in HeLa and U-2 OS cells

For treatment with 1,6-Hexanediol, complete media (DMEM supplemented with 10% FBS, 1% penicillin/streptomycin) was prepared containing 3.5% 1,6-Hexanediol. Media was exchanged and replaced with 3.5% 1,6-Hexanediol or complete media for indicated time periods. HeLa cells were stressed with $NaAsO_2$ (0.5mM) for 1hr. Media was exchanged and replaced with 3.5% 1,6-Hexanediol and $NaAsO_2$ (0.5mM) or $NaAsO_2$ (0.5mM) only containing media at indicated time points. For all experiments, HeLa cells were fixed (4% paraformaldehyde) and co-stained for PABP1 (detected by Cy5 labeled secondary) and G3BP1 (detected by FITC labeled secondary). The same concentrations for staining were used in all experiments for both primary (1:200) and secondary (1:400) antibodies. For U-2 OS experiments, all experiments were conducted with 3.5% 1,6-Hexanediol, cyclohexamide (10 µg/mL), and $NaAsO_2$ (0.5mM) prepared in complete media.

Cellular Viability Measurements

Cellular viability was measured using RealTime-Glo MT Cell Viability Assay (Promega) according to manufacture's protocol. HeLa cells were exposed to complete media containing 3.5% 1,6-Hexanediol for indicated time periods. Following exposure, normal media was exchanged and cells were incubated in presence of Realtime-Glo substrates for 1hr or 14hrs at 37°C. Luminescence measurements were performed using a Victor^3 plate reader (PerkinElmer).

Microscopy and Image Analysis

All yeast and mammalian images were acquired using a DeltaVision Elite microscope with a 100X objective using a PCO Edge sCMOS camera unless otherwise stated. ≥3 images were taken for each experiment comprising of 8 Z-sections each. Stress granule cores were visualized using Deltavision. All images were analyzed using ImageJ. In Figure 2, three independent experiments were performed for each time point. Granule core numbers were calculated using Deltavision, maximal image projections were determined using ImageJ, and thresholding was kept constant across all replicates and time points. Plotted are mean GFP-G3BP1 postiive particles (standard deviation) normalized to particles at time zero. The x-axis was zeroed to 50 to allow for easier comparison within the graph. In Figure 9, percent HeLa cells (N=3, 50 cells/ replicate) with stress granules (G3BP1 and PABP1 double positive foci) were calculated for each condition.

Live cell imaging was performed using a Nikon Spinning Disk Confocal microscope outfitted with an environmental chamber with O_2, CO_2, temperature, and humidity control. All images were acquired using a 100X objective with a 2x Andor Ultra 888 EMCCD camera. For temperature experiments, U-2 OS cells grown at 37°C in 35mm dishes were transferred to environmental chamber at respective temperatures and allowed to equilibrate to respective temperatures for 10 minutes prior to image acquisition. For disassembly experiments, cells were washed twice in normal media and

permitted to recover during image acquisition. All images were acquired by exciting for 200ms using a 488nm laser at a gain of 200 (unless otherwise stated). In Figure 2, GFP-G3BP1 foci were determined from thresholded (Otsu) images acquired at 20s time intervals and plotted as a percentage of peak foci detected for 9 total cells acquired from three independent experiments. In Figure 7, bright field movie was acquired using a Nikon spinning disk microscope using a 20X air objective (HBO arc lamp). Time and event stamps and scale bars were added using ImageJ.

SIM (Structured Illumination Microscopy) was performed using the Nikon N-SIM microscope system run in 3D SIM mode with a Nikon 100X objective at a 1.49NA. Images were acquired using an Andor iXon DU897 EM-CCD camera. GFP-G3BP1 was visualized by exciting the sample using a 488nm laser at a gain of 300. Samples were mounted using ProLong Diamond Mounting media incubated overnight at room temperature. Sample was excited for 600ms/ image. Image reconstruction was performed in Nikon NIS Elements.

CHAPTER 3

THE STRESS GRANULE TRANSCRIPTOME REVEALS PRINCIPLES OF mRNA ACCUMULUATION IN STRESS GRANULES[2]

SUMMARY

Stress granules are mRNA-protein assemblies formed from nontranslating mRNAs. Stress granules are important in the stress response and may contribute to some degenerative diseases. Here we describe the stress granule transcriptome of yeast and mammalian cells through RNA-Seq analysis of purified stress granule cores and smFISH validation. While essentially every mRNA, and some ncRNAs, can be targeted to stress granules, the targeting efficiency varies from <1% to >95%. mRNA accumulation in stress granules correlates with longer coding and UTR regions and poor translatability. Quantifying the RNA-Seq analysis by smFISH reveals only 10% of bulk mRNA molecules accumulate in mammalian stress granules, and only 185 genes have more than 50% of their mRNA molecules in stress granules. These results suggest stress granules may not represent a specific biological program of mRNP assembly, but instead form by condensation of nontranslating mRNPs in proportion to their length and lack of association with ribosomes.

---

[2] This chapter has been modified from the following publication: Khong, A.*, Matheny, T.*, Jain, S., Mitchell, S.F., Wheeler, J.R., Parker, R. (2017). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. Mol. Cell *68*, 808–820. *These authors contributed equally

INTRODUCTION

Stress granules (SGs) are non-membrane bound assemblies of RNA and protein that form when translation initiation is limited (Protter and Parker, 2016; Buchan and Parker, 2009). SGs also share many protein components with neuronal granules, and mutations that increase SG formation or perturb SG clearance are implicated in degenerative diseases such as ALS and multisystem proteinopathy, where aberrant SG-like assemblies form (Buchan et al., 2013; Dewey et al., 2012; Kim et al., 2013b; Li et al., 2013; Mackenzie et al., 2017; Ramaswami et al., 2013). SGs are thought to assemble via nontranslating mRNAs serving as scaffolds for RNA binding proteins, which interact with each other through a variety of protein-protein interactions (Panas et al., 2016; Protter and Parker, 2016).

By super-resolution microscopy, SGs show denser regions of proteins and mRNAs (based on oligo(dT) FISH), referred to as cores, which can be biochemically purified (Jain et al., 2016). Cores are linked together by less concentrated regions of SG components referred to as a “shell”, although whether the composition of the shell is different from cores has not been addressed (Jain et al., 2016). Purification of SG cores revealed the yeast and human SG core proteome is enriched in translation initiation factors, RNA binding proteins, proteins with predicted prion-like domains (PrLDs), and proteins involved in neurodegenerative diseases (Jain et al., 2016). Intriguingly, SGs cores were also found to be composed of many proteins that were neither known to bind RNA nor predicted to contain PrLDs. Thus, this analysis identified many known classes of SG proteins, while providing evidence for additional SG components.

Little is known about the RNA composition of SGs. SGs are known to contain

non-translating mRNA from early reports showing that SGs contain PABP, stain positive for poly(A) mRNA, lack large ribosomal subunits, and are sensitive to drugs that alter the translation (Kedersha et al., 2000, 2002, 1999; McEwen et al., 2005). A few specific mRNAs have been shown to localize to SG (Kedersha and Anderson, 2002; Nelles et al., 2016; Stoecklin et al., 2004; Stöhr et al., 2006), although the full population of mRNAs in SG, and whether SG contain specific non-coding RNAs, or lincRNAs, has not been addressed.

Herein, we adapt our purification of SG cores to determine the transcriptome of both yeast and mammalian SG cores. We find that SGs are composed of over 80% mRNAs, although some lincRNAs and ncRNAs are enriched in SG. SGs contain mRNAs from essentially every expressed gene, with no single mRNA or ncRNA representing more than 1% of the SG RNA molecules. Partitioning of specific mRNAs into SGs is decreased by efficient translation properties and increased by longer coding and UTR regions. By quantifying the composition of the SG transcriptome, we discover only 10% of the bulk mRNA molecules accumulate in SGs, and only 185 genes have more than 50% of their mRNA molecules present in SGs. These results suggest SGs do not represent a specific biological program of mRNP assembly but instead form by condensation of untranslating mRNPs in proportion to their length and lack of association with ribosomes.

Mammalian SG cores are enriched for specific mRNAs

To determine the RNAs in mammalian SG cores, we purified and performed RNA-Seq in triplicate on SG cores isolated from U-2 OS cells after 60' of arsenite exposure **(Figure 3.S1A and Methods),** referred to as $SG_{coreRNA}$. For each sample, 5% of the lysate was extracted for RNA-Seq from the total RNA population, referred to as $Total_{RNA}$. Pairwise correlation coefficients indicate reliable transcriptomes between $SG_{coreRNA}$ library triplicates and $Total_{RNA}$ library triplicates **(Figure 3.S1B).**

Pairwise correlation analysis between $SG_{coreRNA}$ and $Total_{RNA}$ libraries demonstrate SG core transcriptomes are different from total RNA transcriptomes **($R < 0.04$, Figure 3.S1B).** Supporting this analysis, out of the 15689 transcripts identified (restricted to >1 FPKM in $Total_{RNA}$), 1841 transcripts are enriched in SG cores (>2-fold change, $p<0.01$) and 2539 transcripts are depleted from SG cores (<0.5-fold change, $p<0.01$) **(Figure 3.1A and Table 3.S1).** The remainder of the transcripts partitioned similarly between SG cores and the total RNA.

Validation of RNA-Seq by smFISH

To determine if the $SG_{coreRNA}$ transcriptome was accurate, we examined several mRNAs that were, based on RNA-Seq, preferentially enriched, neither enriched nor depleted, or preferentially depleted in SG cores by smFISH. We performed smFISH analysis on both stressed and unstressed U-2 OS cells for the following mRNAs:

AHNAK (fold change (FC) is calculated as $FPKM_{SGcore}/FPKM_{Total}$, FC = 6.92), DYNC1H1 (FC = 6.44), POLR2A (FC = 1.36), TFRC (FC = 1.10), PEG3 (FC=5.69), CDK6 (FC = 3.35), ZNF704 (FC = 4.09), and GAPDH (FC = 0.29) **(Figure 3.1B-D, 3.S1C).**

The smFISH analyses revealed that AHNAK, DYNC1H1, PEG3, ZNF704, and CDK6 mRNAs are more enriched in SGs compared to POLR2A, TFRC, and GAPDH mRNAs (**Figure 3.1B-D, S1C**), consistent with the RNA-Seq results. By quantifying the number of transcripts in SGs and cells detected by smFISH, we observed that on average, 80% of AHNAK, 74% of PEG3, 70% of CDK6, 64% of ZNF704 and 53% of DYNC1H1 mRNAs are in SGs while 17%, 23% and 4% of POLR2A, TFRC, and GAPDH mRNAs are in SGs respectively (**Figure 3.1E**). These proportions parrellel the enrichment of these mRNAs in SG cores by RNA-Seq **(Figure 3.1F)**. These numbers also reflect the SGcore FPKM reads. For example, DYNC1H1 has roughly twice as many FPKM reads as POLR2A (183.8 and 80, respectively) and roughly twice as many average number transcripts in SG in cells as measured by smFISH (22 and 10, respectively) (**Figure 3.1F**). Therefore, the smFISH experiments indicate our transcriptome analysis of SG cores is valid.

The preferential recruitment of specific RNAs to SGs was examined with other SG markers including PABP and Pumilio 2 (PUM2) (**Figure 3.S2A, 3.B**), indicating we are purifying RNAs enriched in bona fide SGs. Moreover, in ΔΔG3BP1/2 U-2 OS cells, which fail to form SGs under arsenite treatment (Kedersha et al., 2016), we no longer

observe the clustering of SG-enriched mRNAs by smFISH (**Figure 3.S2B**). These results argue we are identifying SG-enriched mRNAs.

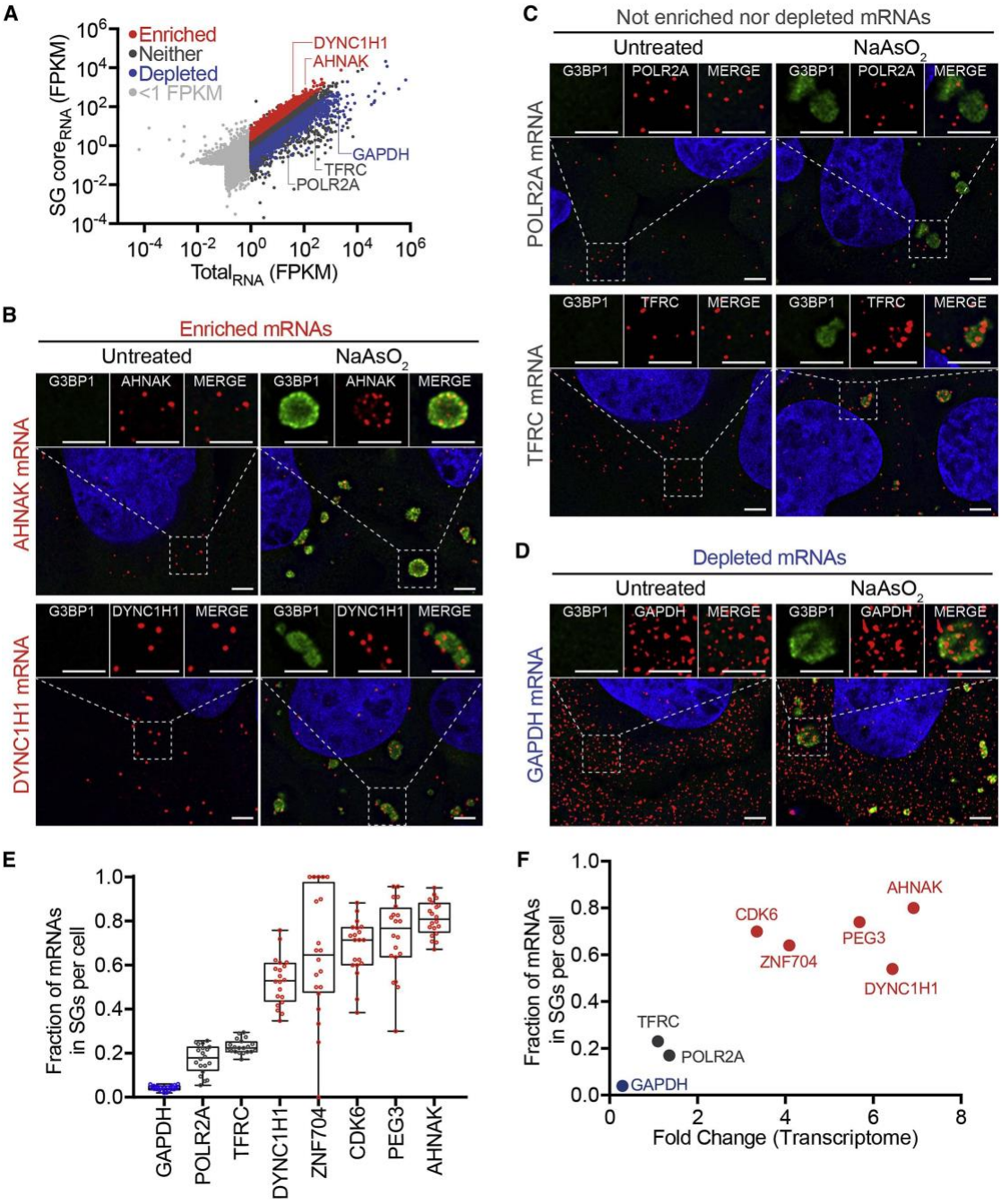


Figure 3.1: mRNAs differ in degree of recruitment to SGs

(A) Scatter plot depicting RNA abundance (FPKM) in SG purified RNA vs Total RNA. Red dots indicate RNAs that are significantly enriched (Fold change > 2 and p < .01) in SG purified RNA compared to Total RNA. Blue dots indicate RNAs that are significantly depleted (Fold change <.05 and p >.01) in SG purified RNA compared to $Total_{RNA}$. Dark gray dots indicate RNAs that are either not significantly enriched or fail to meet the fold change requirement. Light gray dots indicate mRNAs below <1FPKM. (B) smFISH validation of mRNAs enriched in SGs (AHNAK and DYNC1H1). (C) smFISH validation of RNAs that are neither enriched nor depleted from SGs (POLR2A and TFRC). (D) smFISH validation of a mRNA that is depleted from SGs (GAPDH). Scale bar = 2µm. (E) Quantification of the fraction of each indicated transcript in SGs per cell. 20 cells were counted. (F) Scatterplot depicting fraction of each indicated transcript in Sgs vs. FPKM ratios of $SG_{coreRNA}$-Seq: $Total_{RNA}$-Seq data (Fold-change). Twenty cells were counted for each experiment.

Quantitative analysis of the bulk transcriptome and SG transcriptome.

The combination of smFISH, which allowed us to count the number of specific mRNAs in the cell and within SGs, and RNA-Seq data for the total and SG transcriptome allowed us to standardize our RNA-Seq reads to numbers of mRNAs per cell. We observed a linear relationship between individual mRNAs identified per cell by smFISH and FPKM values for each of these mRNAs (**Figure 3.2A**). Linear regression analysis allowed us to estimate the relative number of molecules of every mRNA in the cell based on its FPKM value (**Table 3.S1**). We did not consider genes where the FPKM was less than 1, which allowed us to examine 11195 mRNAs (See STAR methods). From this analysis, we estimate there are an approximate 300,000 total mRNA molecules per cell, with mRNA abundance for different genes spanning from ~3500 to <1 mRNA per cell. These numbers are in the same range as other estimates for mammalian cells (Schwanhaussser et al., 2011)

One possibility is that we are purifying only a subpopulation of mRNAs from SGs by isolating G3BP1 cores and this is not an accurate representation of the complete RNA in SGs. To examine this, we looked at the relationship of the number of mRNA molecules in SGs for eight different mRNAs by smFISH and the FPKM values as determined by the SG core purification **(Figure 3.2B).** We excluded GAPDH mRNAs from this analysis because it is so abundant, we are unable to count all GAPDH mRNAs by smFISH reliably. We observed a linear relationship between the number of mRNAs within SGs identified by smFISH and the FPKM values from RNA-Seq analysis of SG

core-purified RNA with good correlation (**Figure 3.2B, R^2 = 0.759**). At a minimum, this correlation of RNA-Seq data with smFISH implies that the RNA composition of the shell is fairly similar to the core. More simply, we suggest that RNAs detected in the shell region by oligo(dT) FISH (Jain et al., 2016) are attached to, and co-purify with, cores once cells are lysed (see discussion).

With this analysis, we can provide an approximation of the percent enrichment of every mRNA in SGs **(Table 3.S1)**. It should be noted that these estimates are based on extrapolation from the transcripts we have examined by smFISH, and it remains possible that not all mRNAs will follow the same pattern. Nevertheless, we can estimate several features of the RNA content in SGs in 60' arsenite stress. We estimate that there are roughly 42,000 RNA molecules in SGs, of which ~78% are mRNAs (**Figure 3.2C and H**). Only .6% of ncRNAs localize to SGs, while roughly 10% of mRNA molecules localize to SGs **(Figure 3.2D, E)**. Consistent with this finding, we see a similar fraction of oligo(dT) staining (~12%) in SGs (**Figure 3.2F, G),** suggesting that our transcriptome analysis of SG cores captures most of the RNA content of SGs.

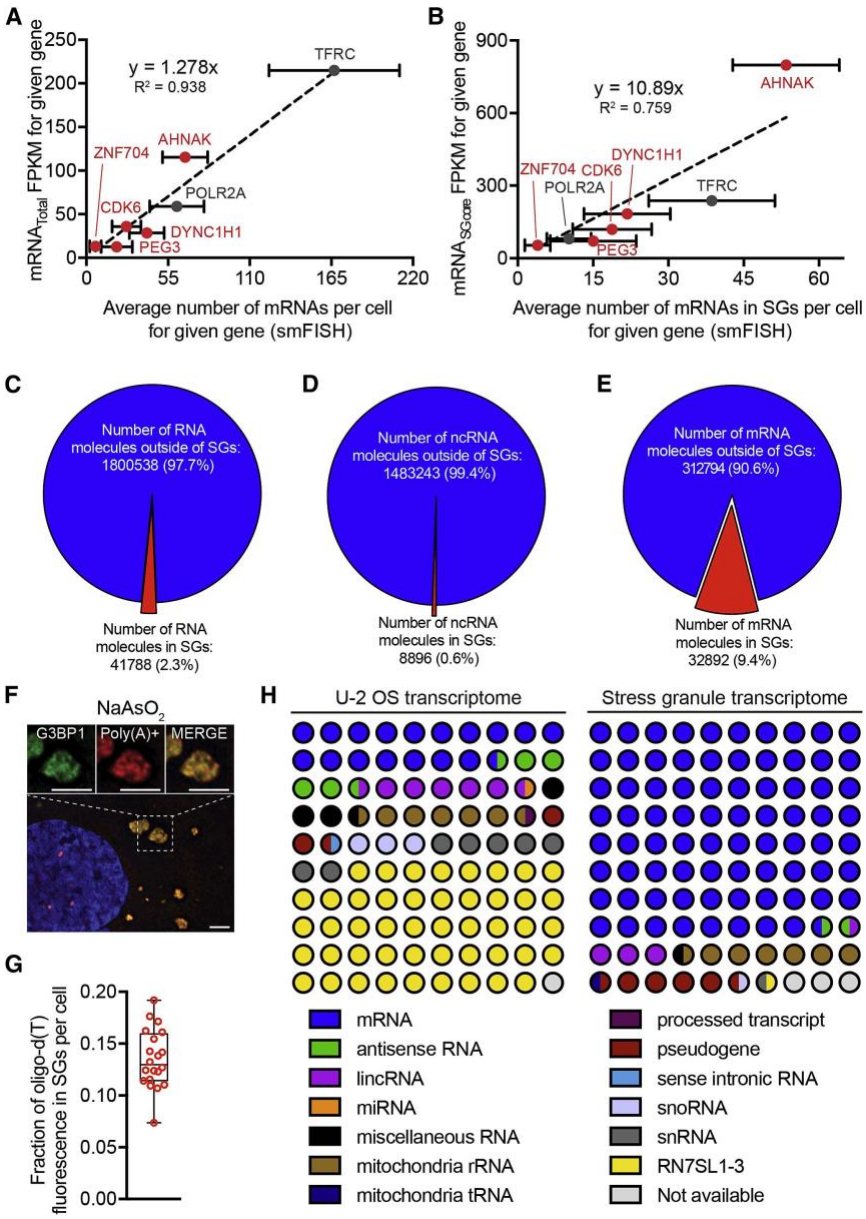

A
B
y = 1.278x
R² = 0.938
y = 10.89x
R² = 0.759
TFRC
AHNAK
ZNF704
POLR2A
CDK6
DYNC1H1
PEG3
mRNA_Total FPKM for given gene
mRNA_SGcore FPKM for given gene
Average number of mRNAs per cell for given gene (smFISH)
Average number of mRNAs in SGs per cell for given gene (smFISH)
C
D
E
Number of RNA molecules outside of SGs: 1800538 (97.7%)
Number of RNA molecules in SGs: 41788 (2.3%)
Number of ncRNA molecules outside of SGs: 1483243 (99.4%)
Number of ncRNA molecules in SGs: 8896 (0.6%)
Number of mRNA molecules outside of SGs: 312794 (90.6%)
Number of mRNA molecules in SGs: 32892 (9.4%)
F
NaAsO2
G3BP1
Poly(A)+
MERGE
G
Fraction of oligo-d(T) fluorescence in SGs per cell
H
U-2 OS transcriptome
Stress granule transcriptome
mRNA
antisense RNA
lincRNA
miRNA
miscellaneous RNA
mitochondria rRNA
mitochondria tRNA
processed transcript
pseudogene
sense intronic RNA
snoRNA
snRNA
RN7SL1-3
Not available

Figure 3.2: Quantitative analysis of SG composition

(A) Scatterplot depicting an average number of mRNAs per cell by smFISH vs. FPKM values from $Total_{RNA}$-Seq data. (B) Scatterplot depicting an average number of mRNAs in SGs per cell by smFISH vs. FPKM values from $SG_{coreRNA}$-Seq data. Error bars indicate 1 strandard deviation. (C) Pie chart depicting the fraction of total RNA inside and outside SGs. (D) Pie chart depicting the fraction of ncRNAs inside and outside SGs. (E) Pie chart depicting the fraction of mRNAs inside and outside SGs. (F) Oligo(dT) staining of sodium arsenite induced SGs. Scale bar = 2µm. (G) Quantitation of the fraction of oligo(dT) staining within SGs per cell. (H) RNA composition of the U2-OS transcriptome and the SGs transcriptome. RNA designations are from ensemble release 90.

mRNA partitioning to SGs is largely similar in diverse stresses.

