CELL-FREE FORMATION OF RNA GRANULES

APPROVED BY SUPERVISORY COMMITTEE

Steven McKnight, Ph.D.

Kim Orth, Ph.D.

Ege Kavalali, Ph.D.

Jose Rizo-Rey, Ph.D.

Hongtao Yu, Ph.D.

CELL-FREE FORMATION OF RNA GRANULES

By

TINA WEI HAN

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The Univeristy of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

May, 2012

CELL-FREE FORMATION OF RNA GRANULES

TINA WEI HAN, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2012

STEVEN L. MCKNIGHT, Ph.D.

Asymmetric RNA localization is a mechanism by which a cell can spatially and temporally regulate the translation of RNAs. This mechanism is essential for many developmental processes such as germ cell formation in *Drosophila* embryos, as well as establishment of cell polarity and synaptic plasticity in the brain. In many instances, asymmetric RNA localization is achieved through transport and sequestration by RNA granules. RNA granules are large, non-membrane bound ribonucleoparticles that have been observed in various biological contexts. Unfortunately, because RNA granules are highly

heterogeneous and weakly associating aggregates, they can be difficult to study biochemically, which constitutes a major impediment for gaining a more detailed understanding of the mechanisms governing RNA granule assembly. Here we describe two *in vitro* models for studying RNA granule assembly. The first method is based on the precipitation activity of a 3,5-disubstituted isoxazole compound that can be used as a quick and efficient pharmacological tool to probe the function and regulation of RNA granules. The second method utilizes a three-dimensional protein-retaining hydrogel formed from a recombinant protein. Polypeptides of low amino acid complexity were found to be the sequence determinants of isoxazole precipitation and hydrogel retention. Next generation sequencing was used to identify RNAs that partitioned with granule components in both isoxazole and hydrogel models and were found to be enriched in mRNAs known to be constituents of neuronal transport granules for dendritic localization. The overrepresented gene ontology categories for these RNAs included cell adhesion, extracellular matrix, and synaptic proteins. The average length of the 3'UTR of these RNAs was found to be longer than the 3'UTRs of RNAs excluded from the cell-free RNA granule preparations. These two *in vitro* models for studying RNA granule assembly offer a novel approach to identify candidate targets recruited to RNA granules by specific RNA-binding proteins.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Han TW, Kato M, Xie S, Wu L, Mirzaei H, Pei J, Chen M, Xie Y, Allen J, Xiao G, McKnight SL. Cell-free formation of RNA granules: identification of RNA components reveals logic of granule assembly. *In press.*.

Kato M, **Han TW**, Xie S, Du X, Wu L, Mirzaei H, Pei J, Schneider J, Tycho R, Eisenberg D, McKnight SL. Cell-free formation of RNA granules: protein components utilize low complexity sequences for phase-transition from soluble to hydrogel-like state. *In press.*

Pieper AA, Wu X, **Han TW**, Estill SJ, Dang Q, Wu LC, Reece-Fincanon S, Dudley CA, Richardson JA, Brat DJ, McKnight SL. The neuronal PAS domain protein 3 transcription factor controls FGF-mediated adult hippocampal neurogenesis in mice. *PNAS* 102: 14052-7 (2005).

Erbel-Sieler C, Dudley C, Zhou Y, Wu X, Estill SJ, **Han T**, Diaz-Arrastia R, Brunskill EW, Potter SS, McKnight SL. Behavioral and regulatory abnormalities in mice deficient in the NPAS1 and NPAS3 transcription factors. *PNAS* 101: 13648-53 (2004)

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Ago	Argonaute
ALS	Amyotrophic lateral sclerosis
AMPA	(2-amino-3-(5-methyl-3-oxo-1,2- oxazol-4-yl) propanoic acid)
BME	Beta-mercaptoethanol
Bps	Base pairs
CPEB	Cytoplasmic polyadenylation element binding protein
CREB	cAMP response element binding protein
DCP	Decapping protein
DDX	DEAD-box helicase
DEAD	Asp-Glu-Ala-Asp aka Walker B motif
DHPG	Dihydroxyphenylglycine
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescence protein
EGFR	Epidermal growth factor receptor
eIF	Eukaryotic initiation factor
ER	Endoplasmic reticulum
FG	Phenylalanine glycine
GLH	Germline helicase
GTP	Guanosine triphosphate
hnRNP	Heterogeneous ribonucleoparticle
IPTG	Isopropyl-β-D-thiogalactopyranoside
Isx	Biotinylated isoxazole
KIF	Kinesin superfamily protein
LCS	Low complexity sequence
MBP	Myelin basic protein

mCh	mCherry
mGluR	Metabotropic glutamate receptor
miRNA	MicroRNA
mRNA	Messenger RNA
NLS	Nuclear localization signal
NPC	Nuclear pore complex
Nts	Nucleotides
NUP	Nucleoporin
Oligo	Oligonucleotide
PABP	Poly-A binding protien
PSD	Postsynaptic density
RNAi	RNA interference
RNase	Ribonuclease
RNP	Ribonucleoparticle
rRNA	Ribosomal RNA
SG	Stress granule
tRNA	Transfer RNA
UTR	Untranslated region

CHAPTER 1:

RNA granules mediate asymmetric localization of RNA

Introduction

Asymmetric localization of messenger RNAs is a mechanism used by cells to restrict protein synthesis to specific subcellular compartments. In many instances, RNA-binding regulatory proteins recruit their cognate RNAs into ribonucleoparticles called RNA granules to facilitate transport and translational control. In this chapter, I will present an overview of the biological roles of localized RNAs. I will next describe the regulation and function of RNA granules that mediate RNA localization. Finally, I will discuss the components that are common to multiple types of granules and that contribute to a unified concept of RNA granule function.

Biological roles of asymmetric RNA localization

Localized RNAs are important to diverse biological processes and cellular states such as segregating determinants of cell fate, asymmetric cell division, cell polarity, and synaptic plasticity. First, uneven distribution of maternal RNAs in *Drosophila* oocytes allows for the establishment of morphogen gradients required for axial patterning and asymmetrical partitioning of cytoplasm determinants into germline precursor cells (Palacios and St. Johnston, 2001). For example, *bicoid*

mRNA is localized to the anterior pole of oocytes and is both necessary and sufficient for the formation of head and thorax structures in the fly (Berleth et al., 1988; Driever et al., 1990). After fertilization, *bicoid* RNA is translated to produce a homeodomain transcription factor that diffuses posteriorly, establishing a morphogen gradient that patterns the anterior embryo. On the other hand, *oskar* mRNA is posteriorly localized and is required for the recruitment of posterior determining factors such as *nanos* mRNA to the pole plasm, a crucial developmental event that gives rise to both abdomen formation and pole cells, precursors of the *Drosophila* germ line (Kim-Ha et al., 1991; Ephrussi et al., 1991; Ephrussi and Lehmann, 1992).

Second, RNA localization has been shown to play a crucial role in asymmetric cell division, most notably in yeast mating type switch. During budding, ASH1 mRNA preferentially accumulates in the daughter cell at the end of anaphase (Long et al., 1997; Takizawa et al., 1997). Ash1p (the protein product of ASH1) represses the expression of HO endonuclease that specifically cleaves the MAT locus, which leads to a mating type switch. Therefore, due to the asymmetric partitioning of Ash1p to the daughter cell, the mother cell inevitably undergoes a mating type switch.

Third, localized translation of RNAs can lead to the establishment and reinforcement of cell polarity. For example, accumulation of β-actin mRNA in the leading edges of chick embryonic fibroblasts directs the site of active actin

polymerization to drive cell migration (Kislauskis et al., 1993; Kislauskis et al., 1994). β-actin mRNA is also transported to the tips of growth cones in developing axons seeking synaptic connections (Zheng et al., 2001). Following treatment with antisense oligonucleotides directed towards the β-actin 3'UTR, neurons experienced reduced motility and, in some cases, retraction of growth cones (Zheng et al., 2001).