Since the protein components of SGs can vary under different stresses (Kedersha et al., 1999, Stoecklin et al., 2004), we examined whether the RNA composition varies under different stresses. We performed smFISH for several mRNAs (AHNAK, DYNC1H1, POLR2A, TFRC, and GAPDH) and quantified their enrichment in SGs in arsenite, thapsigargin, heat shock, or sorbitol-induced SGs (Kedersha et al., 2016; Kim et al., 2007).

We observed that there are similarities and differences between different stresses **(Figure 3.3A-D)**. For example, in all stresses, the GAPDH mRNA is depleted from SGs, while the AHNAK and DYNCH1 mRNAs are enriched. In contrast, the TFRC is only enriched in SG during heat shock with an ~twofold increase in the fraction of mRNAs in SGs compared to other stresses, and the POLR2A mRNA is more highly enriched in SG during heat shock or sorbitol stress as compared to arsenite or thapsigargin stress. These results demonstrate that there are differences in the quantitative enrichment of specific mRNAs in SG dependent on the stress condition, which could be due to differences in the specificity of translation repression during different stresses.

We also examined the localization of AHNAK and POLR2A mRNAs in U-2 OS ΔΔG3BP1/2 cells under heat shock. U-2 OS ΔΔG3BP1/2 cells can not assemble SGs under arsenite and thapsigargin stresses but can assemble SGs under heat shock and sorbitol stresses **(Figure 3.S2B)** (Kedersha et al., 2016). We see enrichment of AHNAK

mRNAs in heat shock-induced SGs in U-2 OS ΔΔG3BP1/2 but not POLR2A mRNAs **(Figure 3.S2B).** This finding implies the RNA content of SGs is at least partially conserved in G3BP1/2 independent SGs.

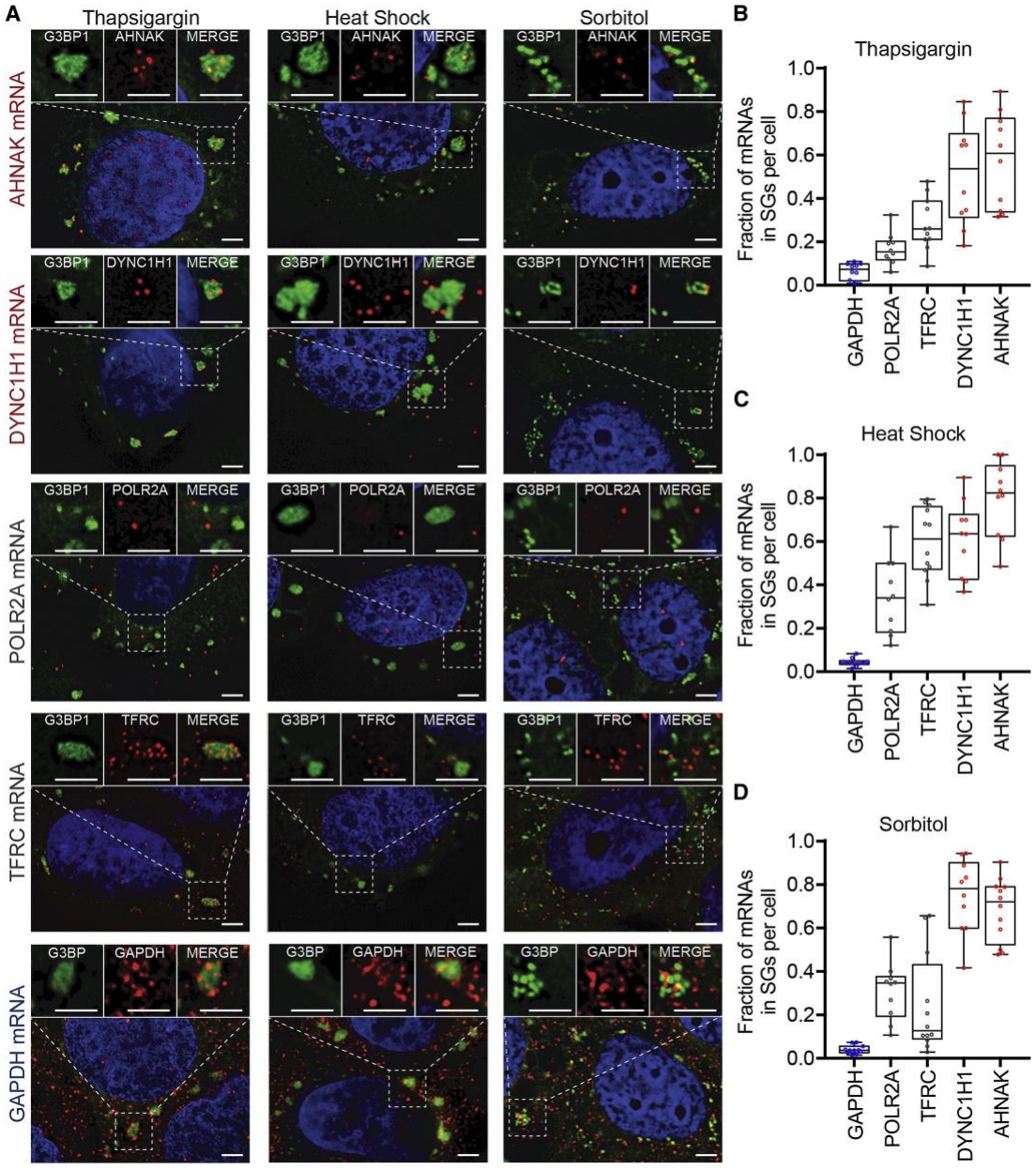


Figure 3.3: mRNA localization is conserved during multiple stresses

(A) smFISH images acquired for five transcripts (AHNAK, DYNC1H1, TFRC, POLR2A and GAPDH) during three different types of stresses (thapsigargin, heat shock, and sorbitol). (B) Quantitation of each transcript's enrichment in SGs. Ten or more cells were counted for each experiment.

mRNAs enriched in mammalian SGs have less ribosome density

To determine how mRNAs are partitioned into SGs, we examined the properites of SG enriched, depleted, and distributed RNAs. We reasoned that translation efficiency might be a major determinant for localization during stress since (1) decreased translation initiation induces SGs, and (2) trapping ribosomes on transcripts with cycloheximide is sufficient to repress SG formation (reviewed in Panas et al., 2016; Protter and Parker, 2016). This rationale predicts that efficiently translated transcripts should be depleted from SGs, while inefficiently translated transcripts preferentially accumulate in SGs.

We compared our sequencing results to a previous report that examined translation efficiency (Sidrauski et al., 2015). For statistical purposes, we restrict our analysis for mRNA with FPKM in total RNA greater than 1 (approximately one transcript per cell) **(Figure 3.2A).** Out of this restricted 11195 mRNAs list, 14.5% were enriched in SGs (> 2-fold change, $p < 0.01$) and 15.9% were depleted from SGs (<0.5 fold-change, $p < 0.01$) **(Figure 3.S3A, B).** We found that transcripts with higher translation efficiency tend to be depleted from SGs, while transcripts with lower translation efficiency tend to localize to SGs (**Figure 3.4A and Methods).** This finding demonstrates that one component that influences mRNA localization to SGs is the translation status of specific mRNAs. Since mRNAs that are poorly translated are generally less abundant and stable (Radhakrishnan and Green, 2016), we also observed mRNAs enriched in SGs have a shorter half life, are less abundant, and have less GC content **(Figure 3.S3C-E).**

Mammalian SG-enriched mRNAs are long

Examination of other metrics revealed SG-localized mRNAs were much longer (average length = 7.1 kb) than SG-depleted mRNAs (average length = 2.5 kb) and mRNAs that showed no preference for SGs (average length = 4.2 kb) (**Figure 3.4B**). Examination of the different contributions of the 5' UTR, ORF, or 3' UTR revealed that the length of the coding region showed the largest difference between enriched and depleted mRNAs in SGs, although the 3' UTR also had an effect (**Figure 3.4C-E**). We only observed slightly longer 5' UTRs on SG-enriched mRNAs (**Figure 3.4D**), which could be because longer 5' UTRs can decrease translation initiation due to RNA structure.

Given the overall dependence on mRNA length for accumulation in SGs, we also examined how the contributions of the 5' and 3' UTRs varied with the length of the coding region. The logic for this analysis was that for mRNAs with short coding regions, the 5' UTR and 3' UTR length might compensate for a shorter coding region. Therefore the 5' and 3' UTR effects might be more obvious at shorter ORF lengths. Thus, we examined the contributions of the 5' and 3' UTR by first binning mRNAs by their ORF length and then determining if there was a difference in the 5' or 3' UTR length for mRNAs accumulating in SGs compared to those depleted from SGs. We observed that 3' length made a significant difference for all coding regions size bins, but with a small effect once the mRNA coding region was over 3000 bases (**Figure 3.S3F**). A similar analysis for 5' UTR length revealed the impact of longer 5' UTR lengths was

insignificant (**Figure 3.S3G)**. Taken together, these results indicate the length of the coding region and 3' UTRs are important metrics for determining SG accumulation.

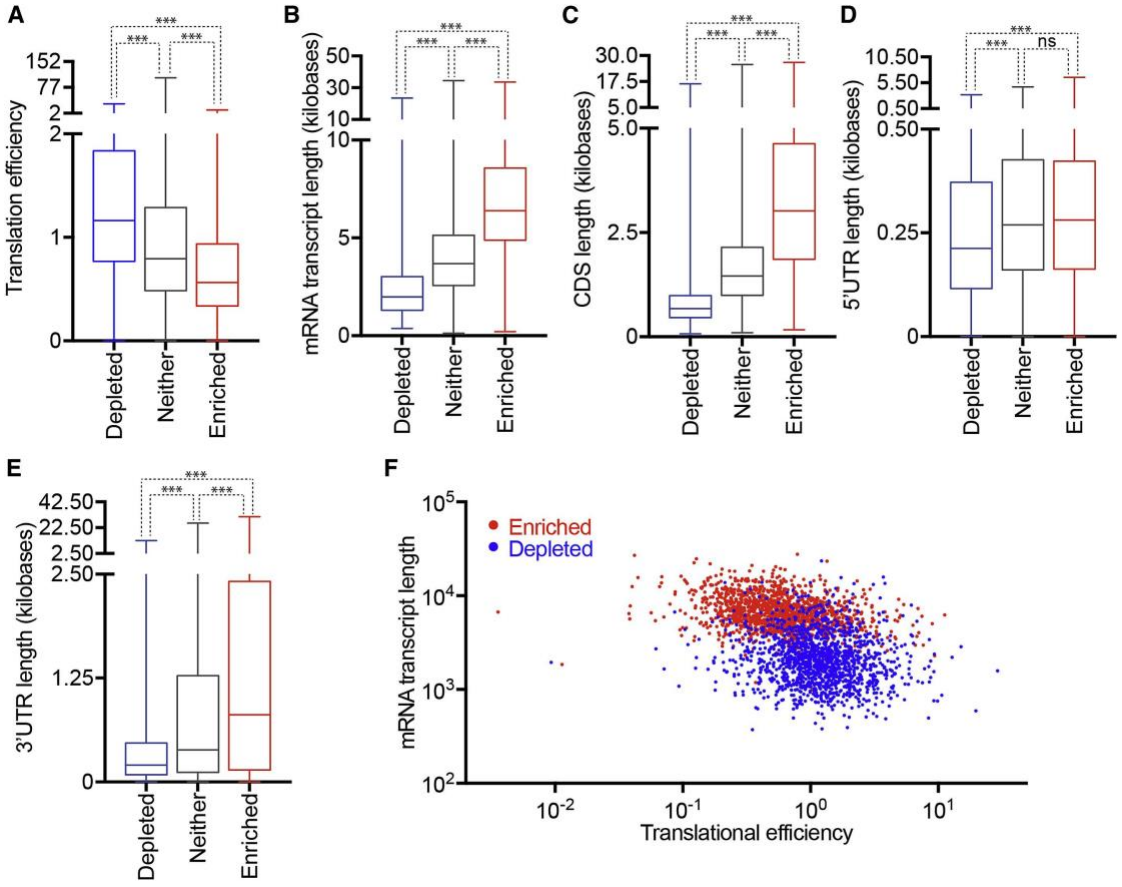


Figure 3.4: Physical basis of differential mRNA recruitment to mammalian SGs

(A-E) Boxplots depicting (A) translational efficiency, (B) total transcript length, (C) CDS length, (D) 5' UTR length, and (E) 3'UTR length for each of the three classes of mRNA localization during stress: SG enriched mRNAs, SG depleted mRNAs, or neither. (F) Scatter plot depicting the mRNA transcript length vs. translational efficiency for SG enriched (red) and SG depleted (blue) mRNAs. n.s. (not significant) = $p \geq .05$ and *** = p value ≤0.001 (Student's t-test).

Mammalian SG-enriched ncRNAs are long

We also examined the ncRNAs that are preferentially recruited to SGs. We restricted our analysis for ncRNA with FPKM in total RNA greater than 1. Out of these restricted 4494 ncRNAs, 4.8% were enriched in SGs (> 2-fold change, $p < 0.01$) and 16.9% were depleted from SGs (<0.5 fold-change, $p < 0.01$) **(Figure 3.5A-B).** Enriched ncRNAs (average length = 1.9 kb) tend to be longer than depleted ncRNA (average length = 0.9 kb) and ncRNA that showed no preference for SGs (average length = 0.9 kb) **(Figure 3.5C).** GC content and abundance of ncRNA are not significant predictors of accumulation in SGs **(Figure 3.5D-E).**

Specific examination of lincRNAs revealed SG-enriched lincRNAs also tend to be longer **(Figure 3.5F).** We validated a lincRNA, NORAD, is enriched in SGs by smFISH **(Figure 3.5G-H).** NORAD is one of the top 10 most enriched lincRNA in SGs **(Table 3.S3A)** and is thought to function as a Pumulio sponge (Lee et al., 2016; Tichon et al., 2016).

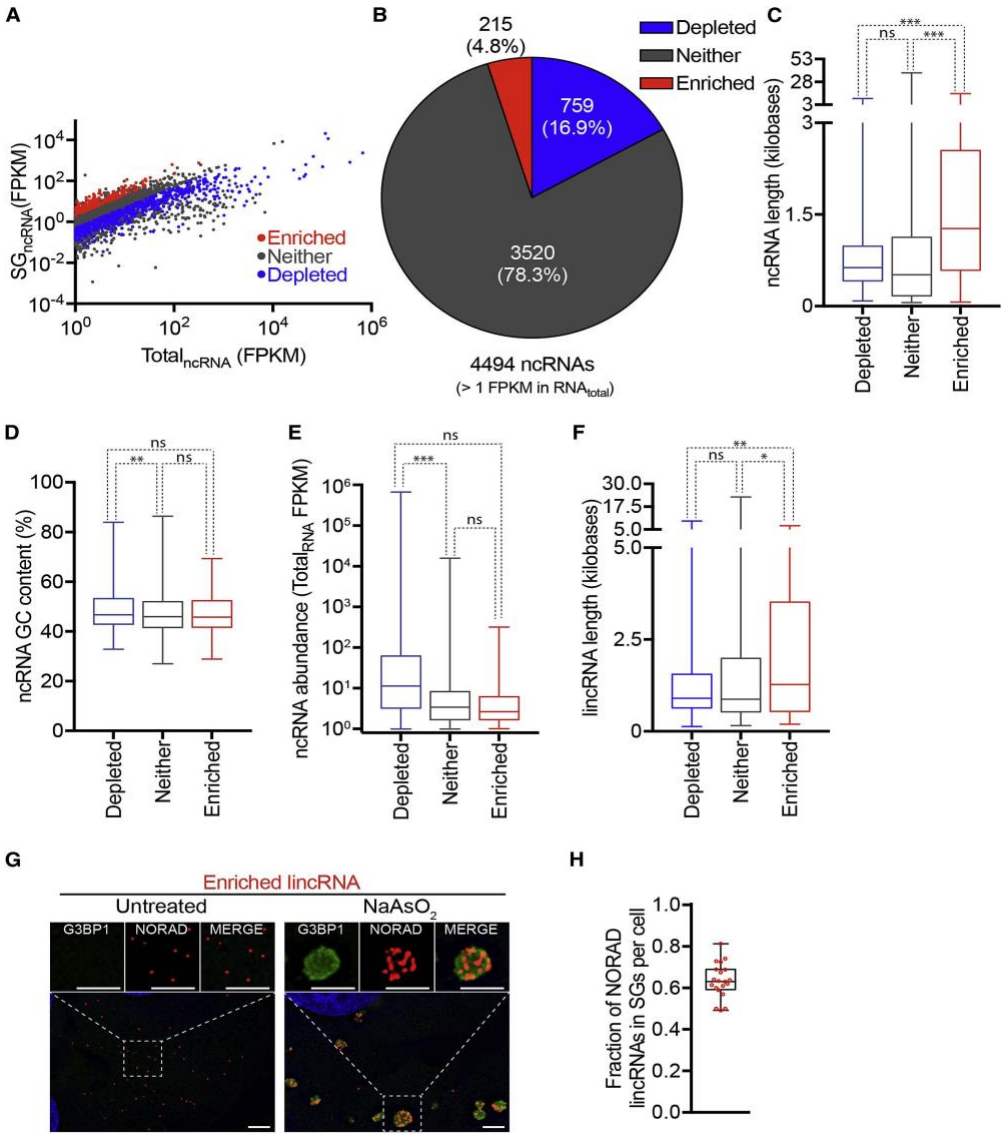

A
B
C
D
E
F
G
H
Enriched
Neither
Depleted
215
(4.8%)
759
(16.9%)
3520
(78.3%)
4494 ncRNAs
(> 1 FPKM in RNA$_{total}$)
SG$_{ncRNA}$ (FPKM)
Total$_{ncRNA}$ (FPKM)
ncRNA length (kilobases)
ncRNA GC content (%)
ncRNA abundance (Total$_{RNA}$ FPKM)
lincRNA length (kilobases)
ns
**
***
*
Enriched lincRNA
Untreated
NaAsO$_2$
G3BP1
NORAD
MERGE
Fraction of NORAD lincRNAs in SGs per cell

Figure 3.5: Physical basis of differential ncRNA recruitment to mammalian SGs

(A) Scatter plot depicting ncRNA abundance (FPKM) in SG purified RNA vs. Total RNA. Red dots indicate ncRNAs that are significantly enriched (Fold change > 2 and p < .01) in SG purified RNA compared to Total RNA. Blue dots indicate ncRNAs that are significantly depleted (Fold change <.05 and p >.01) in SG purified RNA compared to Total RNA. Gray dots indicate ncRNAs that are either not significantly enriched or fail to meet the fold change requirement. (B) Pie chart depicting the relative contribution of each class of ncRNA (SG-enriched, SG-depleted, or neither) to the total transcriptome SG transcriptome (right). (C-F) Boxplot depicting the (C) transcript length, (D) GC content, (E) abundance of ncRNA, and (F) length of lincRNAs for the three classes of localization: SG enriched mRNAs, SG depleted mRNAs, or neither. n.s., *, **, and *** = p value (Student's t-test) > 0.05, ≤0.05, ≤0.01, and ≤0.001 respectively. (G) smFISH images of an SG enriched lincRNA, NORAD, during no stress and sodium arsenite induced stress. Scale bar = 2μm. (H) Quantitation of the fraction of NORAD RNAs in SGs. Twenty cells were counted.

Length and Translation efficiency influence SG targeting of RNAs in yeast

To determine if the principles of mRNA targeting to SGs observed in mammalian cells were conserved, we purified and sequenced the mRNAs within SG cores from *Saccharomyces cerevisiae* after induction of SGs by sodium azide for 30' (Buchan et al., 2011). SG_{RNA} and $Total_{RNA}$ were isolated from three independent biological replicates and sequenced. Reads from biological replicates of SG_{RNA} as well as $Total_{RNA}$ were highly reproducible **(Figure 3.S4A)**.

Yeast RNA-Seq reads from SG_{RNA} and $Total_{RNA}$ showed little similarity to each other indicating a subset of cellular RNA is significantly enriched in SGs **(Figure 3.6A, Table 3.S2, Figure 3.S4A)**. By comparing reads across RNAs in SG_{RNA} and $Total_{RNA}$, we found that 916 mRNAs were significantly (p-value <0.01) enriched in SG_{RNA} as compared to $Total_{RNA}$, while 1107 mRNAs were significantly enriched in $Total_{RNA}$ as compared to SG_{RNA} (**Figure 3.6A**). Thus, during stress in yeast, we observe the same three classes of mRNAs as in mammalian cells: one subset preferentially localized to SGs, another subset preferentially depleted from SGs, and a third subset partitioning similarly between SGs and the cytosol. smFISH for a SG-enriched mRNA, MDM1, validated that it accumulated in SGs (**Figure 3.S4B**).

Multiple lines of evidence suggest that poorly translated mRNAs localize to yeast SGs, while efficiently translated mRNAs are depleted. First, we observed that predicted ribosome density and translation initiation rates for SG-depleted mRNAs are

significantly higher than the predicted rates for SG-enriched mRNAs **(p-value <0.0001, Figure 3.6B and 3.S4C**) (Siwiak and Zielenkiewicz, 2010). Second, we observe codon optimality, which is a metric for translation efficiency (Presnyak et al., 2016), strongly correlates with the distribution of mRNAs. Specifically, SG-enriched mRNAs have an average of 43.9% optimal codons, SG-depleted mRNAs have an average of 59.6% optimal codons **(p-value <0.0001, Figure 3.6C**). Consistent with the observation that SG-enriched mRNAs are composed of non-optimal codons, and the observation that non-optimal codons decrease a transcript's stability (Presnyak et al., 2016), we also found that SG-enriched mRNAs are less stable than SG-depleted mRNAs **(Figure 3.S5D)**.

Similar to mammalian cells, we found that the average SG-enriched mRNA length in yeast (2.7 kb) is significantly longer than SG-depleted mRNAs (1.1kb) (**Figure 3.6D**) (Nagalakshmi et al., 2008). This length dependence in yeast is driven primarily by ORF length, and not by 5' UTR or 3' UTR length, which is consistent with the overall shorter 3' UTR length in yeast (**Figures 3.6F-H**).

Thus, similar to mammalian cells, the major metrics for mRNA accumulation in yeast SGs are length and poor translation efficiency **(Figure 3.4F, 3.6E)**. Indeed, mathematical models made by machine learning and based solely on length and codon optimality (or ribosome density for mammalian mRNAs) are sufficient to accurately predict whether a mRNA will be enriched or depleted from SGs for 75.9% of the yeast mRNAs and 74.1% of mammalian mRNAs (**Figure 3.S5, see Methods).**

Yeast SGs contain non-coding RNAs

Since we observed that mammalian SGs contain ncRNAs, we wanted to examine if yeast SGs contain ncRNA as well. In yeast, we identified some stable untranslated transcripts (SUTs) (Xu et al., 2009) that accumulate in SGs (**Figure 3.S4E**). The SUTs enriched in SGs were significantly longer than SG-depleted SUTs **(Figure 3.S4G)**. In addition, cryptic unstable transcripts (CUTs), which are a class of non-coding RNAs that are related to SUTs (Xu et al., 2009), but are much shorter (200-800nts), are rarely enriched in SGs (**Figure 3.S4F**). Again, SG-enriched CUTs were longer than SG-depleted CUTs (**Figure 3.S4H**). The inclusion of ncRNAs in SGs demonstrates that prior translation *per se* is not a requisite for RNA accumulation in SGs and that length plays a role in targeting both coding and noncoding transcripts to SGs.

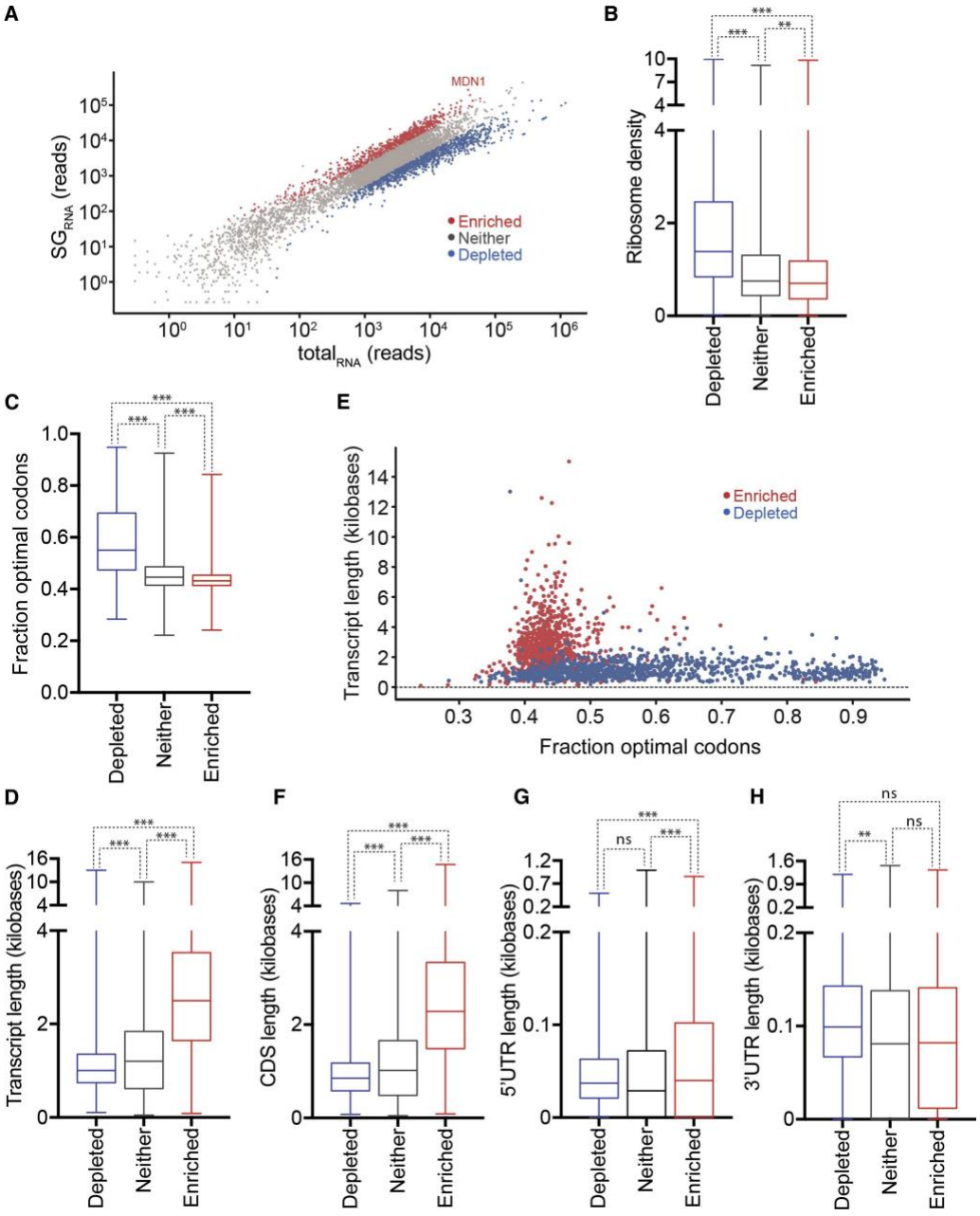

A
MDN1
SG_RNA (reads)
total_RNA (reads)
Enriched
Neither
Depleted
B
Ribosome density
Depleted
Neither
Enriched
C
Fraction optimal codons
D
Transcript length (kilobases)
E
Transcript length (kilobases)
Fraction optimal codons
F
CDS length (kilobases)
G
5'UTR length (kilobases)
H
3'UTR length (kilobases)
ns

Figure 3.6: The physical basis of mRNA localization to SGs is conserved from mammals to yeast

(A) Scatterplot depicting normalized RNA-Seq reads from libraries made from yeast SG purified RNA vs. total RNA. (B and C) Boxplots depicting two metrics of translatability, (B) ribosome density and (C) codon optimality, are correlated with altered localization during stress. (D) Boxplot depicting how transcript length correlates with altered localization. (E) Scatterplot of transcript length vs. fraction optimal codons for SG enriched (red) and SG depleted mRNAs (blue). (F-H) Boxplots depicting how length correlates with localization for (F) the CDS, (G) the 5'UTR, and (H) the 3' UTR. n.s., **, and *** = p value (Student's t-test) > 0.05, ≤0.01, and ≤0.001 respectively.

Membrane association limits SG partitioning of mRNA

Analysis of the SG transcriptome reveals that mRNAs associated with membranes are depleted from SGs, which was suggested by early observations that mRNAs localizing to the ER do not localize to SGs (Unsworth et al., 2010). Specifically, by comparison of the yeast SG-enriched mRNAs to mRNAs associated with the ER through SRP interactions (Chartron et al., 2016), we found that SG-enriched mRNAs are significantly distinct from the subset of mRNAs encoding proteins that localized via SRP **(Figure 3.S6A)**, while SG-depleted mRNAs showed significant overlap with ER-localized mRNAs. Consistent with this observation, we find that human SG-enriched mRNAs are significantly distinct from the subset of mRNAs encoding proteins that are secreted **(Figure 3.S6C).** We also found that SG-enriched mRNAs in both yeast and mammalian cells were significantly distinct from mRNAs that encode proteins that localize to the mitochondria (**Figure 3.S6B, D**) (Chartron et al., 2016), while SG-depleted mRNAs showed significant overlap with mRNAs encoding mitochondria-localized proteins (**Figure S6B, D**).

Taken together, these findings suggest that mRNA localization to other organelles may preclude mRNA from efficiently partitioning into SGs, which could be explained by restricted diffusion when mRNAs are bound to membrane surfaces. It should be noted that we cannot rule out the formal possibility that a subset of SGs form on the ER or mitochondrial surface and such membrane-associated SGs do not purify efficiently in our biochemical preparations.

SG-enriched mRNAs are modestly enriched for SG-resident proteins

One model for mRNA recruitment to SGs is that SG-enriched mRNAs would bind more SG-targeted RNA binding proteins, and the accumulation of those proteins in SG would then partition specific mRNAs into SGs. A prediction of this model is that SG-enriched RNAs would be enriched for RNA-binding sites of known SG-localized RNPs. To test this prediction, we mined previously published eCLIP datasets on both SG resident (n = 36) and non-SG resident (n =51) RNPs by scoring eCLIP binding sites in SG-enriched RNAs **(Figure 3.7A)** (Nostrand et al., (2016)). We observed that RNA binding proteins on average had approximately 22% of their eCLIP sites in mRNAs enriched in SG.

Surprisingly, we observed SG-localized RNA binding proteins were only modestly enriched on SG-enriched mRNAs as compared to non-SG resident RNA binding proteins **(Figure 3.7B, p-value = 0.0484).** Similarly, SG-localized RNA binding proteins were slightly underrepresented in SG-depleted mRNAs as compared to non-SG resident RNA binding proteins **(Figure 3.7C, p-value = 0.0486).** Whether this modest enrichment is because these SG-resident RNA binding proteins play a role in targeting mRNAs into SGs, or are identified as SG components because they show a slight preference for mRNAs that accumulate in SGs will require future experiments.

We also observed that RNA binding proteins found in SGs with prion-like domains (PrLD), or with an association with degenerative diseases such as ALS or FTD, showed an even greater bias to bind to mRNAs enriched in SGs **(Figure 3.7A, D-E)**. The enrichment of PrLD containing proteins involved in disease with SG mRNAs is consistent with a model whereby the prevalent recruitment of these proteins into SGs

may both enhance the formation of aberrant pathological aggregates, and contribute to disease progression (Li et al., 2013; Ramaswami et al., 2013).

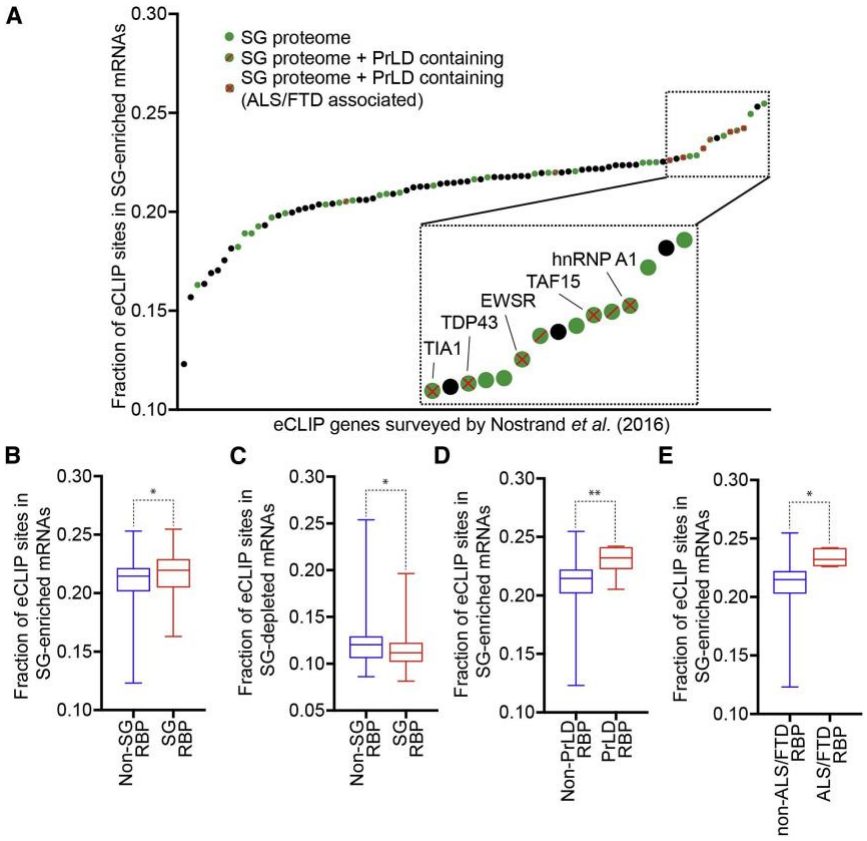
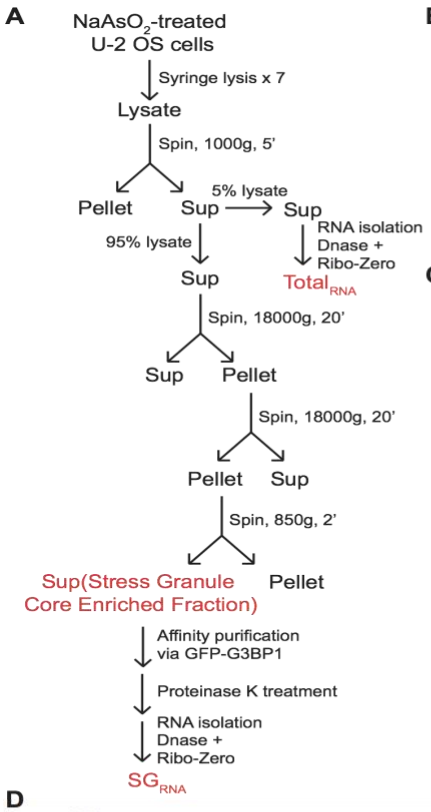


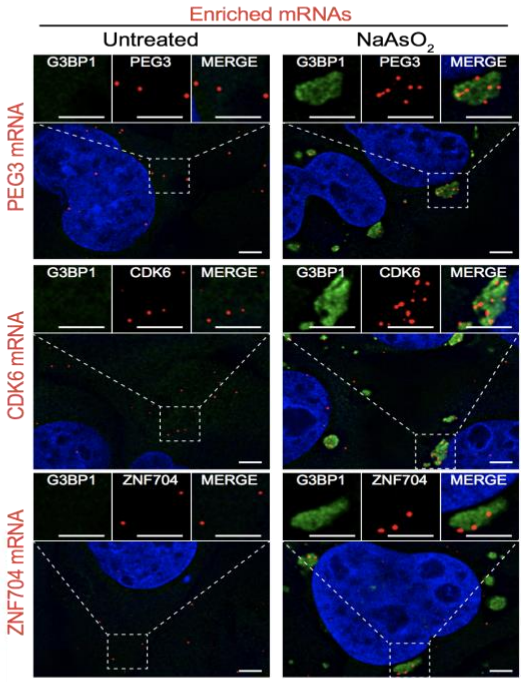
Figure 3.7: SG-enriched mRNAs are enriched for SG-resident proteins

(A) Survey of RNA binding proteins from Nostrand et al. (2016)) plotted against the fraction of eCLIP sites found in SG-enriched mRNAs. Green dots are SG-resident RBPs. Black dots are non-SG resident RBPs. Green dots with a red diagonal line are SG-resident RBPs with PrLD domain. Green dots with a red X are SG-resident RBPs with PrLD domain and are associated with ALS/FTD. Box plot illustrating the fraction of eCLIP sites found in (B) SG-enriched mRNAs or (C) SG-depleted mRNAs from SG-resident RBPs or non-SG RBPs. Box plot depicting the fraction of eCLIP sites found in SG-depleted mRNAs from SG-RBPs or non-SG RBPs. Box plot illustrating the fraction of eCLIP sites found in SG-enriched RNAs from (D) SG-RBPs containing PrLD and (E) RBPs associated with ALS/FTD. * and ** = p value (Student's t-test) ≤0.05, and ≤0.01 respectively.



B

	SG_{RNA1}	SG_{RNA2}	SG_{RNA3}	$Total_{RNA1}$	$Total_{RNA2}$	$Total_{RNA3}$
SG_{RNA1}	1.000	0.947	0.967	0.092	0.092	0.102
SG_{RNA2}	0.947	1.000	0.880	0.065	0.066	0.073
SG_{RNA3}	0.967	0.880	1.000	0.061	0.061	0.071
$Total_{RNA1}$	0.092	0.065	0.061	1.000	0.995	0.995
$Total_{RNA2}$	0.092	0.066	0.061	0.995	1.000	0.998
$Total_{RNA3}$	0.102	0.073	0.071	0.995	0.998	1.000



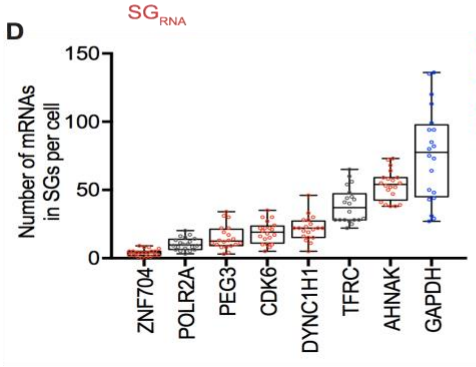


Figure 3.S1 (related to Figure 1): Purification and analysis of the RNA content in mammalian SG cores

(A) Schematic depicting how SG cores were purified. In brief, SG formation was induced by sodium arsenite stress (.5mM, 1 hour) in a G3BP1-GFP tagged U-2 OS cell line. Cells were then subjected to syringe lysis, followed by a series of centrifugation steps. This stress granule core enriched fraction was then affinity purified via GFP-G3BP1 and proteinase K treated. RNA was isolated via trizol extraction and ribodepleted, yielding purified SG core RNA. (B) Pairwise R^2 correlations between individual replicates of SG purified RNA and total RNA. (C) smFISH validation of mRNAs enriched in SGs (PEG3, CDK6, and ZNF704). Scale bar = 2μm. (D) Quantification of the average number of each indicated transcripts that occur in SGs per cell. Twenty cells were counted for each experiment.

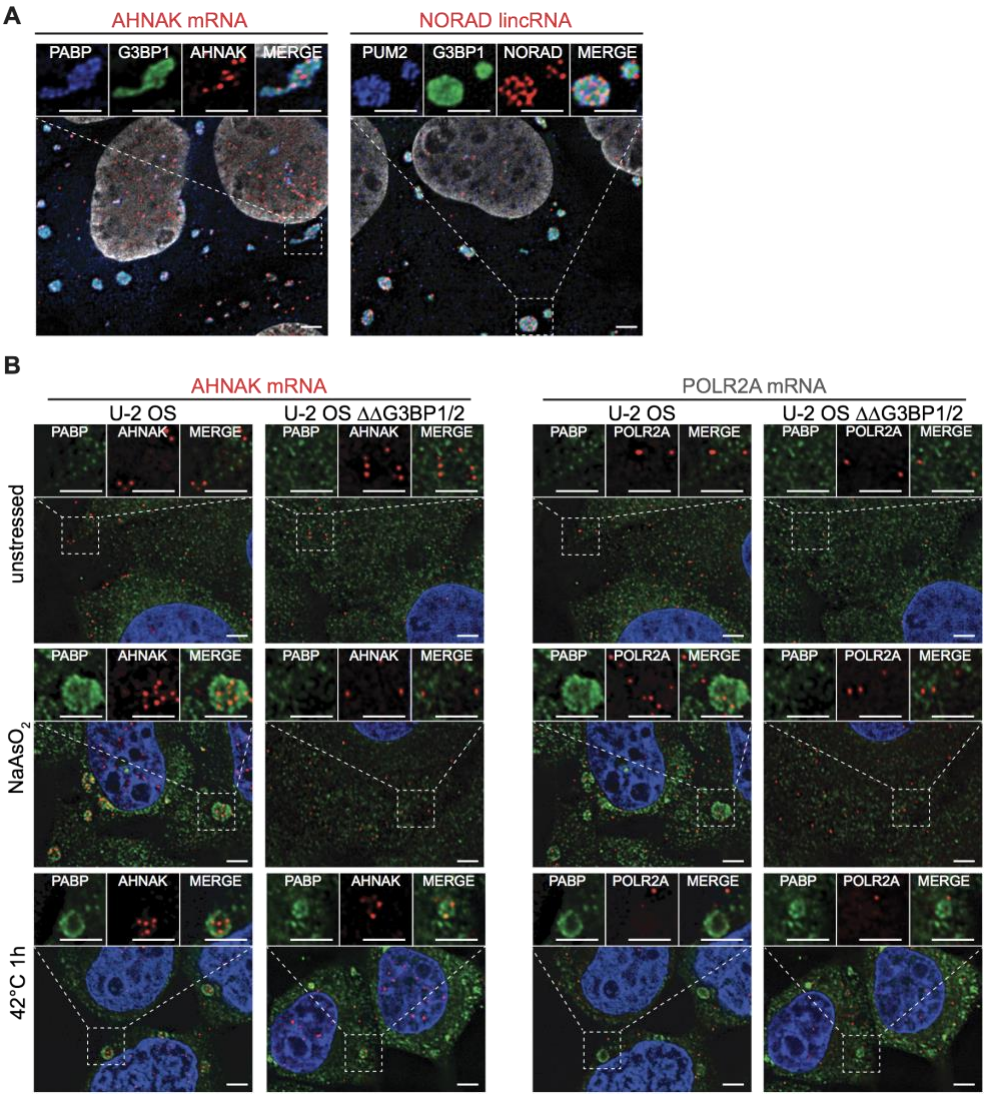


Figure 3.S2 (related to Figure 1): Validation of the RNA-content in bona fide SGs

(A) Arsenite-treated U-2 OS cells co-stained for (Left) PABP protein, G3BP1 protein, and AHNAK mRNA or (Right) PUM2 protein, G3BP1 protein, and NORAD lincRNA. (B) Arsenite-treated wildtype U-2 OS cells and ΔΔG3BP1/2 U-2 OS cells were co-stained for PABP protein, (left) AHNAK mRNA, and (right) POLR2A mRNA. Scale bar = 2μm.

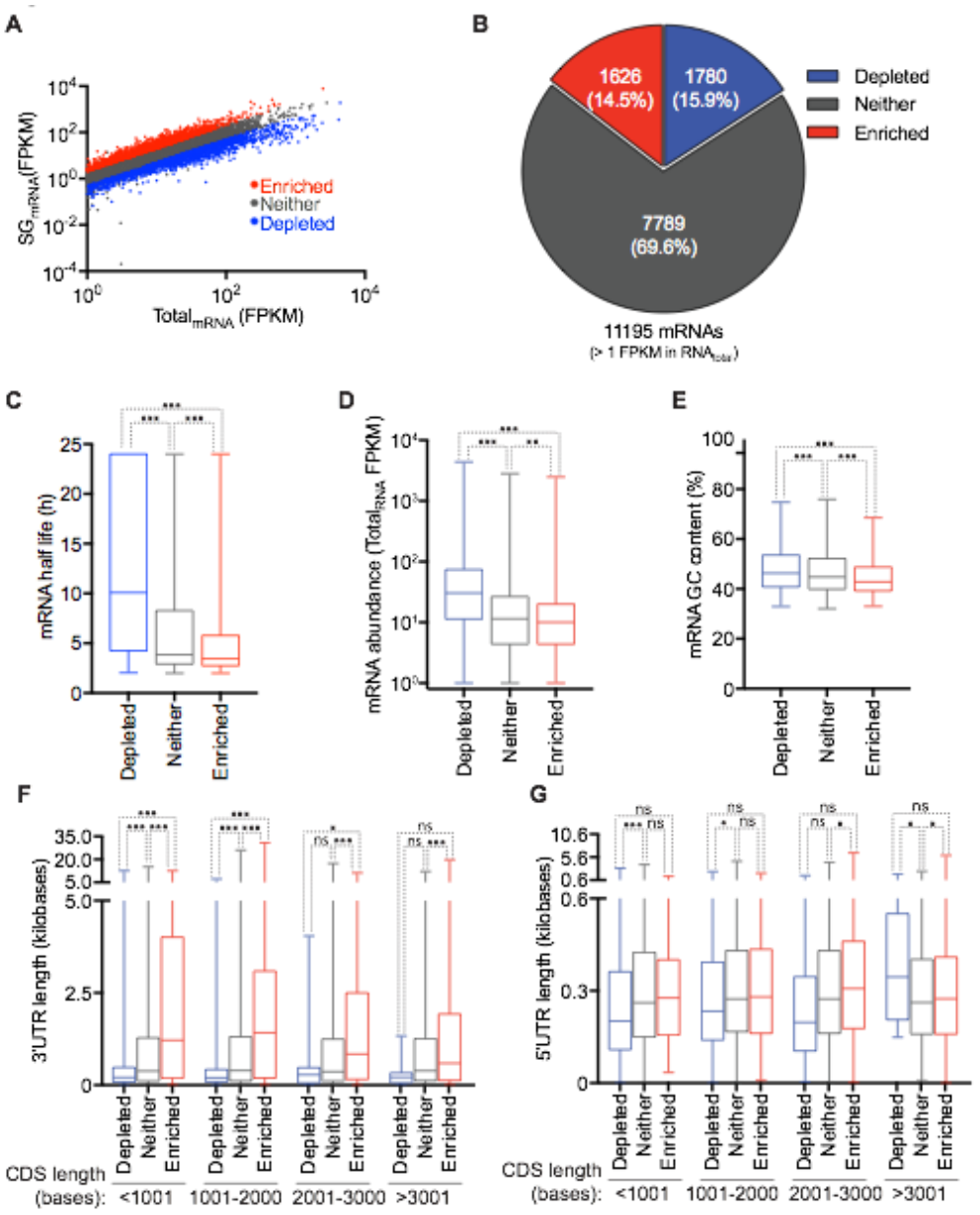

A
B
C
D
E
F
G
Enriched
Neither
Depleted
1626 (14.5%)
1780 (15.9%)
7789 (69.6%)
11195 mRNAs
mRNA half life (h)
mRNA abundance
mRNA GC content (%)
3'UTR length (kilobases)
5'UTR length (kilobases)
CDS length (bases):
<1001
1001-2000
2001-3000
>3001

Figure 3.S3 (related to Figure 4): Properties of differential mRNA recruitment to mammalian SGs

(A) Scatter plot depicting mRNA abundance (FPKM) in SG purified RNA vs. Total RNA. Red dots indicate mRNAs that are significantly enriched (Fold change > 2 and p < .01) in SG purified RNA compared to Total RNA. Blue dots indicate mRNAs that are significantly depleted (Fold change <.5 and p >.01) in SG purified RNA compared to Total RNA. Gray dots indicate mRNAs that are either not significantly enriched or fail to meet the fold change requirement. (B) Pie chart depicting the relative contribution of each class of ncRNA (SG-enriched, SG-depleted, or neither) to the SG transcriptome. (C-E) Boxplots depicting (C) mRNA half-life, (D) mRNA abundance, and (E) GC content for the three classes of localization: SG enriched mRNAs, SG depleted mRNAs, or neither. (F) Boxplots illustrating 3' UTR length at various ORF lengths for each class of mRNA localization. mRNA ORF lengths were binned every kilobase. (G) Same as in F but for 5' UTR lengths. n.s., *, **, and *** = p value (Student's t-test) > 0.05, ≤0.05, ≤0.01, and ≤0.001 respectively.