Lastly, active transport of specialized RNAs to neuronal processes for localized translation in or near synapses is implicated in modulation of synaptic plasticity and establishment of synaptic memories. One of the best characterized dendritically targeted mRNA in mammals is the transcript that encodes calcium/calmodulin-dependent protein kinase II alpha (Camk2a; Burgin et al., 1990). Miller et al. (2002) generated a mutant mouse where the protein coding region of Camk2a was intact, but the 3'UTR of Camk2a was altered in a way that the transcript was retained exclusively in the cell soma. These mutant mice exhibited a reduction of late-phase long-term potentiation in hippocampal field recordings and deficits in spatial and associative behavioral paradigms.

Biology of RNA granules

In many instances, mRNA transcripts are packaged into ribonucleoparticle (RNP) aggregates called RNA granules, which are large, non-membrane-bound cytoplasmic organelles containing heterogeneous mixtures of proteins and RNAs,

and passively tethered or actively transported to specific sites in the cytoplasm (Anderson and Kedersha, 2009). In germ cells, these granules locally store mRNA that is unique to germ line differentiation until needed for development (Voronina et al. 2011). In somatic cells, environmental stresses can induce translational arrest of a select subset of RNAs sequestered into stress granules (Buchan and Parker, 2009). Another class of somatic RNA granules called processing bodies (P bodies) contain the basic enzymatic machinery for mRNA decay and miRNA repression (Buchan and Parker 2009). In post-mitotic neurons, RNA granules are tasked with the unique challenge of delivering translationally silent mRNAs to distal sites in the dendrites for local protein synthesis (Doyle and Kiebler, 2011).

Morphologically, RNA granules are commonly described in similar terms. By electron microscopy, they appear to be roughly spherical granular structures with heterogeneous density that sometimes appears fibrillar (Knowles et al., 1996; Amikura et al., 2001; Souquerre et al., 2009). Although many types of RNA granules have been described in the literature, they all share similar sets of proteins (Fig 1-5), suggesting that these entities may use similar underlying mechanisms for assembly (Anderson and Kedersha 2006, Buchan and Parker 2009, Elvira et al. 2006, Kanai et al. 2004, Voronina et al. 2011).

In the following section, I will review the biological roles of different cytoplasmic RNA granules: neuronal granules, germ granules, stress granules, and processing bodies.

Germ granules

In the early 1900s, Hegner described dark granules in the posterior cytoplasm of chrysomelid beetle oocytes. He found that the beetles failed to develop germ cells as a consequence of ablating this region with a hot needle (Hegner 1908; Hegner 1911). Since then, germ granules have been observed in most if not all species studied and have several different names depending on the species of origins (Eddy, 1975). In *Xenopus*, they are called germinal granules (Kloc et al., 2004), in *Drosophila*, polar granules (Fig 1-1A; Mahould, 1962), in *C. elegans*, P granules (Updike and Strome, 2010), and in mammals, chromatoid bodies (Yokota et al., 2008). Early germ plasm transplantation studies done in *Drosophila* demonstrated that germ granules were likely determinants for germ cell fate. In *Drosophila*, granule components are first produced in germ line feeder cells called nurse cells and deposited into oocytes through intercellular bridges during oogenesis where they then preferentially localize to the posterior pole of the embryo (Mahowald, 1968). Following fertilization, polar granules asymmetrically segregate to pole cells, precursors of the *Drosophila* germ line, which form in the posterior of the embryo (Bashirullah et al., 1998). Because