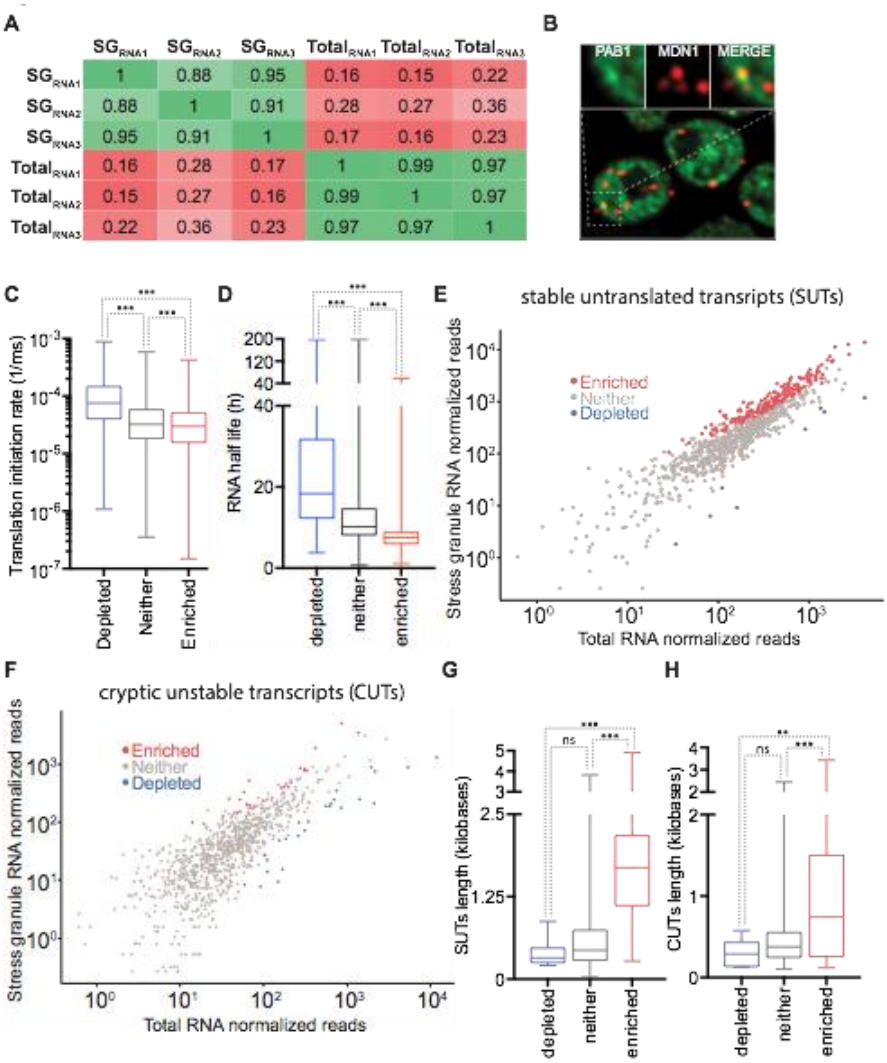

A
SG RNA1
SG RNA2
SG RNA3
Total RNA1
Total RNA2
Total RNA3
1 0.88 0.95 0.16 0.15 0.22
0.88 1 0.91 0.28 0.27 0.36
0.95 0.91 1 0.17 0.16 0.23
0.16 0.28 0.17 1 0.99 0.97
0.15 0.27 0.16 0.99 1 0.97
0.22 0.36 0.23 0.97 0.97 1
B
PABT
MDN1
MERGE
C
Translation initiation rate (1/ms)
Depleted
Neither
Enriched
D
RNA half life (h)
depleted
neither
enriched
E
stable untranslated transripts (SUTs)
Enriched
Neither
Depleted
Stress granule RNA normalized reads
Total RNA normalized reads
F
cryptic unstable transcripts (CUTs)
Enriched
Neither
Depleted
Stress granule RNA normalized reads
Total RNA normalized reads
G
SUTs length (kilobases)
depleted
neither
enriched
H
CUTs length (kilobases)
depleted
neither
enriched

Figure 3.S4 (related to Figure 6): Purification and analysis of the RNA content in yeast SG cores

(A) Pairwise R^2 correlations between individual replicates of SG purified RNA from yeast and total yeast RNA. (B) RNA smFISH of MDN1 after 30 minutes of sodium azide treatment in Pab-GFP expressing yeast. (C) Boxplot showing predicted translation initiation rates for SG enriched RNA, SG depleted RNA, and RNA that shows no preference. (D) Boxplot depicting RNA half-lives for SG enriched RNA, SG depleted RNA, and RNA that shows no preference. (E) Scatterplot illustrating SUTs that are enriched in SGs (red), depleted from SGs (blue), or neither (gray). (F) Same as E but for CUTs. (G) Boxplot depicting transcript length of SUTs for each class of mRNA localization. (H) Same as G but for CUTs. n.s., **, and *** = p value (Student's t-test) > 0.05, ≤0.01, and ≤0.001 respectively.

Figure S5

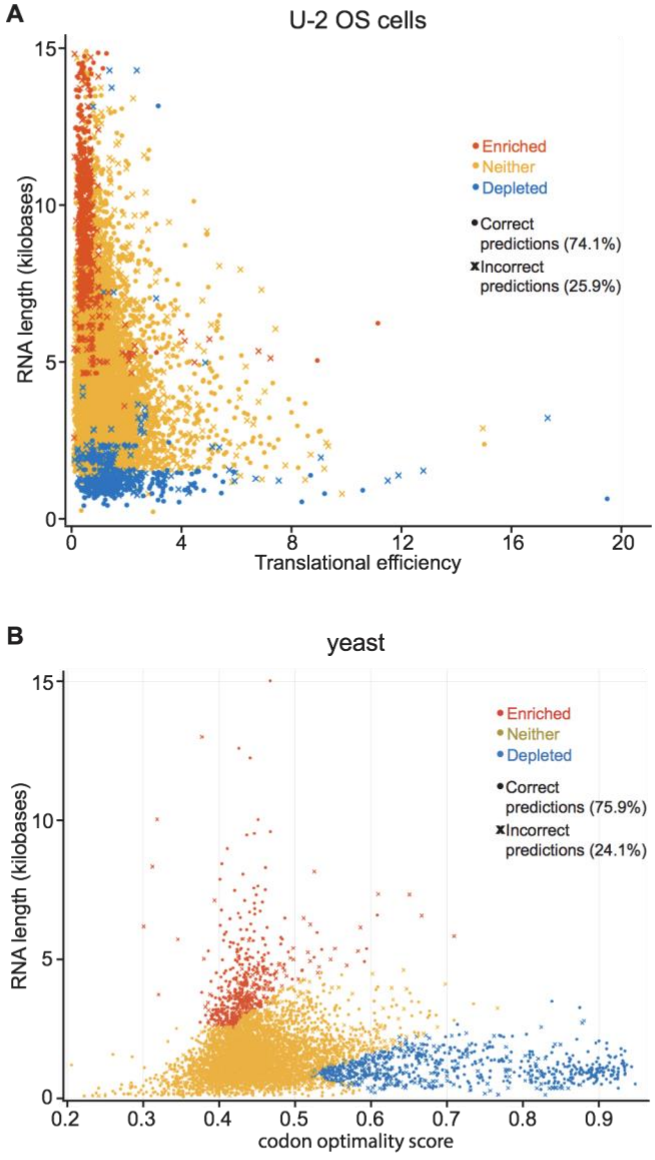

A
U-2 OS cells
RNA length (kilobases)
Translational efficiency
Enriched
Neither
Depleted
Correct predictions (74.1%)
Incorrect predictions (25.9%)
B
yeast
RNA length (kilobases)
codon optimality score
Enriched
Neither
Depleted
Correct predictions (75.9%)
Incorrect predictions (24.1%)

Figure 3.S5 (related to Figures 4 and 6): Machine learning model accurately predicts mRNA localization in SGs

(A) A machine learning model accurately predicts 74.1% of mammalian mRNA localization during stress based on ribosome density and transcript length alone. (B) A machine learning model accurately predicts 75.9% of yeast mRNA localization during stress based on fraction optimal codons and transcript length alone.

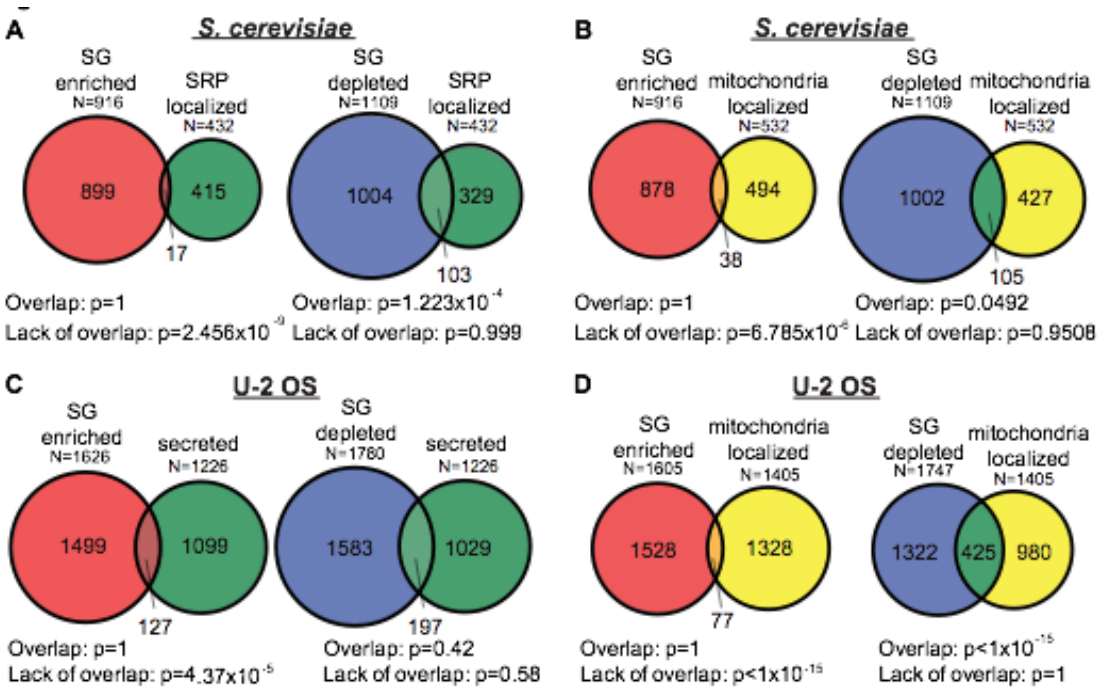

A
S. cerevisiae
SG enriched N=916
SRP localized N=432
899
415
17
Overlap: p=1
Lack of overlap: p=2.456x10-9
SG depleted N=1109
SRP localized N=432
1004
329
103
Overlap: p=1.223x10-4
Lack of overlap: p=0.999
B
S. cerevisiae
SG enriched N=916
mitochondria localized N=532
878
494
38
Overlap: p=1
Lack of overlap: p=6.785x10-6
SG depleted N=1109
mitochondria localized N=532
1002
427
105
Overlap: p=0.0492
Lack of overlap: p=0.9508
C
U-2 OS
SG enriched N=1626
secreted N=1226
1499
1099
127
Overlap: p=1
Lack of overlap: p=4.37x10-5
SG depleted N=1760
secreted N=1226
1583
1029
197
Overlap: p=0.42
Lack of overlap: p=0.58
D
U-2 OS
SG enriched N=1605
mitochondria localized N=1405
1528
1328
77
Overlap: p=1
Lack of overlap: p<1x10-15
SG depleted N=1747
mitochondria localized N=1405
1322
425
980
Overlap: p<1x10-15
Lack of overlap: p=1

Figure 3.S6 (related to Figures 4 and 6): Membrane targeting affects SG localization

(A) Venn diagram depicting the overlap between *(Left)* yeast SG-enriched mRNAs and mitochondrial-localized genes or *(Right)* yeast SG-depleted mRNAs and mitochondrial-localized genes. (B) Venn diagram illustrating the overlap between *(Left)* yeast SG-enriched mRNAs and SRP localized genes or *(Right)* SG-depleted mRNAs and SRP localized genes. (C) Venn diagram illustrating the overlap between *(Left)* human SG-enriched mRNAs and mitochondrial-localized genes or *(Right)* SG-depleted mRNAs and mitochondrial-localized genes. (D) Venn diagram illustrating the overlap between *(Left)* human SG-enriched mRNAs and SRP localized genes or *(Right)* SG-depleted mRNAs and SRP localized genes.

Table S1A: Most enriched cytosolic mRNAs in SGs. Related to Figures 1-2.

Gene symbol	Gene description	SG enrichment	mRNA length (bases)	Number of mRNAs in SGs/cell	Number of mRNAs in cell	% of mRNAs in SGs
AHNAK2	AHNAK nucleoprotein 2	11.00	18,335	5.1	4.0	>95*
RP11-438J1.1	Not available	10.84	2,831	1.4	1.1	>95*
BCL9L	B-cell CLL/lymphoma 9 like	9.50	10,005	11.5	10.4	>95*
MAP1A	microtubule associated protein 1A	9.46	10,553	5.9	5.3	>95*
MACF1	microtubule-actin crosslinking factor 1	9.40	24,828	34.7	31.4	>95*
CHD7	chromodomain helicase DNA binding protein 7	9.17	11,568	11.6	10.8	>95*
MAML1	mastermind like transcriptional coactivator 1	8.51	6,559	4.1	4.1	>95*
CDR1	cerebellar degeneration related protein 1	8.43	2,467	2.1	2.1	>95*
KMT2D	lysine methyltransferase 2D	8.33	19,419	11.2	11.4	>95*
HERC2	HECT and RLD domain containing E3 ubiquitin protein ligase 2	7.90	15,337	15.9	17.1	>95*

*% of mRNAs in SGs greater than 95%

Table S1B: Most abundant cytosolic mRNAs in SGs. Related to Figures 1-2.

Gene symbol	Gene description	SG enrichment	mRNA length (bases)	Number of mRNAs in SGs/cell	Number of mRNAs in cell	% of mRNAs in SGs
ACTB	actin beta	0.44	2,245	178.2	3440	5.2
CHD4	chromodomain helicase DNA binding protein 4	3.56	6,554	159.2	381	41.8
EIF4G2	eukaryotic translation initiation factor 4 gamma 2	1.34	6,769	154.3	981	15.7
THBS1	thrombospondin 1	2.78	7,775	149.2	457	32.7
EEF2	eukaryotic translation elongation factor 2	0.89	3,164	142.9	1364	10.5
HSPA8	heat shock protein family A (Hsp70) member 8	0.98	2,463	106.3	929	11.4
PABPC1	poly(A) binding protein cytoplasmic 1	1.95	3,485	101.5	444	22.9
ERBB2IP	erbb2 interacting protein	4.76	8,647	100.9	181	55.9
PKM	pyruvate kinase, muscle	0.66	3,016	98.5	1264	7.8
COL1A1	collagen type I alpha 1 chain	2.08	6,727	77.2	317	24.4

Table S2: Most enriched mRNA in yeast. Related to Figure S6.

Gene symbol	Description	SG enrichment	RNA length (bases)
FIG2	Factor-induced gene	8.82	5721
DAN4	Delayed ANaerobic	8.28	3544
BNI1	Bud Neck Involved	6.64	6265
HKR1	Hansenula mrakii Killer toxin Resistant	5.83	5582
FLO1	FLOculation	5.83	3228
NFT1	New Full-length MRP-type Transporter	5.32	3657
SAG1	Sexual AGglutination	5.25	2091
PRP8	Pre-mRNA Processing	5.21	7320
PMD1	Paralog of MDS3	5.12	5447
HSL1	Histone Synthetic Lethal	4.71	4912

Table S3A: Most enriched lincRNA in mammalian SGs. Related to Figure 5.

Gene symbol	SG enrichment	RNA length (bases)	Number of RNAs in SGs/cell	Number of RNAs in cell	% of RNAs in SGs
RP11-1100L3.8	8.10	687	0.79	0.83	95
RP11-123O10.4	5.05	192	0.99	1.64	59
CTC-444N24.11	4.99	4,258	1.38	2.35	58
CTC-459F4.1	4.99	3,666	0.47	0.81	58
LINC00657 (NORAD)	4.62	5,339	11.1	20.5	54
AC004449.6	4.34	433	0.65	1.29	51
CITF22-49D8.1	4.19	4,364	0.64	1.31	49
RP11-521L9.1	4.13	260	0.92	1.90	48
RP11-284F21.9	4.08	3,222	0.72	1.51	48
BAIAP2-AS1	4.07	4,465	3.30	6.97	48

Table S3B: Most enriched SUTs in yeast SGs. Related to Figure 6.

Gene symbol	SG enrichment	Average SG reads	RNA length (bases)
ST4580	5.95	11511	4280
ST4796	5.29	991	1344
ST6823	4.99	2226	1880
ST6969	4.71	12772	3280
ST2060	4.67	2963	2728
ST6465	4.36	2678	1824
ST2371	4.08	4158	1824
ST1657	3.99	6353	3528
ST1280	3.82	2303	1664
ST5030	3.77	5183	2248

DISCUSSION

Based on determining the SG transcriptome in yeast and mammalian cells we demonstrate several principles of RNA accumulation in SGs. One key observation is that SGs are primarily accumulations of diverse mRNAs. For example, analyses in mammalian cells showed SG are >78% mRNAs, although some ncRNAs can accumulate within SGs **(Figure 2, 5)**. Similarly, in yeast RNA-Seq genes indicates ~95% of the molecules in SGs are mRNAs. More strikingly, there is no overly abundant RNA within SGs, with the most abundant mRNAs in mammalian SGs, actin, only representing ~0.5% of the SG mRNAs **(Figure 2E, Table S1B)**. Thus, SGs are a true composite of many mRNAs, which is also supported by the fact that essentially every mRNA is present in SGs to some extent. This is not an artifact of contaminating abundant RNA molecules since smFISH for GAPDH, an abundant mRNA depleted from SGs, affirms that 4% of the GAPDH mRNA molecules are present within SGs (**Figure 1**). The fact that essentially every mRNA can accumulate in SGs suggests that the interactions required for SG accumulation are generic and are not limited to a specific subset of mRNAs.

Two key observations argue that we are effectively identifying the majority of RNAs within SGs. The first observation that we effectively identify the full spectrum of SG RNAs is that smFISH analysis of individual mRNAs, which identifies mRNAs anywhere within SGs, and the RNA-Seq analysis, which by definition is only sequencing RNAs present in the core region, show a strong linear correlation **(Figure 2B)**. The second key observation is that based on the quantification of the RNA-Seq of SG cores,

we calculate that ~9.4% of the total mRNAs in the cell accumulate in SGs, which is similar to the 13% of total oligo(dT) staining that resides in SGs under the same conditions **(Figure 2F, G)**. Thus, while it is possible we are selectively losing a subset of RNAs loosely associated with SGs, these two observations suggest we are efficiently identifying the vast majority of mRNAs associated with SGs. One simple explanation is that the staining of both proteins and RNAs in the shell region of SGs (Jain et al., 2016) is due to mRNAs of sufficient length that they have portions residing within the core region, and thus co-purify with cores after cell lysate. However, two caveats should be noted. First, since we are purifying SG cores using G3BP1 as an affinity ligand, it is possible we are missing some mRNAs in SG, particularly if they accumulate in SG cores that are lacking in G3BP. Second, we may be underestimating the fraction of some RNAs in SGs if specific mRNAs in the shell region region are lost during the purification process.

Two observations led to the striking conclusion that only 10-12% of the total mRNA molecules in U-2 OS cells accumulate in SG. First, 12-13% of the poly(A)+ RNA signal accumulates in SG based on oligo(dT) FISH (**Figure 2F, G**). Second, by quantifying the distribution of the mRNA population, we calculated that approximately 9-10% of the mRNA molecules in the cell assemble into SG (**Figure 2E**). Moreover, based on our standardization of the RNA-Seq data by smFISH analysis (**Figure 2A, B**), we estimate that there are only 185 genes in which greater than 50% of the mRNA molecules are present in SGs. Thus, only a minority of the bulk mRNA population enters SGs and only for <2% of the analyzed genes are the majority of the transcripts within

SGs.

The diversity and overall minor amount of mRNA within SGs force us to reconsider their nature and possible function. First, the small amount of mRNA within SGs suggests that these structures will not have large global effects on bulk mRNA. This is consistent with observations that cell lines lacking G3BP, which fail to form SG, still efficiently repress global translation of mRNAs (Kedersha et al., 2016), and cells with defective SG formation do not show altered mRNA degradation (Bley et al., 2015). Thus, if SGs directly affect the function of mRNAs within them, SG formation will only have a substantial impact on the few mRNAs with a high percentage of their total molecules present in SG. It remains possible that sequestering ~10-12% of the mRNAs in the cell has other impacts on cell physiology, such as affecting signaling pathways (Kedersha et al., 2013).

The transcriptome of SGs argues that every SG will be different in composition. Since total SG volume in U-2 OS cells is 10 μm^3, and each core is ~0.0066 μm^3 (based on the size of cores using nanoparticle tracking (Jain et al. 2016)), we estimate that there are between 300 and 1500 SG cores in U-2 OS cells after 60' of arsenite stress if SGs consist of 20% (estimated in Jain et al., 2016) or 100% (a maximal boundary condition) cores, respectively. Since actin mRNA, the most abundant molecule in SGs, has only 160 molecules within SGs/cell and there are at least 300 cores per cell by our most conservative estimate, core RNA composition must be heterogeneous. Moreover, at 1 hour of arsenite stress, we estimate there are 32,000 mRNAs in SGs (Figure 2E).

Thus, each core should have approximately 21 to 106 mRNA molecules, and by necessity cores would need to be composed of a heterogeneous mixture of transcripts to explain our transcriptome data.

We identify three parameters that correlate with the partitioning of specific mRNAs into SGs. First, both yeast and mammalian analyses show inefficient translation correlates with higher partitioning of mRNAs into SGs **(Figure 4, 5)**. This is consistent with prior models that mRNAs engaged with ribosomes are restricted from entering SGs and provides additional evidence that mRNAs need to be nontranslating to accumulate within SG (Kedersha et al., 2000) (Kedersha and Anderson, 2002). Second, association with ER or mitochondrial membranes correlates with mRNAs being excluded from SGs, perhaps due to membrane tethering limiting the mRNPs from diffusing into SGs **(Figure S6)**. Third, a strong and predominant metric of mRNA accumulation in SGs is the overall length of the coding region and the 3'UTR in both yeast and mammalian SG transcriptomes and in both organisms is strongly linked to the length of the ORF (**Figure 4, 6**).

In principle, a longer mRNA could be more efficiently localized to SGs by two potentially overlapping mechanisms. First, a longer coding region could increase SG partitioning by providing multiple binding sites for RNA binding proteins that may target mRNAs into SGs**.** Although it has been tacitly assumed that mRNAs would enter SGs through the interactions of specific RNA binding proteins, to our knowledge there is no direct evidence in support of this model. Indeed, although actin mRNA has been

localized to SG, its association with SG is independent of the ZBP1 RNA binding protein (Stöhr et al., 2006). Moreover, we observe only a very modest bias in the binding of SG RNA binding proteins to mRNAs enriched in SG **(Figure 7B)**. Thus, while mRNA binding proteins can clearly play a role in the overall assembly of SGs (Panas et al., 2016; Protter and Parker, 2016), whether they dictate the specific mRNAs localized to SGs remains to be determined. An alternative model to long ORFs simply providing more RBP sites is that the condensation of mRNAs into SGs would be promoted by RNA-RNA interactions between the exposed coding regions, and any 3' UTR regions exposed due to the off-rate of mRNA binding proteins and their redistribution into exposed coding regions. In this model, the longer the coding region and/or the 3' UTR, the more RNA-RNA interactions that could occur. Thus, a key area of future research will be in determining what interactions define the SG transcriptome, and how that affects mRNA function.

MATERIALS AND METHODS

U-2 OS growth conditions and reagents

Human osteosarcoma U-2 OS cells expressing G3BP1-GFP (Figley et al., 2014), U-2 OS cells and U-2 OS ΔΔG3BP1/2 (Kedersha et al., 2016), maintained in DMEM with high glucose, 10% fetal bovine serum, and 1% penicillin/streptomycin at 37°C/5% CO_2, were used in all experiments.

Isolation of RNA from U-2 OS cells and SG Cores for RNA-sequencing

U-2 OS cells expressing G3BP1-GFP were grown to 85% confluency in three 500 cm^2 square TC-treated culture dishes (Thermo Fisher Scientific, 07-200-599). One hour prior to stress, cell culture media was exchanged with fresh media. Cells were then treated with 500 µM $NaAsO_2$ for one hour at 37°C/5% CO_2. After stress, cells were washed once with media, transferred to falcon tubes, and pelleted at 1,500 x g for 3 mins. Upon aspirating the media, the pellets were immediately flash-frozen in liquid N_2 and stored at -80°C until isolation of mammalian SG cores was performed.

The isolation of SG cores was adapted from two papers (Jain et al., 2016; Wheeler et al. 2017). Briefly, the pellet was thawed on ice for 5 mins, re-suspended in 1 mL SG lysis buffer (50 mM TrisHCl pH 7.4, 100 mM KOAc, 2 mM MgOAc, 0.5 mM DTT, 50 µg/mL Heparin, 0.5% NP40, cOmplete mini EDTA-free protease inhibitor (Sigma Aldrich, 11836170001), 1 U/µL of RNasin Plus RNase Inhibitor (Promega, N2615) and passed through a 25 gauge 5/8 needle attached to a 1 ml syringe 7 times. After lysis, the lysates were spun at 1000 x g for 5 min at 4°C to pellet cell debris. 50 µL and 950 µL of the supernatants were transferred to new microcentrifuge tubes for isolating total and SG RNAs respectively (**Figure S1A**). For isolating total RNA, Trizol LS reagent (Thermo Fisher Scientific, 10296-028) was added and RNA was extracted following the manufacturers protocol. Following isopropanol precipitation, the RNA pellet was re-suspended in 50 µL RNase-free H_2O.

The following steps were performed to isolate mammalian SG cores and extract its RNA: 1) The 950 µL supernatant was spun at 18,000 x g for 20 mins at 4°C to pellet SG cores. 2) The resulting supernatant was discarded and the pellet was re-suspended in 1 mL SG lysis buffer. 3) Steps 1 and 2 were repeated to enrich for SG cores. 4) The resulting pellet was then re-suspended in 300 µL of SG lysis

buffer and spun at 850 x g for 2 mins at 4°C. 5) The supernatant which represents the mammalian SG core enriched fraction was transferred to a new tube. 6) The enriched fraction was pre-cleared twice by adding 60 µL equilibrated DEPC-treated Protein A Dynabeads (Thermo Fisher Scientific, 10002D) and nutating at 4°C for 30 mins. Dynabeads were removed using a magnet. 7) 20 µL of anti-GFP antibody (Life technologies, A1112) was added to the enriched fraction and nutated at 4°C for 1 hr to affinity purify SG cores. 8) The solution was spun at 18,000 x g for 20 mins at 4°C and the supernatant was discarded to remove any unbound antibody. 9) The pellet was then re-suspended in 500 µL SG lysis buffer and 60 µL equilibrated DEPC-treated Protein A Dynabeads. 10) The sample was nutated for 3 hr at 4°C. 11) The Dynabeads were washed three times with wash buffer 1 (20 mM Tris HCl pH 8.0, 200 mM NaCl, 1 U/µL of RNasin Plus RNase Inhibitor) for 5 mins, once with wash buffer 2 (20 mM Tris HCl pH 8.0, 500 mM NaCl, 1 U/µL of RNasin Plus RNase Inhibitor) for 5 mins, and once with wash buffer 3 (SG lysis buffer + 2M Urea, 1 U/µL of RNasin Plus RNase Inhibitor) for 2 mins at 4°C. 12) The beads were resuspended in 200 µL of 100 µg/mL Protease K solution (1X TE buffer, 2M Urea, 1 U/µL of RNasin Plus RNase Inhibitor) and incubated for 15 mins at 37°C. 13) Trizol LS reagent (Thermo Fisher Scientific, 10296-028) was added to the samples and RNA was extracted following the manufacturers protocol. Following isopropanol precipitation, the RNA pellet was re-suspended in 20 µL RNase-free H_2O.

Library construction and RNA-sequencing

After isolating RNA from cells and stress granule cores 1 and 20 µL of RNA isolated from cells and SG cores, respectively, were treated with rDNase 1 following manufacturer's protocol (Thermo Fisher Scientific, AM1906). Immediately, the samples were ribosomal-depleted using Ribo-Zero rRNA removal kit following manufacturer's protocol (Illumina, MRZH11124). Afterwards, the quality of the RNA was inspected using high sensitivity RNA ScreenTape (Agilent, 5067-5579) on Agilent TapeStation 2200 instrument (University of Colorado-Boulder BioFrontiers Next-Generation Sequencing Facility).

cDNA libraries were prepared following manufacturer's protocol from three replicates of 10 ng ribosomal-depleted RNA from cells and SG cores using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina (New England Biolabs, E7420S). The qualities and amounts of the cDNA libraries were assessed using high sensitivity D1000 ScreenTape (Agilent, 5067-5584) on Agilent Tapestation 2200

instrument and the Qubit (Thermo Fisher Scientific), respectively (University of Colorado-Boulder BioFrontiers Next-Generation Sequencing Facility) The cDNA libraries (3 from cells and 3 from SG cores) were pooled and sequenced on the Illumina NextSeq 500 platform using the High Output 150 cycles kit (paired-end reads, single index) (University of Colorado-Boulder BioFrontiers Next-Generation Sequencing Facility).

Yeast Library Construction Libraries were constructed using the Bioo Scientific NEXTflex Rapid Illumina Directional RNA-Seq Library Prep Kit. We followed the manufacturer supplied directions. Prior to making the library, the RNA was ribosome depleted using the Ribo-Zero rRNA removal kit, also according to manufacturer's directions. The resulting RNA was incubated with Turbo DNase at 37 C for 1 hour and then cleaned up using a Zymo Research RNA Clean and Concentrator kit also according to the manufacturer's directions.

Bioinformatics analysis of mammalian SG cores

Yeast Sequencing Data Analysis

Read quality was assessed using fastqc. Illumina adaptors were trimmed using Trimmomatic 0.32 in single end (SE) mode(Bolger et al., 2014). An index genome was built with Bowtie 0.12.7 using the 'build' command and the S288C reference genome version (R64-1-1 .fasta file), which was acquired through Saccharomyces Genome Database (SGD)(Langmead et al., 2009). Reads were aligned using Bowtie 0.12.7 using the following options: -S -v 2 -m 3 –best.

Reads mapping to each gene were counted using HTSeq with parametes –t gene –I Name –f sam –s reverse, using the R64-1-1 annotation file(Anders et al., 2015). Normalization and differential expression analysis were then performed using DESeq 1.22.1(Anders et al., 2010).

SUTs and CUTs were analyzed using the same method, but with a different annotation file gene annotation file that contained these transcript annotations (REF).

Yeast Length, Translation, and Stability analysis

Length: Data for the length analysis was obtained from a previous report (Nagalakshmi et al., 2008).

Codon Optimality: The fraction of optimal codon for each gene was calculated using a custom script

based on codon stabilization coefficients presented in a previous report (Presnyak et al., 2016). In brief, the amount of each codon was tallied on a gene by gene basis. Optimal codons were assigned a value of one while non-optimal codons were assigned a value of zero. The total number of optimal codons was then divided by the total number of codons in order to find fraction of optimal codons. Predicted translation initiation rates and ribosome density data was obtained from a previous study (Siwiak and Zielenkiewicz, 2010) *Stability:* RNA half-lives were obtained from a previous report (Neymotin et al., 2014).

All boxplots were generated using PRISM. Scatterplots were created using PRISM of Tableau.

Mammalian Sequencing Data Analysis

Read quality was assessed using fastqc. Illumina adaptors were trimmed using Trimmomatic 0.32 in paired and (PE) mode (Bolger et al., 2014). An index genome was acquired from NCBI build 37.2. Reads were aligned using Tophat (version 2.0.6) and Bowtie2 (version 2.0.2) and the following parameters: --b2 –fast –microexon-search (Kim et al., 2013; Langmead and Salzberg, 2012). Differential expression analysis was performed using Cuffdiff (version 2.2.1) with the default parameters (Trapnell et al., 2013).

Long non-coding RNAs were analyzed in a similar manner, however a different a different annotation file from GENCODE (Release 19 GRCh37.p13) (Harrow et al., 2012).

Human Length, Translation, and Stability analysis

All length data (5'UTR, ORF, 3' UTR and total) was acquired using Ensemble's Biomart tool (Aken et al., 2016; Kinsella et al., 2011). Ribosome protected fragment measurements were acquired from a previous report, which examined ribosome protected fragment during arsenite induced stress as well as a no stress control (Andreev et al., 2015). Ribosome density was then calculated by normalizing read counts by dividing by total transcript length. GC content data was acquired using Ensemble's Biomart tool. Half-life data was acquired from a previous report (Tani et al).

CLIP Data analysis

All CLIP data was analyzed from a previous report, which analyzed eCLIP targets of over 80 proteins (Van Nostrand et al., 2016). We examined eCLIP data from 11 known RBPs in the SG proteome (Jain et al., 2016). In brief, a bed file for each eCLIP experiment was obtained from the ENCODE data repository, which contained the genomic coordinates of all CLIP peaks. Next, a bed file was created for the human genome using NCBI's gtf annotation file. The program BEDOPS v2.4.26 was used performed to convert the gtf file into a bed file, using the command 'gtf2bed' (Neph et al., 2012). The human genome bed file was then compared to each eCLIP bed file using the 'intersect' command in bedtools version 2.16.2, in order to see which genes contained eCLIP peaks (Quinlan and Hall, 2010).

Machine learning

A mathematical model for mRNA localization during stress was built using MATLAB's "Classification Learner" application. For yeast mRNA localization, two parameters (length and fraction optimal codons), and a response (mRNA localization) were imported into the classification learner. The model used 90% of the data for training and held back 10% for testing. This process was performed for multiple iterations until every mRNA had been used for both model building and model testing. The course nearest neighbor model provided the best results, reaching 75.9% accuracy. In brief, the course nearest neighbor model makes localization predictions of a given mRNA based on the localization of its 100 nearest neighbors on a scatterplot of length vs codon optimality. A similar approach was used for modeling mammalian mRNA localization, but instead using ribosome density instead of codon optimality. The course nearest neighbor model was 74.5% accurate in predicting mammalian RNA localization.

Statistical Tests

R^2 values were calculated for pairwise comparison of sequencing replicates by using the squaring MATLAB's 'corr' function (which reports the Pearsno's correlation coefficient). To calculate the statistical significance in boxplots, a 2-sample t-test was performed using PRISM software. In order to calculate the statistical significance of the overlap between SG RNA and CLIP targets, a 2 population proportion test was performed in R using the option "alternative=greater". A 2 population proportion test was also used to calculate the overlap of yeast SG enriched and SG depleted mRNA with SRP-localized and

mitochondria-localized mRNA.

Stress conditions

To examine mRNA localization during other stresses we performed the used the following stress conditions. Prostaglandin J2 was added to a final concentration of 50uM for 1 hour. Thapsigargan was added to a final concentration X for 1 hour. For heat shock experiments, cells were placed at 42 degrees Celsius for 1 hour. For osmotic stress, cells were stressed in .5M sorbitol for 2.5 hours. For smFISH cells were fixed after the completion of each stress as previously described below.

Single molecule FISH

Ship ready Stellaris® FISH Probes recognizing TFRC and POLR2A transcripts labeled with Quasar 570 Dye and GAPDH labeled with Quasar 670 Dye (SMF-2006-1, SMF-2003-1, and SMF-2026-1, respectively, Biosearch Technologies, Inc., Petaluma, CA) and Custom Stellaris® FISH Probes designed against AHNAK and DYNC1H1 transcripts labeled with Quasar 670 dye by utilizing the Stellaris® RNA FISH Probe Designer (Biosearch Technologies, Inc., Petaluma, CA), available online at www.biosearchtech.com/stellarisdesigner (version 4.2), were hybridized to fixed U-2 OS G3BP1-GFP cells by following the manufacturer's instructions available online at www.biosearchtech.com/stellarisprotocols.

Microscopy and image analysis

Fixed U-2 OS G3BP1-GFP cells stained with Stellaris® FISH Probes and DAPI were imaged using a DeltaVision Elite microscope with a 100X objective using a PCO Edge sCMOS camera. At least 12 images were taken for each experiment comprising of 5 Z-sections each. All images shown in the manuscript are one Z-plane and image processed by ImageJ with FIJI plugin and Adobe Photoshop. Quantification of single molecule FISH spots in SGs, in cells and area of SGs were determined using Imaris Image Analysis Software (Bitplane) (University of Colorado-Boulder, BioFrontiers Advanced Light Microscopy Core). 20 cells were counted for each replicate; Therefore, a total of 60 cells were counted (n=3).

CHAPTER 4

RNA PARTITIONING INTO STRESS GRANULES IS INDEPENDENT OF G3BP

SUMMARY

Stress granules (SGs) are stress induced RNA-protein assemblies formed from untranslating mRNPs. Mammalian SGs are non-uniform and contain “cores”, which are stable in lysates and heterogenous in protein composition. The interactions defining RNA recruitment into SGs, and whether the RNA composition varies between different cores remains unclear. Herein, we determine the SG transcriptome through purification of both PABPC1 and G3BP SG cores during both arsenite and sorbitol stress. We observe that similar RNAs are recruited to SGs independent of the protein used for core purification, suggesting individual RNA-binding proteins (RBPs) may play a limited role in defining the SG transcriptome. Analysis of the sorbitol-induced SG transcriptome reveals G3BP1/2 has little effect on RNAs recruited to SGs suggesting RNAs localize to SGs independent of individual RBPs. Taken together, we suggest that partitioning of RNAs to SGs is independent of at least some proteins, consistent with SGs forming through intermolecular RNA-RNA interactions.

INTRODUCTION

Stress granules (SGs) are non-membranous ribonucleoprotein assemblies that form during stress when translation is limited (Buchan and Parker, 2009; Protter and Parker, 2016). SGs are of interest because they are cellular markers of translation, are related to similar ribonucleoprotein (RNP) granules in neurons and embryos, share components with toxic aggregates observed in degenerative disease, and play roles in viral response as well as cancer progression (Anderson et al., 2015; Kedersha et al., 2013; Kim et al., 2013b; Li et al., 2013; Reineke and Lloyd, 2013; Somasekharan et al., 2015).

Based on super-resolution microscopy and single molecule analysis, mammalian SGs formed during arsenite stress are non-uniform in nature and contain local regions of protein concentration, referred to as SG cores (Jain et al., 2016; Niewidok et al., 2018; Wheeler et al., 2016). SG cores are stable sub-assemblies in cell lysates (Jain et al., 2016; Wheeler et al., 2016). Interestingly, SG cores have been shown to be heterogeneous in composition by staining of SG cores in cell lysates, which showed that PABPC1 and G3BP, two standard SG components, only show approximately 40% colocalization in cores (Jain et al., 2016). Moreover, super resolution microscopy has shown that PABP and G3BP do not show perfect colocalization (Jain et al., 2016). This core heterogeneity raises the possibility that different types of cores, as defined by their

protein composition, would recruit different RNAs into SGs.

SG cores have been purified from GFP-G3BP expressing U2OS cells through differential centrifugation and GFP pulldown (Khong et al., 2017a; Wheeler et al., 2017). This SG purification strategy allowed for the discovery of a diverse SG proteome composed of numerous RNA binding proteins forming a dense protein-protein interaction network (Jain et al., 2016). More recently, a similar SG protein composition was detected by proximity labeling methods (Youn et al., 2018; Markmiller et al., 2018). The fact that SGs are composed of a dense protein-protein interaction network is consistent with the recruitment of proteins to SGs through protein-protein interactions, which might then recruit specific mRNAs bound to those RNA binding proteins.

To discern the mechanisms that recruit RNAs into SGs, we previously purified and sequenced the SG transcriptome in U2OS cells expressing GFP-G3BP via differential centrifugation and immunoprecipitation using a GFP antibody and RNA-Seq (Khong et al., 2017, 2018). This analysis revealed that long, poorly translated transcripts preferentially enrich in SGs (Khong et al., 2017). While we found that SG-enriched transcripts showed a modest bias for eCLIP sites for SG RNA binding proteins, this analysis was not nearly as predictive as transcript length alone (i.e. no single RNA binding seems to strongly correlate with RNA recruitment to SGs). This raised the possibility that RNA recruitment to SGs might be affected by other aspects of RNA and may not simply be due to protein-protein interactions within SGs. This hypothesis is also supported by the observation that the NORAD RNA, which has at least 17 Pumilio

binding sites, is recruited to SGs independently of the Pumilio RBP (Lee et al., 2016; Namkoong et al., 2018).

An alternative possibility is that RNA-RNA interactions might play a role in defining the SG transcriptome. Consistent with this model, the transcriptome of protein-free *in vitro* RNA assemblies, formed from total yeast RNA, shows remarkable overlap with the transcriptome of yeast SGs (Van Treeck et al., 2018). This suggests a model wherein RNAs can be recruited to SGs by both RNA-RNA and protein-protein interactions. However, the relative importance of these interactions, and how they might contribute to the mRNAs and their organization within SGs, is unknown.

Herein, we examine how proteins influence the recruitment of RNAs to SGs by (1) examining the recruitment of RNAs to G3BP and PABPC1 cores and (2) examining how deletion of major SG RNA binding proteins, G3BP1 and G3BP2, affects RNA recruitment to SGs. We report that both sorbitol and arsenite-induced SG cores purified by differential centrifugation and immunoprecipitation with PABPC1 share a similar RNA composition to SGs purified via immunoprecipitation of GFP-G3BP. Moreover, we present evidence that the RNA composition of SGs is largely unaffected by the presence or absence of G3BP, suggesting individual RBPs may play a limited role in defining which RNAs localize to SGs. Interestingly, under sorbitol-induced hyperosmotic stress, SGs form a more punctate structure than observed during arsenite-induced stress, with groups of interacting proteins condensing into tight puncta, surrounded by mRNAs. Taken together, we suggest that SGs can recruit RNAs in a

protein independent manner, and proteins associated with SG RNAs can subsequently strengthen interactions and also rearrange to form a complex substructure.

RESULTS

G3BP and PABPC1 stress granule cores share similar transcriptomes

In order to determine if the protein composition of different SG cores correlated with a specific subset of mRNAs, we purified and sequenced SG cores during arsenite stress using an antibody to PABPC1, instead of GFP-G3BP1 as done previously (Khong et al., 2017**)(Figure 4.1A)**. This experiment is based on the observation that during arsenite stress SG cores can be heterogeneous in composition, since only approximately 40% of G3BP1-purified SG cores were previously found to contain PABPC1 (Jain et al., 2016). Since SG cores can be heterogeneous in protein composition, we sought to test whether the RNA composition of SG cores changed if we pulled down with a different SG protein, PABPC1.

Based on three biological replicates, both total RNA and SG core purification via PABPC1 pulldown yielded reproducible transcriptomes **(Figure 4.1B)**. SG RNA transcriptomes based on the purification of SG cores with PABPC1 antibody showed little similarity to total RNA transcriptomes **(Figure 4.1B),** with 1189 transcripts enriched in SGs ($p < .01$, Fold Change (FC) > 2, FPKM > 1), and 1785 depleted from SGs ($p<.01$, FC < .05, FPKM > 1) **(Figure 4.1C)**. Thus, similar to the analysis of G3BP1-purified cores, PABPC1-purified SG cores defined a specific set of enriched and

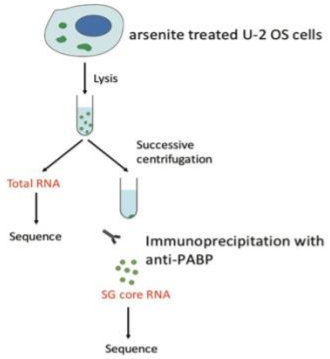
depleted RNAs.

Comparison of the RNAs enriched in PABPC1 SG cores showed that these RNAs were very similar to the SG transcriptome defined by GFP-G3BP core pulldown. Specifically, we found that SG core enrichment scores from PABPC1-purified cores and GFP-G3BP-purified cores showed a strong linear correlation **(Figure 4.1D R = .83)**. Thus, PABPC1 pulldown identifies a population of RNAs that strongly overlaps with the GFP-G3BP pulldown of the SG transcriptome. Consistent with PAPBPC1 and GFP-G3BP purification identifying similar RNAs, we observed that in both cases, RNAs enriched in SGs are biased towards long, poorly translated RNAs (Khong et al., 2017 & **(Figure 4.1E-G)**.

We reasoned that poly(A) tail length might also play a role in localizing RNAs to SGs, since PABPC1 is a component of SGs. By comparing our data to a recent study which examined poly-A tail length in HeLa cells (Subtelny et al., 2014), we found that RNAs that are depleted from SGs tend to have shorter poly-A tails **(Figure 4.1H)**. A bias in the poly(A) tail length of mRNAs in SG could be a yet to be explained effect of the poly(A) tail, or could be because more stable mRNAs tend to be efficiently translated and have shorter poly(A) tails at steady state (Lima et al., 2017).

Taken together our findings suggest that PABPC1 and G3BP SG cores share a similar RNA composition and that the RNAs that co-purify with PABPC1 and G3BP cores have similar physical properties.

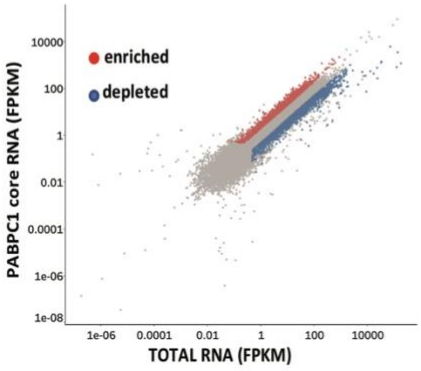
A



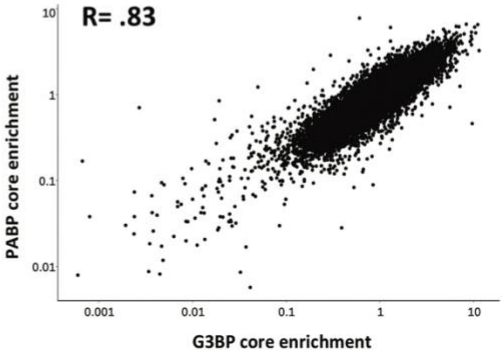
B

r^2	SGrep1	SGrep2	SGrep3	TOTrep1	TOTrep2	TOTrep3
SGrep1	1	0.953	0.935	0.156	0.159	0.152
SGrep2	0.953	1	0.934	0.192	0.194	0.186
SGrep3	0.935	0.934	1	0.161	0.137	0.137
TOTrep1	0.156	0.192	0.161	1	0.941	0.937
TOTrep2	0.159	0.194	0.137	0.941	1	0.966
TOTrep3	0.152	0.186	0.137	0.937	0.966	1

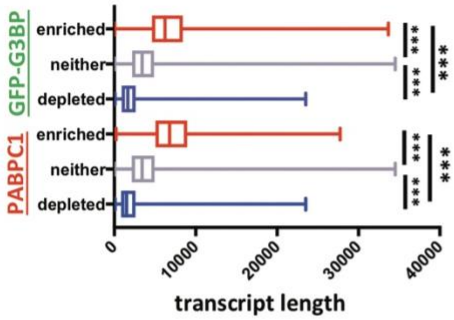
C



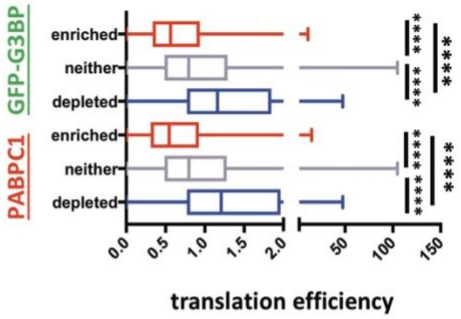
D



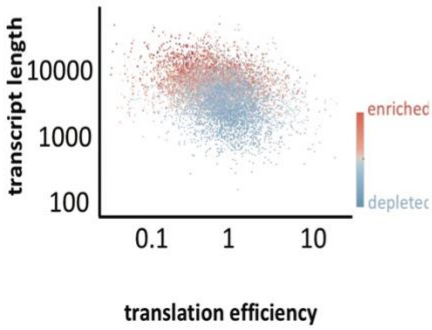
E Lorem ipsum



F



G



H

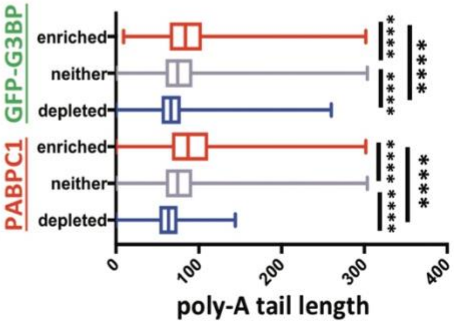


Figure 4.1: RNA-Seq of the PABPC1 stress granule core transcriptome

(A) Schematic of PABPC1 SG core purification. **(B)** Table showing pairwise r^2 values for comparisons across SG RNA and total RNA replicates. **(C)** Scatterplot of RNA-Seq analysis of PABPC1 SG core purified RNA and total RNA. **(D)** Scatterplot of RNA enrichment in PABPC1 cores vs enrichment in G3BP cores. **(E)** Boxplots showing transcript length distributions for RNAs purified from PABPC1 and G3BP SG cores. **(F)** Same as in E but for translation efficiency. **(G)** Scatterplot of transcript length vs. translation efficiency. Color-coded by degree of enrichment in SGs. **(H)** Same as E & F, but for poly-A tail length.

The stress granule transcriptome is conserved during different stresses

To determine if PABPC1 and G3BP1 define similar SG transcriptomes in different stresses, we also purified and sequenced SG cores based on PABPC1 and G3BP-GFP purification during hyperosmotic stress induced by sorbitol, which is known to induce SG formation (Dewey et al., 2011; Kedersha et al., 2016).