polar granules were observed to be highly enriched with RNAs and associated with pole cells, Mahowald (1968) hypothesized that they regulate a select subset of maternal transcripts that are specifically translated in the early embryo to give rise to the germ line. Accordingly, UV irradiation of egg posterior, which would likely damage granule RNAs, disrupts germ cell development, but pole cells are restored through transplantation of germ plasm from non-irradiated eggs (Okada et al., 1974). Furthermore, transplantation of germ plasm to the anterior pole of *Drosophila* oocytes induced ectopic germ cell development (Illmensee and Mahowald, 1974), indicating that the germ plasm RNAs were necessary and sufficient for germ cell development. After fertilization of the egg, polar granules were observed to shed electron density corresponding to RNA content (Mahowald, 1971), which coincided with the formation of pole cells, suggesting that perhaps these RNAs had been released for translation. Additionally, poly A+ RNA isolated from late stage oocytes but not late stage embryos can rescue pole cell development in UV irradiated embryos (Okada and Kobayashi, 1987) suggesting that RNAs localized to the germ plasm and regulated by polar granules might be germ cell determinants.

Interestingly, P granules in developing gonad cells of adult worms are exclusively perinuclear in distribution and appear to tightly associate with clusters of nuclear pore complexes (Fig 1-2B; Pitt et al., 2000). Using a genomewide RNAi screen, Updike and Strome (2009) found that 14 out of 20 nuclear pore

components were required for the perinuclear distribution of P granules. Furthermore, a *C. elegans* homolog of Nup98 was not only identified as a P granule component itself, but also found to be required for granule integrity as knockdown of Nup98 in worms caused the cytoplasmic dispersion of constitutive P granule components (Voronina and Seydoux, 2010). RNA helicases in P granules GLH-1, GLH-2, and GLH-4 have phenylalanine-glycine (FG) repeats that are similar to those found in nucleoporins, including Nup98 (Updike et al., 2011). In the nuclear pore complex, FG repeat-containing nucleoporins are thought to be involved in the selective gating of macromolecules traversing through the nuclear pore channel (Wälde and Kehlenbach, 2010). It is interesting to speculate that the FG-containing helicases might interact with FG nucleoporins to form the basis of the perinuclear localization of P granules.

Neuronal transport granules

Discovery of neuronal transport granules began with the observation that RNA pulse-labeled with ^{3}H-uridine migrated from the cell body of cultured hippocampal neurons into dendritic processes (Davis et al., 1987). Ainger et al. (1993) microinjected RNA encoding myelin basic protein (MBP) into cultured oligodendrocytes and observed aggregation of RNA into large particles that were heterogeneous in size, but all roughly circular. Using electron microscopy to visualize SYTO14-stained RNA, Knowles et al., 1996 showed that these

endogenous RNA were packaged into large granular particles along with clusters of ribosomes (Fig 1-2A). Due to the highly polarized nature of post-mitotic neurons, the key distinguishing feature of neuronal granules is the ability to transport RNA over long distances. Considering that in some cases the time it would take a protein to traverse the length of an axon might be longer than the half-life of the protein itself, a mechanism to facilitate local protein synthesis must exist (Campenot and Eng 2000). Early experiments indicated that dendritic targeting of RNAs depended on cytoskeletal elements (Davis et al., 1987; Ainger et al., 1993). Moreover, treatment with microtubule depolymerizing drugs like nocodazole reduced levels of RNAs and RNA-binding proteins in neuronal processes, but disruption of actin polymerization by pharmacological means seemed to have little or no effect (Knowles et al., 1996; Köhrmann et al., 1999; Bannai et al., 2004). Observed velocities of granule movement (ranging from 0.1-1 um/s) suggested that RNA transport relied on molecular motors such as dyneins and kinesins (Davis et al., 1987; Ainger et al., 1993; Knowles et al., 1996; Köhrmann et al., 1999). Accordingly, knockdown of conventional kinesin heavy chain (Kif5) impaired the ability of microinjected MBP RNA to move into oligodendrocytic processes (Ainger et al., 1993) Furthermore, Kanai et al., 2004 immunoprecipitated RNA granules associated with Kif5 and found 42 RNA-associated proteins and known dendritic mRNAs Camk2a and Arc.