First, we sequenced the transcriptome of sorbitol-induced SGs by pulldown of GFP-G3BP. We found that purification of GFP-G3BP SGs yielded 646 significantly enriched and 1513 significantly depleted transcripts **(Figure 4.2A)**. In agreement with the previous observations that the SG transcriptome is conserved between multiple stresses (Khong et al., 2017; Namkoong et al., 2018), we found that GFP-G3BP1 pulldown of arsenite-induced SGs showed a very similar transcriptome to GFP-G3BP pulldown of sorbitol-induced SGs **(Figure 4.2B)**. Similarly, the transcriptome of PABPC1 cores, like G3BP cores, is largely conserved between arsenite and sorbitol stress **(Figure 4.2C & D)**. These observations argue that the RNA composition of SGs remains largely consistent between different stresses.

Moreover, the SG enrichment scores of individual mRNAs in either PABPC1 cores and G3BP cores purified from sorbitol treated cells showed a strong correlation **(Figure 4.2E),** providing additional evidence that PABPC1 and G3BP cores contain a similar population of RNAs.

Taken together, our results indicate that the SG transcriptome is largely conserved between stress conditions, and that the mRNAs pulled down are

independent of the protein used for the affinity purification. This observation raises the possibility that proteins may play a minor role in determining the RNAs recruited into SG (see discussion).

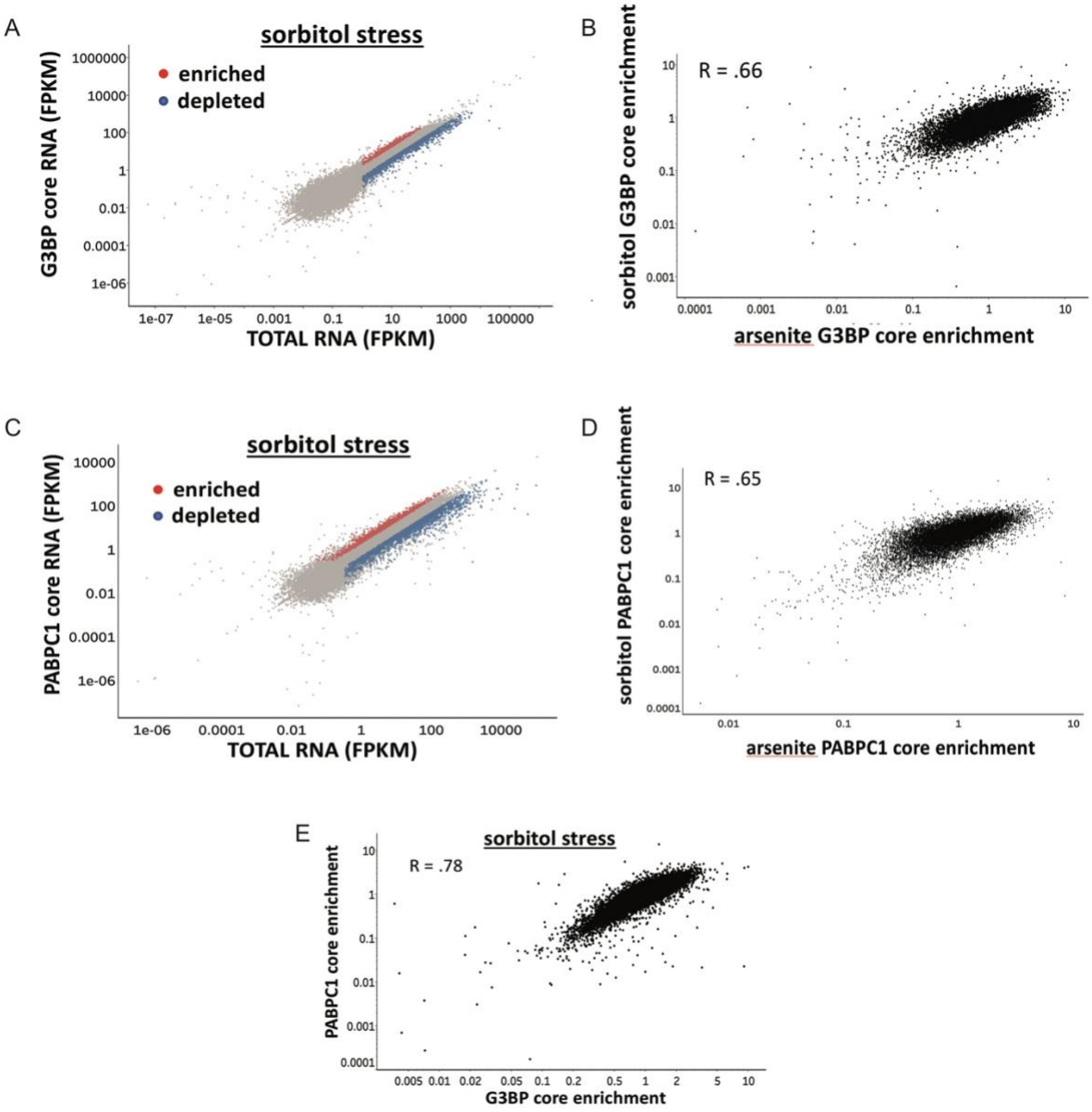


Figure 4.2: RNA localization to stress granule cores is conserved during stress

(A) Scatterplot of RNA-Seq analysis of GFP-G3BP SG core purified RNA and total RNA during sorbitol stress. **(B)** Scatterplot showing correlation between G3BP core enrichment scores in sorbitol vs arsenite stress. **(C)** Scatterplot of RNA-Seq analysis of PABPC1 SG core purified RNA and total RNA during sorbitol stress. **(D)** Scatterplot showing correlation between PABPC1 enrichment scores in sorbitol vs arsenite stress. **(E)** Scatterplot showing correlation between PABPC1 and G3BP enrichment scores in SGs during sorbitol stress.

RNA localization to stress granule cores is independent of G3BP

To test whether an RNA binding protein can affect the targeting of specific mRNAs into SGs, we focused on the highly paralogous proteins, G3BP1 and G3BP2, which are among the most abundant RNA binding proteins within SGs (Tourrière et al., 2003a). G3BP1 and G3BP2 are required for SG assembly during arsenite exposure, but are not required for SG formation during sorbitol stress (Kedersha et al., 2016; Tourrière et al., 2003b). If protein-protein interactions can recruit mRNAs into SGs, we would predict that in cell lines lacking G3BP1 and G3BP2, some mRNAs bound by these proteins would no longer be recruited to SGs. Conversely, if RNA-RNA interactions dominate the recruitment of RNAs into SGs, then we would expect the same SG transcriptome in both normal and G3BP1/2 deleted cells. To test these possibilities, we purified SG cores using PABPC1 antibodies from a U2OS G3BP1/2 double knockout cell line (ΔΔG3BP1/2), which forms SGs during sorbitol stress (Kedersha et al., 2016).

We found that sorbitol-induced SGs contain a similar transcriptome regardless of whether cells expressed or lacked G3BP. Indeed, the enrichment scores for these transcriptomes showed a strong linear correlation (R = .93), suggesting that the enrichment of RNA is largely independent of G3BP **(Figure 4.3A)**. We interpret this observation to argue G3BP1/2 does not generally affect the mRNAs recruited to SGs.

A subtler possibility is that the presence of G3BP bound to specific mRNAs might act in concert with other features of the mRNA, such as overall length which can

contribute to RNA enrichment in SG in an RNA based mechanism (Van Treeck et al., 2018; Khong et al., 2017). In this view, G3BP would only affect the recruitment of mRNAs to SGs that were on the cusp of SG enrichment.

To examine this possibility, we looked at the ratio of WT SG enrichment scores to G3BP knockout SG enrichment scores as a function of RNA length. If G3BP acted in concert with other features of RNA to drive localization to SGs, one would anticipate that G3BP would only have an observable effect for transcripts of shorter lengths that are on the verge of accumulating in SGs. We first looked globally at the ratio of WT/G3BP knockout enrichment scores at different lengths. Globally we do not observe a significant change in SG enrichment for small RNAs between WT and ΔΔG3BP1/2 cell lines **(Figure 4.3B)**.

We also performed a similar analysis for only the transcripts that were identified as significantly enriched in WT SGs. Once again, we observed a limited effect of G3BP in changing the partitioning of transcripts to SGs **(Figure 4.3C)**. In fact, we see a consistent WT/G3BP knockout SG enrichment score of less than 1 for short transcripts, suggesting that these transcripts might enrich in SGs even more in G3BP knockout cells **(Figure 4.3C)**.

To further examine the role of G3BP on RNAs within SGs, we also compared our SG-enrichment data to a recent study that examined G3BP1/2 targets by PAR-CLIP (Edupuganti et al., 2017). Consistent with G3BP generally not targeting RNAs into SGs,

we found no significant overlap of G3BP1/2 targets and SG-enriched RNAs **(Figure 4.3D)**. Moreover, in ∆/∆G3BP1/2 cells, G3BP1/2 target transcripts do not drastically change their localization to SGs **(Figure 4.3E)**. Taken together, our results indicate that at least G3BP does not play a significant role in defining the RNA composition of SGs and suggests that individual RNA binding proteins may play a limited role in defining the SG transcriptome.

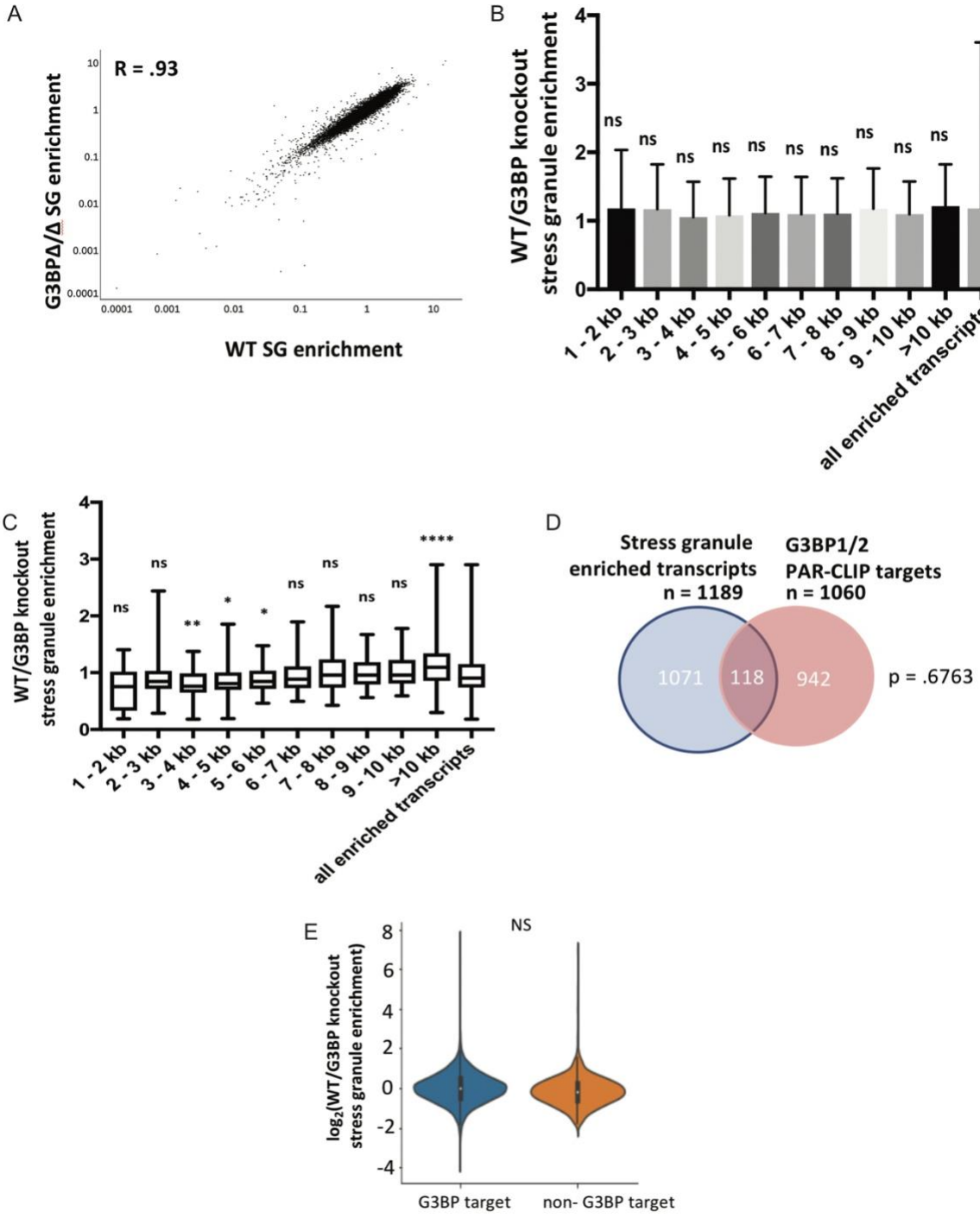

A
R = .93
G3BPΔ/Δ SG enrichment
WT SG enrichment
B
WT/G3BP knockout stress granule enrichment
ns
1 - 2 kb
2 - 3 kb
3 - 4 kb
4 - 5 kb
5 - 6 kb
6 - 7 kb
7 - 8 kb
8 - 9 kb
9 - 10 kb
>10 kb
all enriched transcripts
C
WT/G3BP knockout stress granule enrichment
ns
**
*
****
D
Stress granule enriched transcripts n = 1189
G3BP1/2 PAR-CLIP targets n = 1060
1071
118
942
p = .6763
E
NS
log2(WT/G3BP knockout stress granule enrichment)
G3BP target
non- G3BP target

Figure 4.3: RNA localization to stress granule cores is independent of G3BP

(A) Scatterplot depicting the enrichment of transcripts in PABPC1 stress granule cores in ΔΔG3BP1/2 vs WT U-2 OS cells. **(B)** Boxplots depicting WT/G3BP knockout stress granule enrichment ratios for all transcripts, binned by length. **(C)** Same as B, but only for SG-enriched RNAs. **(D)** Venn diagram depicting overlap between stress granule enriched transcripts and G3BP1/2 PAR-CLIP targets. **(E)** Violin plot showing log2 WT/G3BP knockout stress granule enrichment ratios for G3BP target transcripts and non-G3BP target transcripts.

Hyperosmotic stress reveals specific substructure within stress granules

During our characterization of the sorbitol SG transcriptome, we observed that during sorbitol stress, GFP-G3BP forms a peculiar punctate phenotype that is much more compact than arsenite-induced SGs **(Figure 4.4A)**. To determine if this change in morphology was unique to G3BP, or true for other RBPs, we examined the localization of other SG components relative to G3BP during sorbitol stress.

We first tested the localization of PABPC1 during sorbitol-induced stress. Interestingly, PABPC1 localization does not show a punctate pattern like G3BP, but rather shows a more diffuse staining pattern similar to arsenite-induced SGs **(Figure 4.4B).** FISH for poly(A) tails revealed a similar pattern to PABPC1 (**Figure 4.4C**). This suggests that during sorbitol stress, G3BP forms a more compact and distinct subassembly within a larger SG, defined by PABPC1 and poly(A) tails, than during arsenite stress.

We hypothesized that the tight G3BP puncta might represent a specific subassembly of SGs. A recent analysis of the proximity of SG components to each other based on BioID revealed that SGs are composed of at least two different protein interaction groups, termed Group 1 (which contains PABPC1) and Group 9 (which contains G3BP1 & G3BP2) (Youn et al., 2018). Given these two distinct subgroups of proteins, we hypothesized that the tight puncta of G3BP1 might represent concentrations of Group 9 proteins. To test this prediction, we examined the localization

of additional proteins in Group 9 and Group 1 by immunofluorescence microscopy to see if they overlapped or were distinct from the G3BP1 puncta.

An important result was that both Caprin and FXR1, which are in Group 9, showed a punctate staining pattern that directly colocalizes with G3BP **(Figure 4.4D-F)**. In contrast, three proteins in Group 1 (TIA1, FAM120A, and PUM2), did not perfectly colocalize with G3BP, which is consistent with these proteins preferentially being localized in other regions of SGs. **(Figure 4.4G-I)**. Similarly, in a 3-color fluorescence microscopy experiment, we observed that Group 1 proteins G3BP1 and FXR1 show clear colocalization, but do not colocalize as strongly with FAM120A **(Figure 4.4 J&K)**. Taken together, our analysis provides *in vivo* evidence for at least two distinct granule substructures that are consistent with previous interactome studies (Youn et al., 2018).

Since G3BP1 and G3BP2 are important modulators of SG assembly under certain stress conditions, we hypothesized that they might be required for the formation of the Group 9 protein puncta, and in some stresses those assemblies would be required for SG assembly, while in other stresses, the interactions of the Group 1 proteins would be sufficient. This model would predict that in the ΔΔG3BP1/2 cell lines, SG would still form in sorbitol stress, but they would lack the Group 9 protein puncta. We tested this hypothesis by examining the localization of Caprin in sorbitol-induced SG in the ΔΔG3BP1/2 cell lines. We observed that Caprin still formed tight puncta indicating that G3BP1 and G3BP2 are not required for the formation of the Group 9 puncta **(Figure 4.4L)**. This raises the possibility that these puncta of group 9 proteins

form through a set of multiple interactions between these proteins and thereby can form independently of G3BP (see discussion).

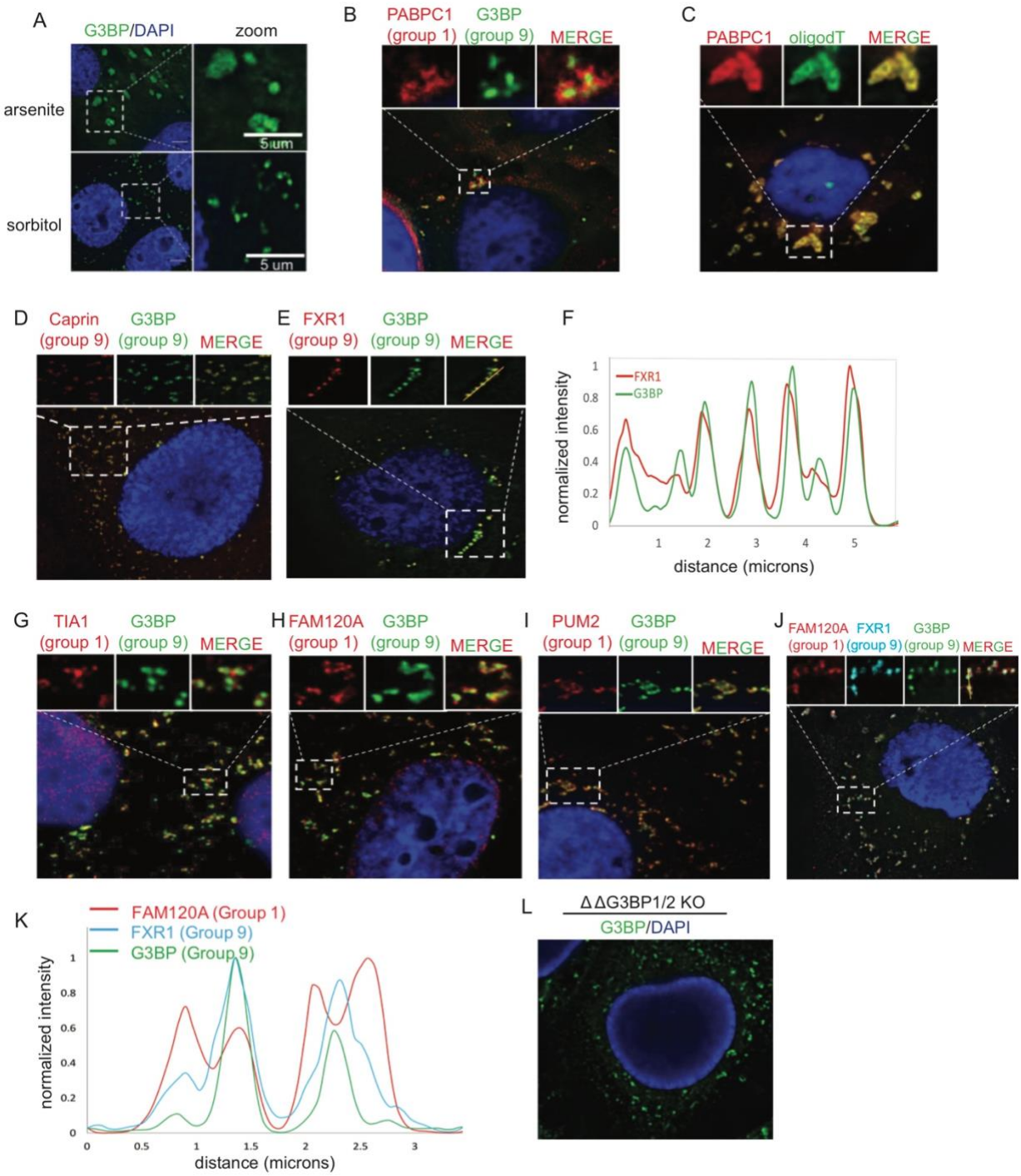


Figure 4.4: Hyperosmotic stress reveals specific substructure within stress granules

(A) Fluorescence microscopy imaging of arsenite and sorbitol induced SGs **(B**) Immunofluorescence microscopy of GFP-G3BP expressing U2OS cells stained for PABPC1 after 2.5 hours of sorbitol stress. **(C)** Immunofluorescence imaging of PABP and oligodT FISH after 2.5 hours of sorbitol stress. **(D and E)** Same as B, but stained for CAPRIN **(D)** and FXR1 **(E)**, interaction group 9 proteins. **(F)** Intensity plot profile for FXR1 and G3BP1 (see gold line in top right panel of E). **(G-I)** Same as B, but stained for TIA1 **(G)**, FAM120A **(H)**, and PUM2 **(I)**, interaction group 1 proteins. **(J)** Immunofluorescence microscopy of G3BP-GFP expressing U2OS cells stained for FAM120A and FXR1 after 2.5 hours of sorbitol stress. **(K**) Intensity plot of FAM120A, FXR1, and G3BP (see gold line in top right panel of G). **(L)** Immunofluorescence microscopy of Caprin in ΔΔG3BP1/2 cells.

Stress granule mRNAs localize adjacent to G3BP cores

We observed that SGs are composed of a heterogenous mixture of punctate cores during sorbitol stress. However, we also observed that purification of SGs by different RBPs yields a similar transcriptome. To reconcile these observations, we hypothesized that RNAs may act as a linker between protein cores, in which case individual RNAs might lie at the periphery of tight G3BP puncta observed in sorbitol stress. Thus, we examined where RNAs localize with respect to group 9 proteins during sorbitol stress.

We examined the localization of AHNAK transcript relative to Caprin cores during sorbitol-induced stress by smFISH. Interestingly, we find that the AHNAK transcript localizes to the periphery of Caprin SG cores **(Figure 4.5 A and B)**. Similarly, we find that CENPN transcripts localizes to the periphery of GFP-G3BP cores **(Figure 4.5C and D)**. Finally, we found that AHNAK localization to CAPRIN cores is unaffected in G3BPΔ/Δ cells by smFISH **(Figure 4.5 E and F)**.

Taken together these data support a model in which RNAs act as scaffolds for RNA binding protein cores and may serve as bridges between different RNA binding protein cores (see discussion).

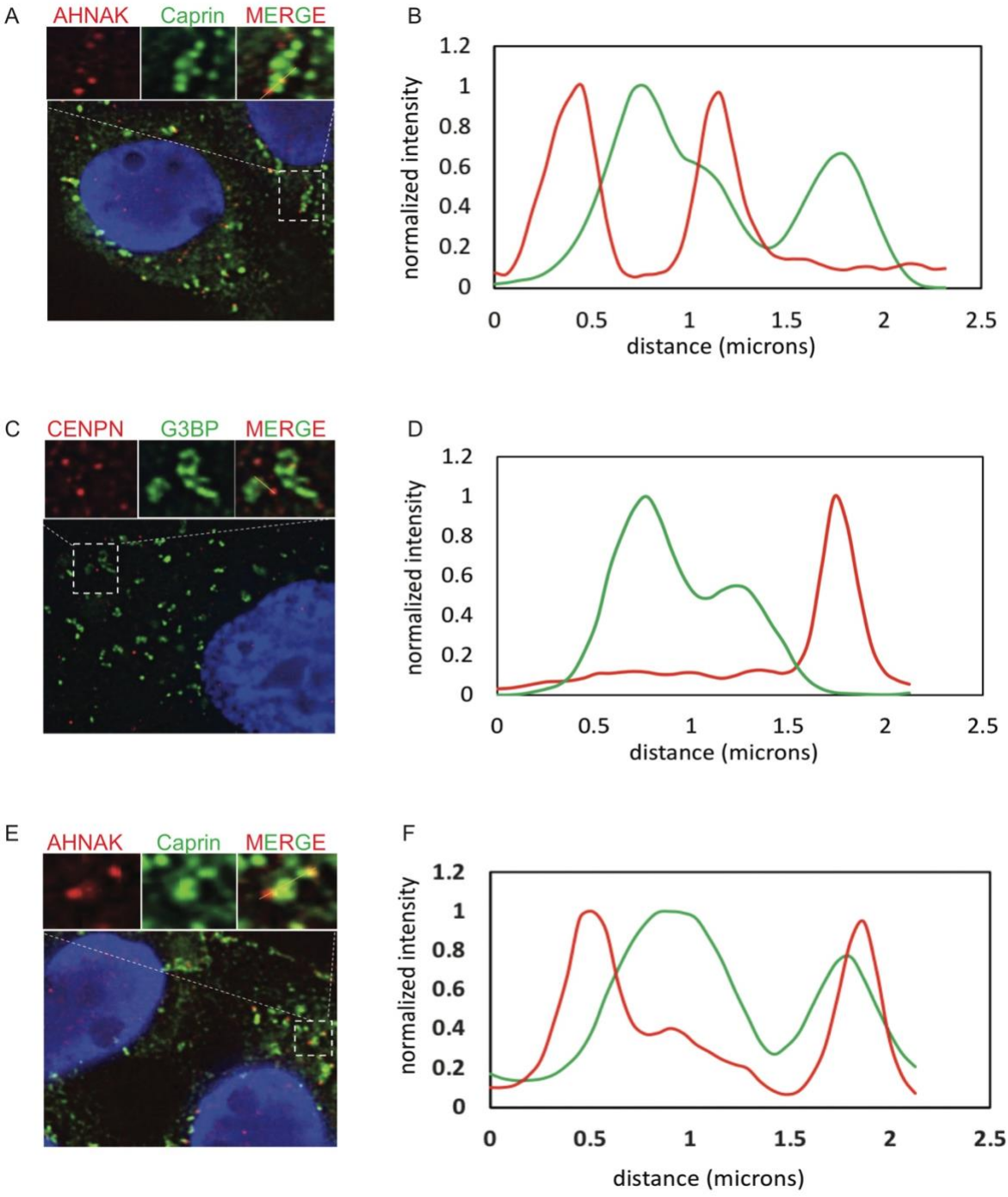

A
AHNAK
Caprin
MERGE
B
normalized intensity
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
2.5
distance (microns)
C
CENPN
G3BP
MERGE
D
normalized intensity
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
2.5
distance (microns)
E
AHNAK
Caprin
MERGE
F
normalized intensity
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
2.5
distance (microns)

Figure 4.5: RNAs localize to the periphery of stress granule cores during sorbitol stress

(A) smFISH of AHNAK and imaging of GFP-G3BP during sorbitol stress in WT cells. **(B)** Intensity plot for AHNAK and Caprin. **(C)** smFISH of CENPN in G3BP-GFP cells during sorbitol stress. **(D)** Intensity plot for CENPN and G3BP. **(E)** Same as A, but in G3BP1/2 knockout cells stained for Caprin. **(F)** Intensity plot for AHNAK and Caprin in G3BP1/2 knockout cells.

Transcripts that are upregulated during stress show lowered localization to granules

During conditions of stress, cells must alter their gene expression in order to initiate a stress response. We reasoned that stress response genes that are transcriptionally upregulated during stress are also likely to be translated in order to make functional proteins to initiate a stress response. Since actively translated genes are thought to escape sequestration into SGs, we hypothesized that genes that are upregulated during stress would be depleted from SGs. To test this hypothesis, we compared RNA-Seq libraries from arsenite-treated and untreated U2-OS cells, and examined the partitioning of induced RNAs into SGs.

As anticipated, gene ontology for upregulated transcripts during arsenite stress yielded numerous stress response terms such as ‘regulation of cell death’, ‘response to cytokine’, ‘cellular response to oxidative stress’, and ‘PERK-mediated unfolded protein response’ **(Figure 4.6A)**. Transcripts that were upregulated after arsenite stress tended to show lower levels of enrichment in SGs **(Figure 4.6B)**. We also compared stress induced gene expression with SG enrichment scores via scatterplot and observed a very weak negative correlation globally **(Figure 4.6C)**. While we only observe a very weak correlation globally, we do note that the most upregulated transcripts tend to be depleted from granules, while the most downregulated transcripts tend to be enriched in granules. Thus, we interpret these results to suggest that highly stress-induced

transcripts tend to have lower levels of accumulation into SGs.

We also performed a similar analysis to observe whether sorbitol-induced transcripts show reduced accumulation into SGs. Similar to arsenite treatment we find that during sorbitol stress, upregulated transcripts tend to have lower SG enrichment scores, than those that are downregulated **(Figure 4.6D)**. Consistent with this finding, we observe a weak global negative correlation between stress induced gene expression and SG enrichment **(Figure 4.6E)**. Finally, we also observe that arsenite and sorbitol treatment yield similar gene expression changes during stress, suggesting that these two differing stresses may generate a similar general stress response **(Figure 4.6F)**. Taken together, these data suggest that mRNAs induced during stress responses show reduced partitioning into SGs.

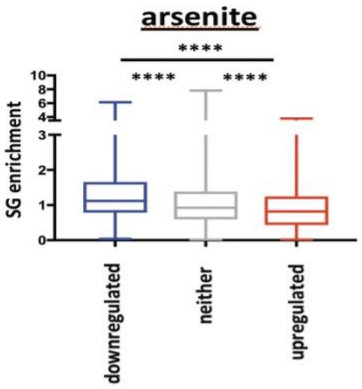
Since RNA length correlates with SG enrichment, the reduced SG partitioning of stress induced RNAs could simply be due to these transcripts being shorter in length. Indeed, we observed that length and stress-induced RNAs showed a negative correlation **(Figure 4.6G**), consistent with stress induced RNA being shorter on average**.** To determine if the reduced accumulation of stress induced RNAs was simply due to their length, we compared SG enrichment across length matched samples **(Figure 4.6H).** We observed that transcripts that are upregulated during stress consistently have lower SG enrichment scores than downregulated transcripts across a broad spectrum of lengths, suggesting that upregulated transcripts have additional mechanisms for evading sequestration into SGs, possibly due to increased translation.

Taken together, our results are consistent with the hypothesis that stress induced genes escape sequestration into SGs, presumably through a combination of their translation status and/or length.

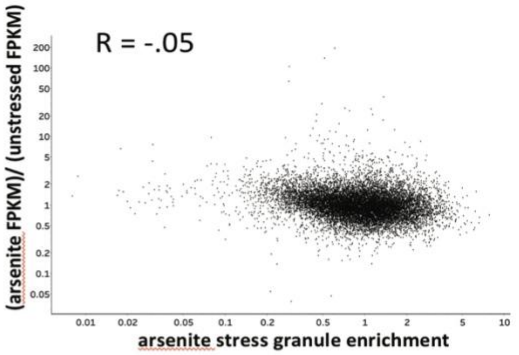
A

GO TERMS	p-value
regulation of cell death (GO:0010941)	1.20E-09
response to cytokine (GO:0034097)	4.76E-09
cellular response to oxidative stress (GO:0034599)	6.13E-03
PERK-mediated unfolded protein response (GO:0036499)	4.91E-03

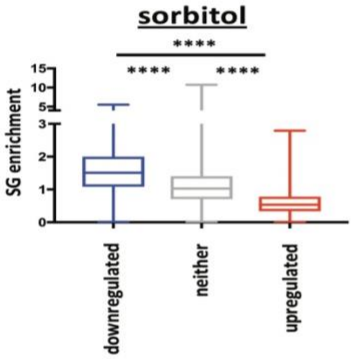
B



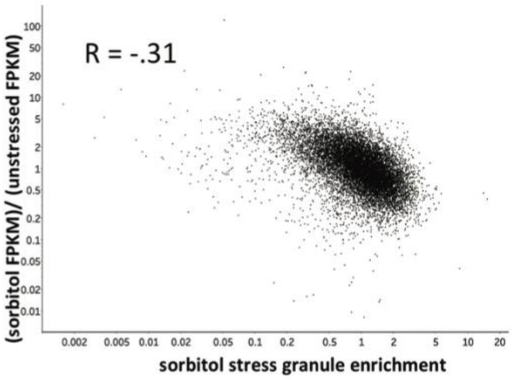
C



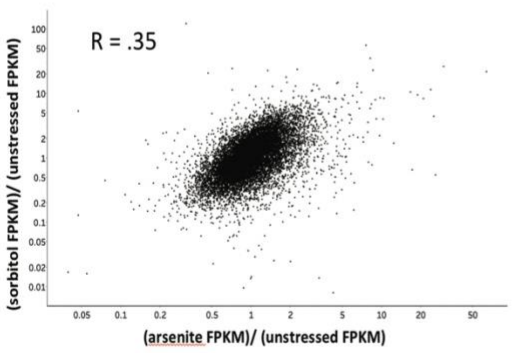
D



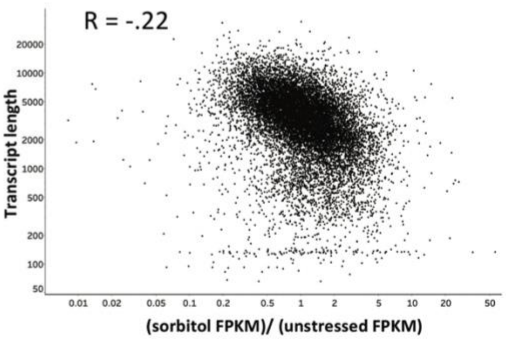
E



F



G



H

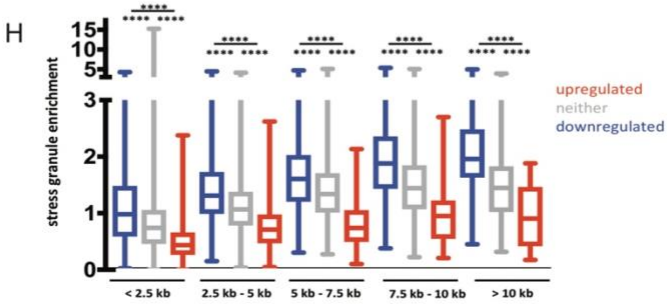


Figure 4.6: RNAs upregulated during stress are depleted from SGs

(A) Table of GO-terms for arsenite induced transcripts. **(B)** Boxplot showing SG enrichment scores for upregulated and downregulated RNAs after arsenite treatment. (**C)** Scatterplot showing arsenite induced transcripts vs. arsenite SG enrichment. **(D)** Boxplot showing SG enrichment scores for upregulated and downregulated RNAs after sorbitol treatment. **(E)** Scatterplot showing sorbitol induced transcripts vs. sorbitol SG enrichment. **(F)** Scatterplot showing correlation between sorbitol induced transcripts vs arsenite induced transcripts **(G)** Scatterplot showing correlation between length vs sorbitol induced genes. **(H)** Boxplot showing SG enrichment scores for upregulated and downregulated RNAs after sorbitol treatment, binned by length every 2.5 kilobases.

DISCUSSION

In this work, we present evidence that the RNAs accumulating within SGs are largely conserved between sorbitol and arsenite induced stress. The key observation is that RNAs enriched in SGs are similar in sorbitol or arsenite induced stress, and this is independent of whether the SG cores are purified by PABPC1 or G3BP affinity methods **(Figure 4.2)**. This finding is consistent with earlier smFISH experiments where five test mRNAs were shown to be recruited similarly to SGs in thapsigargin, heat shock, or sorbitol stress (Khong et al., 2017). Moreover, analysis of mRNA localization to RNA granules (based on pelleting heavy complexes) suggests that similar mRNAs are recruited during heat shock, ER stress, and arsenite induced stress (Namkoong et al., 2018). Thus, RNA localization to SGs is only modestly impacted by the type of stress.

The fact that the RNA composition of SGs is largely conserved between different stresses has two implications. First, since some proteins are only recruited to SGs under specific stresses (Buchan and Parker, 2009; Markmiller et al., 2018), it implies that the stress-specific recruitment of proteins will occur by the modulation of the ability of those proteins to bind RNA and/or other components of SGs, and not by the recruitment of specific RNAs. Second, this finding suggests that SGs are part of a generic stress response pathway in which long, poorly translated RNAs preferentially accumulate in SGs, which could be a general property of these RNAs to form multivalent RNA-RNA interactions and co-assemble into SGs (Van Treeck et al., 2018).

We also present evidence that RNA partitioning into SGs can be independent of at least some RNA binding proteins. The first line of evidence that suggests RNA localization to SGs is independent of proteins is that PABPC1 SG cores contain a similar RNA composition to G3BP SG cores. If RBPs played a highly specialized role in recruiting distinct RNA substrates to SGs, one would anticipate that pull-down of SG cores via different RNA binding proteins would yield different transcriptomes. More strikingly, we observed that SGs isolated from G3BP1/2 knockout cells during sorbitol induced stress share the same RNA composition as SGs isolated from WT cells **(Figure 4.3)**. Consistent with RBPs playing a limited role in SG partitioning of RNAs, mRNAs enriched in SG are not enriched for binding sites of RBPs in SG (**Figure 4.3D**; Khong et al., 2017), and the recruitment of the NORAD RNA, which has at least 17 Pumilio binding sites, to SGs has been shown to be independent of the Pumilio RBP (Namkoong et al., 2018,Lee et al., 2016).

The observation that G3BP, an abundant RBP in SGs, does not affect the partitioning of RNA into SGs can be interpreted in two general manners. First, it could be that other RBPs, or a combination of RBPs, determine what RNAs are targeted to SGs, although this model is inconsistent with the limited enrichment of binding sites for SG components in RNAs recruited to SGs (Khong et al., 2017). Second, these results are consistent with a model wherein intermolecular RNA-RNA interactions dominate the assembly of SGs and are simply more prevalent between longer RNAs. This latter model is also supported by the observation that the yeast SG transcriptome can be

largely recapitulated by the *in vitro* assembly of RNA in the absence of any proteins (Van Treeck et al., 2018). In this model, proteins can be recruited to stress granules by interactions with the resident RNAs and, in cases of multivalent RNA binding proteins, contribute to cross linking between individual RNPs and thereby increase stress granule assembly.

Several observations suggest that during sorbitol stress SGs have a defined organization with the following features. First, a set of interacting proteins, Group 9 proteins (Youn et al., 2018), form discrete puncta within a larger SG **(Figure 4.4)**. Proteins in a second interaction group such as PABPC1 and FAM120 are distributed outside these puncta **(Figure 4.4)**. Since poly(A)+ FISH signal and smFISH for specific mRNAs shows these mRNAs are generally outside of the Group 9 protein puncta **(Figure 4.4 & 5)**, we hypothesize that the puncta are formed from Group 9 proteins on different mRNAs interacting with each other with their associated mRNAs on the periphery of the puncta **(Figure 4.7)**. While this organization is based on our observations in sorbitol stress, we speculate that a similar organizational principle occurs in other stress conditions.

This organizational pattern allows us to understand how PABPC1 and other RBPs can form distinct substructures from G3BP during sorbitol stress, yet pulldown of PABPC1 SG cores yields a similar transcriptome to G3BP cores. We propose a model in which RNAs organize this substructure **(Figure 4.7).** Given its size, an individual RNA molecule likely has RBP sites for many different RBPs. Since many RBPs are

capable of forming homotypic interactions through dimerization and other interactions, it is reasonable to postulate that these regions may come into close proximity, due to protein-protein interactions forming between individual RBPs and the ability of RNA to fold in 3-dimensional space. In our model, interaction Group 9 proteins may be able to form at a given location on a RNA molecule, while at another point on the same RNA molecule interaction Group 1 proteins may be able to form distinct clusters. Each of these interaction group clusters could then go on to form interactions with protein clusters on other RNA molecules, leading to the SG substructure we observe during sorbitol stress. Indeed, this model is consistent with the observations that we see a similar RNA composition of PABPC1 cores and G3BP cores, but also observe differences PABPC1 and G3BP localization.

Finally, we present evidence that stress-induced RNAs tend to avoid sequestration into SGs in both sorbitol and arsenite induces stress. Moreover, arsenite and sorbitol stress have similar gene expression changes **(Figure 4.6)**, suggesting that a general stress response pathway is being activated in response to both of these stresses. While short transcripts tend to be produced during stress response **(Figure 4.6)**, we find that length alone does not explain the reduced SG enrichment of stress-induced genes **(Figure 4.6)**. Indeed, if we compare length-matched transcripts, stress-induced transcripts still show lowered partitioning to SGs, suggesting that there are additional mechanisms that may act to lower their partitioning to SGs. This may, in part, be explained by a model in which stress induced transcripts are also translationally upregulated during stress. However, it remains a possibility that these transcripts are

modified in some unknown way during stress and that this unknown modification may aid in helping these transcripts evade sequestration into SGs.

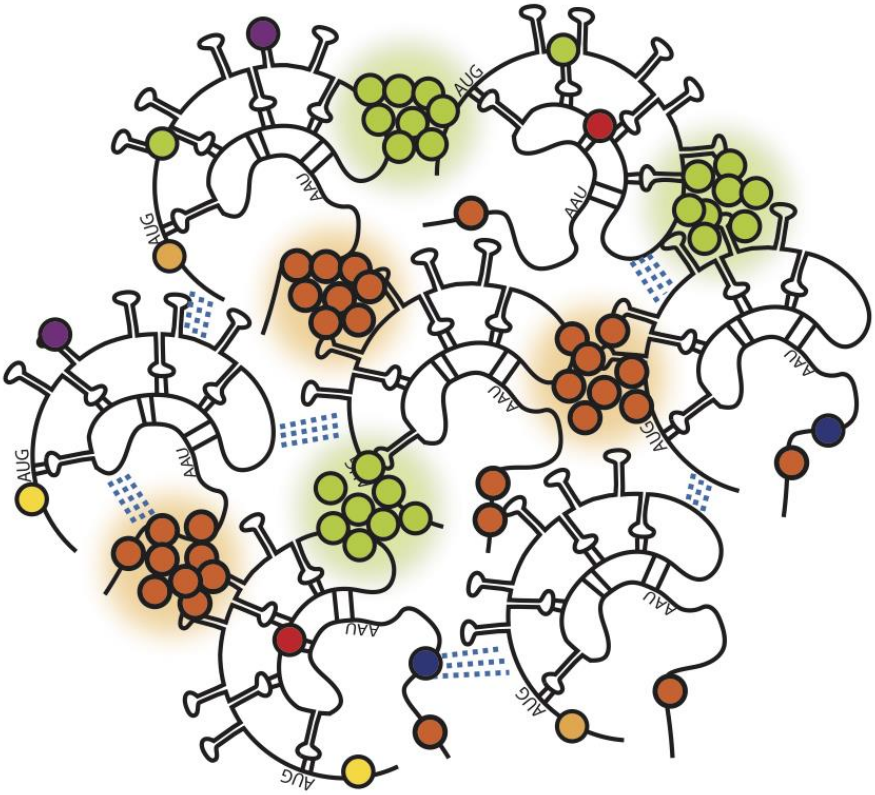


Figure 4.7: Model for RNA and protein arrangement in SGs

Green circles represent interaction group 9 proteins, while red circles represent interaction group 1 proteins (other colored circles represent other RBPs). In our model, RNA-RNA interactions and protein-protein interactions act synergistically to form SG architecture. Initially RNA would bring various RBPs into close proximity. Protein cores (red and green halos) would then self-assemble from intra and inter molecular interactions. In this model, RNA may act as a bridge between protein cores.

METHODS

U-2 OS growth conditions and reagents

Human osteosarcoma U-2 OS cells expressing G3BP1-GFP (Paul Taylor Lab), U-2 OS cells and U-2 OS ΔΔG3BP1/2 (Kedersha et al., 2016) were used in all experiments. All cells were maintained in DMEM with high glucose, 10% fetal bovine serum, and 1% penicillin/streptomycin at 37°C/5% CO_2, were used in all experiments.

Isolation of RNA from U-2 OS cells and SG cores for RNA-sequencing

Parental and ΔΔG3BP1/2 U-2 OS cells expressing G3BP1-GFP were grown to 85% confluency in three 500 cm^2 square TC-treated culture dishes (Thermo Fisher Scientific, 07-200-599). One hour prior to stress, cell culture media was exchanged with fresh media. Cells were then treated stressed with either NaAsO2 or sorbitol (see 'Stress Conditions'). After stress, cells were washed once with media, transferred to falcon tubes, and pelleted at 1,500 x g for 3 mins. Upon aspirating the media, the pellets were immediately flash-frozen in liquid N_2 and stored at -80°C until isolation of mammalian SG cores was performed. PABPC1 SG cores were purified as previously described for the purification of G3BP cores, however instead of using anti-GFP for the pulldown of SG cores, 20 uL of anti-PABPC1 was used (Khong et al., 2017, 2018). For all experiments, biological triplicates were acquired for both SG-purified RNA and total RNA except in the case of PABPC1 SG RNA purified during sorbitol stress, in which duplicates were acquired (One of the replicates was excluded because it showed little similarity to the other biological replicates).

Stress conditions

To examine mRNA localization during stress, we used the following stress conditions. For arsenite stress experiments, cells were treated with .5mM sodium arsenite (Sigma-Aldrich S7400) for 1 hour. For osmotic stress, cells were stressed in .5M D-sorbitol (Sigma-Aldrich S1876) for 2.5 hours. Cells were fixed after the completion of each stress with 4% paraformaldehyde (Fisher Scientific NC0179595).

Immunofluorescence and smFISH

Immunofluorescence and smFISH experiments were performed utilizing the Stellaris manufacturers protocol. Specifically, the sequential immunofluorescence and smFISH protocol was used during experiments in which both IF and smFISH was performed. In order to maintain consistency, the same protocol was utilized in IF only experiments, however the portions of the protocol calling for smFISH were omitted. Antibodies that were used include TIA1 (Abcam ab40693), PUM2 (Abcam ab10361), PABP (Abcam ab21060), G3BP (Abcam ab56574), FXR1 (Abcam ab51970), FAM120A (Sigma-Aldrich HPA019734) and CAPRIN (50-554-357). In all imaging experiments at least 20 cells were imaged and biological triplicates were performed.

smFISH probes were labeled based on a recent protocol (Gaspar et al., 2017). Olionucleotides used for smFISH were designed using the Stellaris Probe Designer (https://www.biosearchtech.com/support/education/stellaris-rna-fish). Briefly, oligos were labeled by incubating adding 4 uL of 200 uM pooled oligonucleotides (Integrated DNA Technologies), 1 uL of TdT (Thermo Fischer Scientific EP0161) and 6 uL of ddUTP

(Axxora JBS-NU-1619-633**)** fluorophore for eight hours. After 8 hours, another 1 uL of TdT enzyme was added and the reaction was allowed to continue overnight. Probes were than ethanol precipitated by adding 164.5 uL of nuclease free water, .5 uL of .5 mg/ml linear acrylamide, 20 uL of 3M sodium acetate (pH 5.5) and 800 uL of pre-chilled 100% ethanol. This mixture was then placed at -80 C for at least 20 minutes. Labeled oligos were then pelleted by centrifugation at 16,000 x g at 4 C. Probes were then washed 3 times by adding 1 mL 80% ethanol and centrifugation at 16,000 x g at 4 C. After washing, the pellet was allowed to air dry for 10 minutes and probes were resuspended in 25uL of nuclease free water.

Library construction and RNA-sequencing

RNA quality was assessed by TapeStation analysis at the Biofrontiers Institute Sequencing Facility. Paired-end cDNA libraries were prepared at the Biofrontiers Institute Sequencing Facility using the KAPA HyperPrep with RiboErase. cDNA libraries were sequenced on a NextSeq High output 150 cycle (2x75).

Sequencing data analysis

Read quality was assessed using fastqc. Illumina adaptors were trimmed using Trimmomatic 0.32 in paired and (PE) mode (Bolger et al., 2014). An index genome was acquired from GENCODE (Release 19 GRCh37.p13). Reads were aligned using Tophat (version 2.0.6) and Bowtie2 (version 2.0.2) and the following parameters:–b2 –fast –microexon-search (Kim et al., 2013a; Langmead and Salzberg, 2012). Differential expression analysis was performed using Cuffdiff (version 2.2.1) with the default

parameters (Trapnell et al., 2013). Gene Ontology analysis was performed using the gene ontology consortium (http://www.geneontology.org/). Transcript lengths were acquired from Ensembl's Biomart Tool. All sequencing data can be found at NCBI GEO GSE119977.

Statistical Analyses and Plotting

Statistics for overlap in Venn Diagrams were performed using a two population proportion test in R. All p-values for boxplots were calculated using an ordinary one-way ANOVA test in Prism version 7.0d. All boxbplots were created in Prism version 7.0d. All scatterplots were generated using Tableau version 10.0. Intensity Plot Profiles were generated using ImageJ version 2.0.0.

REFERENCES

CHAPTER 1 REFERENCES

Ainger, K., Avossa, D., Morgan, F., Hill, S.J., Barry, C., Barbarese, E., and Carson, J.H. (1993). Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. J. Cell Biol. *123*, 431–441.

Anderson, P., and Kedersha, N. (2002). Stressful initiations. J. Cell Sci. *115*, 3227–3234.

Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M., and Heyer, W.D. (1997). A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J. Cell Biol. *136*, 761–773.

Bond, C.S., and Fox, A.H. (2009). Paraspeckles: nuclear bodies built on long noncoding RNA. J. Cell Biol. *186*, 637–644.

Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. Science (80-.). *310*, 486–489.

Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol. Cell *36*, 932–941.

Buchan, J.R., Muhlrad, D., and Parker, R. (2008). P bodies promote stress granule assembly in Saccharomyces cerevisiae. J. Cell Biol. *183*, 441–455.

Courel, M., Clement, Y., Foretek, D., Vidal, O., Yi, Z., Kress, M., Vindry, C., Benard, M., Bossevain, C., Antoniewski, C., et al. (2018). GC content shapes mRNA decay and storage in human cells. bioRxiv 373498.

Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb. Perspect. Biol. *4*, 1–16.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. J. Cell Biol. *179*, 437–449.

Didiot, M.-C., Subramanian, M., Flatter, E., Mandel, J.-L., and Moine, H. (2009). Cells Lacking the Fragile X Mental Retardation Protein (FMRP) have Normal RISC Activity but Exhibit Altered Stress Granule Assembly. Mol. Biol. Cell *20*, 428–437.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol. Cell. Biol. *27*, 3970–3981.

El Fatimy, R., Davidovic, L., Tremblay, S., Jaglin, X., Dury, A., Robert, C., De Koninck, P., and Khandjian, E.W. (2016). Tracking the Fragile X Mental Retardation Protein in a Highly Ordered Neuronal RiboNucleoParticles Population: A Link between Stalled Polyribosomes and RNA Granules. PLoS Genet. *12*, e1006192.

Fox, A.H., and Lamond, A.I. (2010). Paraspeckles. Cold Spring Harb. Perspect. Biol. *2*, a000687.

Fox, A.H., Nakagawa, S., Hirose, T., and Bond, C.S. (2018). Paraspeckles: Where Long Noncoding RNA Meets Phase Separation. Trends Biochem. Sci. *43*, 124–135.

Galganski, L., Urbanek, M.O., and Krzyzosiak, W.J. (2017). Nuclear speckles: molecular organization, biological function and role in disease. Nucleic Acids Res. *45*, 10350–10368.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress Granule Assembly Is Mediated by Prion-like Aggregation of TIA-1. Mol. Biol. Cell *15*, 5383–5398.

Hanson, G., and Coller, J. (2017). Codon optimality, bias and usage in translation and mRNA decay. Nat. Rev. Mol. Cell Biol. *19*, 20–30.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib,

R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol. Cell *68*, 144–157.e5.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

Jiang, L., Shao, C., Wu, Q.-J., Chen, G., Zhou, J., Yang, B., Li, H., Gou, L.-T., Zhang, Y., Wang, Y., et al. (2017). NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat. Struct. Mol. Biol. *24*, 816–824.

Kedersha, N., Cho, M.R., Li, W., Yacono, P.W., Chen, S., Gilks, N., Golan, D.E., and Anderson, P. (2000). Dynamic Shuttling of TIA-1 Accompanies the Recruitment of mRNA to Mammalian Stress Granules 7. *151*, 1257–1268.

Kedersha, N., Chen, S., Gilks, N., Li, W., Miller, I.J., Stahl, J., and Anderson, P. (2002). Evidence That Ternary Complex (eIF2-GTP-tRNA iMet)– Deficient Preinitiation Complexes Are Core Constituents of Mammalian Stress Granules. *13*, 195–210.

Kedersha, N., Panas, M.D., Achorn, C.A., Lyons, S., Tisdale, S., Hickman, T., Thomas, M., Lieberman, J., McInerney, G.M., Ivanov, P., et al. (2016). G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. J.

Cell Biol. *212*, 845–860.

Kedersha, N.L., Gupta, M., Li, W., Miller, I., and 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*, 1431–1442.

Khong, A., and Parker, R. (2018). mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. J. Cell Biol. jcb.201806183.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., Parker, R., Khong, A., Matheny, T., Jain, S., Mitchell, S.F., et al. (2017). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules Resource The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. Mol. Cell *68*, 808–820.e5.

Kiebler, M.A., and Bassell, G.J. (2006). Neuronal RNA Granules: Movers and Makers. Neuron *51*, 685–690.

Kimball, S.R., Horetsky, R.L., Ron, D., Jefferson, L.S., and Harding, H.P. (2003). Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. Am. J. Physiol. Cell Physiol. *284*, C273-84.

Lam, Y.W., and Trinkle-Mulcahy, L. (2015). New insights into nucleolar structure and function. F1000Prime Rep. *7*, 48.

Liu-Yesucevitz, L., Bassell, G.J., Gitler, A.D., Hart, A.C., Klann, E., Richter, J.D., Warren, S.T., and Wolozin, B. (2011). Local RNA translation at the synapse and in disease. J. Neurosci. *31*, 16086–16093.

McInerney, G.M., Kedersha, N.L., Kaufman, R.J., Anderson, P., and Liljeström, P. (2005). Importance of eIF2α Phosphorylation and Stress Granule Assembly in Alphavirus Translation Regulation. Mol. Biol. Cell *16*, 3753–3763.

Mollet, S. (2008). Translationally Repressed mRNA Transiently Cycles through Stress Granules during Stress. Mol. Biol. Cell *19*, 4469–4479.

Moon, S.L., Morisaki, T., Khong, A., Lyon, K., Parker, R., and Stasevich, T.J. (2018). Imaging of single mRNA translation repression reveals diverse interactions with mRNP granules. bioRxiv 332692.

Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., Lee, J.H., Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., and Lee, J.H. (2018). Granulation Resource Systematic Characterization of Stress-Induced RNA Granulation. Mol. Cell *70*, 175–187.e8.

Niewidok, B., Igaev, M., Pereira, A., Strassner, A., Lenzen, C., Richter, C.P., and

Piehler, J. (2018). Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. *217*, 1303–1318.

Nover, L., Scharf, K.D., and Neumann, D. (1983). Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. Mol. Cell. Biol. *3*, 1648–1655.

Nover, L., Scharf, K.D., and 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*, 1298–1308.

Parker, R., and Sheth, U. (2007). P Bodies and the Control of mRNA Translation and Degradation. Mol. Cell *25*, 635–646.

Savino, T.M., Gébrane-Younès, J., De Mey, J., Sibarita, J.B., and Hernandez-Verdun, D. (2001). Nucleolar assembly of the rRNA processing machinery in living cells. J. Cell Biol. *153*, 1097–1110.

Sheth, U., and Parker, R. (2003). Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. Science (80-.). *300*, 805–808.

Sidrauski, C., Mcgeachy, A.M., Ingolia, N.T., and Walter, P. (2015). The small molecule ISRIB reverses the effects of eIF2 α phosphorylation on translation and stress granule assembly. 1–16.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly (A) -tail profiling reveals an embryonic switch in translational control. Nature *508*, 66–71.

Taliaferro, J.M., Vidaki, M., Oliveira, R., Olson, S., Zhan, L., Saxena, T., Wang, E.T., Graveley, B.R., Gertler, F.B., Swanson, M.S., et al. (2016). Distal Alternative Last Exons Localize mRNAs to Neural Projections. Mol. Cell *61*, 821–833.

Teixiera, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA *11*, 371–382.

Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J. (2003). The RasGAP-associated endoribonuclease G3BP assembles stress granules. J. Cell Biol. *160*, 823 LP-831.

Wheeler, J.R., Matheny, T., Jain, S., Abrisch, R., and Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. Elife *5*, 1–25.

Woo, Y.M., Kwak, Y., Namkoong, S., Kristjánsdóttir, K., Lee, S.H., Lee, J.H., and Kwak, H. (2018). TED-Seq Identifies the Dynamics of Poly(A) Length during ER Stress. Cell Rep. *24*, 3630–3641.e7.

Yang, Y.-C.T., Di, C., Hu, B., Zhou, M., Liu, Y., Song, N., Li, Y., Umetsu, J., and Lu, Z.J. (2015). CLIPdb: a CLIP-seq database for protein-RNA interactions. BMC Genomics *16*, 51.

CHAPTER 2 REFERENCES

Brangwynne, C.P. (2013). Phase transitions and size scaling of membrane-less organelles. J. Cell Biol. *203*, 875–881.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., and Hyman, A.A. (2009). Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation. Science *324*, 1729–1732.

Brangwynne, C.P., Tompa, P., and Pappu, R.V. (2015). Polymer physics of intracellular phase transitions. Nat. Phys. *11*, 899–904.

Buchan, J.R. (2014). mRNP granules. Assembly, function, and connections with disease. RNA Biol. *11*, 1019–1030.

Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol. Cell *36*, 932–941.

Buchan, J.R., Kolaitis, R.-M., Taylor, J.P., and Parker, R. (2013). Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. Cell *153*, 1461–1474.

Chernov, K.G., Barbet, A., Hamon, L., Ovchinnikov, L.P., Curmi, P.A., and Pastré, D. (2009). Role of Microtubules in Stress Granule Assembly J. Biol. Chem. *284*, 36569–

36580.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. J. Cell Biol. *179*, 437–449.

Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M.E., Mackenzie, I.R.A., Capell, A., Schmid, B., et al. (2010). ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. EMBO J. *29*, 2841–2857.

Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C.-H., Eckmann, C.R., Myong, S., and Brangwynne, C.P. (2015). The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc. Natl. Acad. Sci. *112*, 7189–7194.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. Cell *165*, 1686–1697.

Frey, S., and Görlich, D. (2007). A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. Cell *130*, 512–523.

Fujimura, K., Katahira, J., Kano, F., Yoneda, Y., and Murata, M. (2009). Microscopic dissection of the process of stress granule assembly. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1793*, 1728–1737.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L.M., and Anderson, P. (2004). Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol. Biol. Cell *15*, 5383–5398.

Han, T.W., Kato, M., Xie, S., Wu, L.C., Mirzaei, H., Pei, J., Chen, M., Xie, Y., Allen, J., Xiao, G., et al. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. Cell *149*, 768–779.

Hülsmann, B.B., Labokha, A.A., and Görlich, D. (2012). The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. Cell *150*, 738–751.

Hyman, A.A., and Simons, K. (2012). Cell biology. Beyond oil and water--phase transitions in cells. Science *337*, 1047–1049.

Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-Liquid Phase Separation in Biology. Annu. Rev. Cell Dev. Biol. *30*, 39–58.

Ivanov, P.A., Chudinova, E.M., and Nadezhdina, E.S. (2003). Disruption of microtubules

inhibits cytoplasmic ribonucleoprotein stress granule formation. Exp. Cell Res. *290*, 227–233.

Izumi, R., Warita, H., Niihori, T., Takahashi, T., Tateyama, M., Suzuki, N., Nishiyama, A., Shirota, M., Funayama, R., Nakayama, K., et al. (2015). Isolated inclusion body myopathy caused by a multisystem proteinopathy-linked hnRNPA1 mutation. Neurol. Genet. *1*, e23.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

Kato, M., Han, T.W., Xie, S., Shi, K., Du, X., Wu, L.C., Mirzaei, H., Goldsmith, E.J., Longgood, J., Pei, J., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. Cell *149*, 753–767.

Kedersha, N., Tisdale, S., Hickman, T., and Anderson, P. (2008). Real-time and quantitative imaging of mammalian stress granules and processing bodies. Methods Enzymol. *448*, 521–552.

Kedersha, N., Ivanov, P., and Anderson, P. (2013). Stress granules and cell signaling: more than just a passing phase? Trends Biochem. Sci. *38*, 494–506.

Kim, H.J., Kim, N.C., Wang, Y.-D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S.,
Molliex, A., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature *495*, 467–473.

King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. *1462*, 61–80.

Kroschwald, S., Maharana, S., Mateju, D., Malinovska, L., Nüske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. eLife *4*, e06807.

Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L., and McKnight, S.L. (2013). Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. Cell *155*, 1049–1060.

van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., et al. (2014). Classification of Intrinsically Disordered Regions and Proteins. Chem. Rev. *114*, 6589–6631.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J. Cell Biol. *201*, 361–372.

Lin, Y., Protter, D.S.W., Rosen, M.K., and Parker, R. (2015). Formation and Maturation

of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol. Cell *60*, 208–219.

Loschi, M., Leishman, C.C., Berardone, N., and Boccaccio, G.L. (2009). Dynein and kinesin regulate stress-granule and P-body dynamics. J. Cell Sci. *122*, 3973–3982.

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell *163*, 123–133.

Murakami, T., Qamar, S., Lin, J.Q., Schierle, G.S.K., Rees, E., Miyashita, A., Costa, A.R., Dodd, R.B., Chan, F.T.S., Michel, C.H., 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*, 678–690.

Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., et al. (2015). Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. Mol. Cell *57*, 936–947.

Ohn, T., Kedersha, N., Hickman, T., Tisdale, S., and Anderson, P. (2008). A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. Nat. Cell Biol. *10*, 1224–1231.

Pak, C.W., Kosno, M., Holehouse, A.S., Padrick, S.B., Mittal, A., Ali, R., Yunus, A.A., Liu, D.R., Pappu, R.V., and Rosen, M.K. (2016). Sequence Determinants of Intracellular Phase Separation by Complex Coacervation of a Disordered Protein. Mol. Cell *63*, 72–85.

Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M., et al. (2015). A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell *162*, 1066–1077.

Patel, S.S., Belmont, B.J., Sante, J.M., and Rexach, M.F. (2007). Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. Cell *129*, 83–96.

Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell *154*, 727–736.

Reijns, M.A.M., Alexander, R.D., Spiller, M.P., and Beggs, J.D. (2008). A role for Q/N-rich aggregation-prone regions in P-body localization. J. Cell Sci. *121*, 2463–2472.

Ribbeck, K., and Görlich, D. (2002). The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. EMBO J. *21*, 2664–2671.

Schmidt, H.B., and Görlich, D. (2016). Transport Selectivity of Nuclear Pores, Phase

Separation, and Membraneless Organelles. Trends Biochem. Sci. *41*, 46–61.

Toretsky, J.A., and Wright, P.E. (2014). Assemblages: Functional units formed by cellular phase separation. J. Cell Biol. *206*, 579–588.

Updike, D.L., Hachey, S.J., Kreher, J., and Strome, S. (2011). P granules extend the nuclear pore complex environment in the C. elegans germ line. J. Cell Biol. *192*, 939–948.

Walters, R.W., Muhlrad, D., Garcia, J., and Parker, R. (2015). Differential effects of Ydj1 and Sis1 on Hsp70-mediated clearance of stress granules in Saccharomyces cerevisiae. RNA N. Y. N *21*, 1660–1671.

Weber, S.C., and Brangwynne, C.P. (2012). Getting RNA and Protein in Phase. Cell *149*, 1188–1191.

Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.

Zhang, H., Elbaum-Garfinkle, S., Langdon, E.M., Taylor, N., Occhipinti, P., Bridges, A.A., Brangwynne, C.P., and Gladfelter, A.S. (2015). RNA Controls PolyQ Protein Phase Transitions. Mol. Cell *60*, 220–230.

CHAPTER 3 REFERENCES

Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Banet, J.F., Hourlier, T., Billis, K., Garc, C., et al. (2016). Database update The Ensembl gene annotation system. 1–19.

Anders, S., Huber, W., S, A., and W, H. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Anders, S., Pyl, P.T., and Huber, W. (2015). Genome analysis HTSeq — a Python framework to work with high-throughput sequencing data. *31*, 166–169.

Andreev, D.E., Connor, P.B.F.O., Fahey, C., Kenny, E.M., Terenin, I.M., Dmitriev, S.E., Cormican, P., Morris, D.W., Shatsky, I.N., and Baranov, P. V (2015). Translation of 5 ′ leaders is pervasive in genes resistant to eIF2 repression. *4*, 1–21.

Bley, N., Lederer, M., Pfalz, B., Reinke, C., Fuchs, T., Glaß, M., Birgit, M., and Stefan, H. (2015). Stress granules are dispensable for mRNA stabilization during cellular stress. *43*.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Genome analysis Trimmomatic : a flexible trimmer for Illumina sequence data. *30*, 2114–2120.

Buchan, J.R., Yoon, J., and Parker, R. (2011). Stress-specific composition , assembly and kinetics of stress granules in Saccharomyces cerevisiae. 228–239.

Chartron, J.W., Hunt, K.C.L., and Frydman, J. (2016). Cotranslational signal-independent SRP preloading during membrane targeting. Nature *536*, 224–228.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE : The reference human genome annotation for The ENCODE Project. 1760–1774.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

Kedersha, N., Panas, M.D., Achorn, C.A., Lyons, S., Tisdale, S., Hickman, T., Thomas, M., Lieberman, J., McInerney, G.M., Ivanov, P., et al. (2016). G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. J. Cell Biol. *212*, 845–860.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2 : accurate alignment of transcriptomes in the presence of insertions , deletions and gene fusions. 1–13.

Kim, W.J., Kim, J.H., and Jang, S.K. (2007). Anti-inflammatory lipid mediator 15d-PGJ2 inhibits translation through inactivation of eIF4A. *26*, 5020–5032.

Kinsella, R.J., Ka, A., Spudich, G., Almeida-king, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., and Flicek, P. (2011). Original article Ensembl BioMarts : a hub for data retrieval across taxonomic space. *2011*, 1–9.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, 1–10.

Lee, S., Kopp, F., Chang, T., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell *164*, 69–80.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J. Cell Biol. *201*, 361–372.

McEwen, E., Kedersha, N., Song, B., Scheuner, D., Gilks, N., Han, A., Chen, J.-J., Anderson, P., and Kaufman, R.J. (2005). Heme-regulated Inhibitor Kinase-mediated Phosphorylation of Eukaryotic Translation Initiation Factor 2 Inhibits Translation, Induces Stress Granule Formation, and Mediates Survival upon Arsenite Exposure. J. Biol. Chem. *280*, 16925–16933.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science (80-.). *320*.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Meth *13*, 508–514.

Panas, M.D., Ivanov, P., and Anderson, P. (2016). Mechanistic insights into mammalian stress granule dynamics. *215*, 313–323.

Presnyak, V., Alhusaini, N., Chen, Y., Martin, S., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Brenton, R., and Coller, J. (2016). Codon optimality is a major determinant of mRNA stability. *160*, 1111–1124.

Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. Trends Cell Biol. *26*, 668–679.

Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell *154*, 727–736.

Siwiak, M., and Zielenkiewicz, P. (2010). A comprehensive, quantitative, and genome-

wide model of translation. PLoS Comput. Biol. *6*, 4.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotech *31*, 46–53.

Unsworth, H., Raguz, S., Edwards, H.J., Higgins, C.F., and Yagüe, E. (2010). mRNA escape from stress granule sequestration is dictated by localization to the endoplasmic reticulum. FASEB J. *24*, 3370–3380.

Wheeler, J.R., Matheny, T., Jain, S., Abrisch, R., and Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. Elife *5*, 1–25.

CHAPTER 4 REFERENCES

Anderson, P., Kedersha, N., and Ivanov, P. (2015). Stress granules , P-bodies and cancer. BBA Gene Regul. Mech. *1849*, 861–870.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Genome analysis Trimmomatic : a flexible trimmer for Illumina sequence data. Bioinformatics. *30*, 2114–2120.

Buchan, J.R., and Parker, R. (2009). Eukaryotic Stress Granules: The Ins and Outs of Translation. Mol. Cell *36*, 932–941.

Dewey, C.M., Cenik, B., Sephton, C.F., Dries, D.R., Iii, P.M., Good, S.K., Johnson, B.A.,

Herz, J., and Yu, G. (2011). TDP-43 Is Directed to Stress Granules by Sorbitol , a Novel Physiological Osmotic and Oxidative Stressor. Mol Cell Biol. *31*, 1098–1108.

Edupuganti, R.R., Geiger, S., Lindeboom, R.G.H., Shi, H., Hsu, P.J., Lu, Z., Wang, S., Baltissen, M.P.A., Jansen, P.W.T.C., Rossa, M., et al. (2017). N6 -methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol. *24*, 870-878

Gaspar, I., Wippich, F., and Ephrussi, A. (2017). Enzymatic production of single-molecule FISH and RNA capture probes. RNA. 1582–1591.

Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. Cell *164*, 487–498.

Kedersha, N., Ivanov, P., and Anderson, P. (2013). Stress granules and cell signaling : more than just a passing phase? Trends Biochem. Sci. *38*, 494–506.

Kedersha, N., Panas, M.D., Achorn, C.A., Lyons, S., Tisdale, S., Hickman, T., Thomas, M., Lieberman, J., McInerney, G.M., Ivanov, P., et al. (2016). G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. J. Cell Biol. *212*, 845–860.

Khong, A., Jain, S., Matheny, T., Wheeler, J.R., and Parker, R. (2017a). Isolation of mammalian stress granule cores for RNA-Seq analysis. Methods 1–6.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., Parker, R. (2017b). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules Resource The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. Mol Cell. *68*, 808–820.e5.

Khong, A., Jain, S., Matheny, T., Wheeler, J.R., and Parker, R. (2018). Isolation of mammalian stress granule cores for RNA-Seq analysis. Methods. *137*, 49–54.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013a). TopHat2 : accurate alignment of transcriptomes in the presence of insertions , deletions and gene fusions. 1–13.

Kim, H.J., Kim, N.C., Wang, Y., Scarborough, E.A., Moore, J., Diaz, Z., Maclea, K.S., Freibaum, B., Li, S., Molliex, A., et al. (2013b). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature *495*, 467–473.

Kim, N.C., Tresse, E., Kolaitis, R.M., Molliex, A., Ruth, E., Alami, N.H., Wang, B., Joshi, A., Smith, R.B., Ritson, G.P., et al. (2014). VCP Is Essential for Mitochondrial Quality Control by PINK1 / Parkin and this Function Is Impaired by VCP Mutations. Neuron. *78*,

65–80.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *9*, 357–360.

Lee, S., Kopp, F., Chang, T., Yu, H., Xie, Y., and Mendell, J.T. (2016). Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. Cell *164*, 69–80.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J. Cell Biol. *201*, 361–372.

Lima, S.A., Chipman, L.B., Nicholson, A.L., Chen, Y.-H., Yee, B.A., Yeo, G.W., Coller, J., and Pasquinelli, A.E. (2017). Short poly(A) tails are a conserved feature of highly expressed genes. Nat. Struct. & Mol. Biol. *24*, 1057.

Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., Lee, J.H., Namkoong, S., Ho, A., Woo, Y.M., Kwak, H., and Lee, J.H. (2018). Granulation Resource Systematic Characterization of Stress-Induced RNA Granulation. Mol. Cell. *70*, 175–187.e8.

Niewidok, B., Igaev, M., Pereira, A., Strassner, A., Lenzen, C., Richter, C.P., and Piehler, J. (2018). Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. J. Cell Biol. *217*, 1303–1318.

Protter, D.S.W., and Parker, R. (2016). Principles and Properties of Stress Granules. Trends Cell Biol. *26*, 668–679.

Reineke, L.C., and Lloyd, R.E. (2013). Diversion of stress granules and P-bodies during viral infection. Virology *436*, 255–267.

Somasekharan, S.P., El-Naggar, A., Leprivier, G., Cheng, H., Hajee, S., Grunewald, T.G.P., Zhang, F., Ng, T., Delattre, O., Evdokimova, V., et al. (2015). YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. J. Cell Biol. *208*, 913–929.

Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly (A) -tail profiling reveals an embryonic switch in translational control. Nature *508*, 66–71.

Tourrière, H., Chebli, K., Zekri, L., Courselaud, B., Blanchard, J.M., Bertrand, E., and Tazi, J. (2003a). The RasGAP-associated endoribonuclease G3BP assembles stress granules. J. Cell Biol. *160*, 823 LP-831.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotech *31*, 46–53.

VanTreeck, B. Van, Protter, D.S.W., Matheny, T., Khong, A., Link, C.D., and Parker, R. (2018). RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. 1–6.

Wheeler, J.R., Matheny, T., Jain, S., Abrisch, R., and Parker, R. (2016). Distinct stages in stress granule assembly and disassembly. Elife *5*, 1–25.

Wheeler, J.R., Jain, S., Khong, A., and Parker, R. (2017). Isolation of yeast and mammalian stress granule cores. Methods *126*, 12–17.

Youn, J., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M., Angers, S., Morris, Q., Fabian, M., Cote J., Gingra, A. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies Resource High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol Cell. *69* 517–532.