Ultrastructural studies of neuronal synapses hinted at the potential for localized mRNA translation in or near dendritic spines. Using electron microscopy, Steward and Levy, 1983 observed polyribosomes clustering preferentially at the base of spines and postsynaptic density-like mounds. Only about 12% of the dendritic spines were associated with polyribosomes (Steward and Levy, 1983), which suggests that perhaps only a subset of spines would be competent for local protein synthesis. Microdissected dendrites are able to incorporate radiolabelled amino acids (Torre and Steward, 1992) and synthesize GFP-tagged proteins (Aakaln et al., 2002; Job and Eberwine, 2001), which demonstrated that dendrites harbored functional translational machinery. Even if dendrites were competent for mRNA translation, a structure performing ER-like functions would still be required for the synthesis of integral membrane proteins and posttranslational modifications necessary for the maturation of newly synthesized proteins such as glycosylation. About half of the polyribosomes localized to dendritic spines were also associated with tubular cisterns (Steward and Reeves, 1988) which Pierce et al., 2000 showed to stain positive for the Sec61a complex which is required for the translocation of newly synthesized proteins through the ER. Microdissected dendrites are able to translate mRNAs encoding transmembrane proteins and insert them into the plasma membrane (Kacharmina et al., 2000), suggesting that some organelle competent for ER functions resided in neuronal processes. By EM immunogold staining, Pierce et

al., 2000 also showed that the tubular cisterns were also positive for Golgi markers, but this seems to be a controversial issue as Horton and Ehlers, 2003 localized Golgi markers to the dendritic shaft instead. The spine appartus itself has been speculated to be a mini-Golgi apparatus, but so far it has not been implicated in any Golgi functions (Steward and Reeves, 1988).

Localized translation of mRNAs at the synapse has been proposed as a mechanism for modulating synaptic plasticity and establishing synaptic "memories" (Fig 1-2B; Kiebler and DesGroseillers, 2000). In other words, one might imagine that activity at a synapse can trigger translation of proteins to strengthen the fruitful connections and facilitate future activity. Silence at a neighboring synapse might tip the balance towards protein turnover and pruning of inefficient connections. So far, a significant subset of identified dendritically targeted RNAs encode proteins important for building the postsynaptic density such as Camk2a, Arc, Shank1/3, AMPA receptor subunits GluR1/2, and PSD95 (Burgin et al., 1990; Link et al., 1995; Lyford et al., 1995; Böckers et al., 2004; Grooms et al., 2006; Muddashetty et al., 2007). Pharmacological manipulations that simulate activate synapses such as neurotrophin, BDNF or global neuronal depolarization induced by KCl and electroconvulsive seizures promote granule transport into the dendrites (Zheng et al., 2001; Rook et al., 2000; Wallace et al., 1998). In addition, KCl treatment of partially purified granules appear to relax and unwind (Krichevsky and Kosik, 2001), which could be interpreted as

releasing mRNAs for translation in response to synaptic activation. Finally, depletion of neuronal granule proteins such as Pumilio2, FMRP, Staufen1/2, CPEB1 and FUS leads to the development of abnormal spine morphology in cultured hippocampal neurons (Vessey et al., 2010; Comery et al., 1997; Vessey et al., 2008; Goetze et al., 2006; McEvoy et al., 2007; Fujii et al., 2005), which implicates neuronal granule function as important for the development and maintenance of dendritic spine structures and, by extension, synaptic connectivity.

Stress granules

Stress granules (SGs) were first observed in cultured tomato plant cells treated with mild heat shock (Nover et al., 1983). These cytoplasmic aggregates were found to accumulate small heat shock proteins and translationally repressed mRNAs (Nover et al., 1983; Nover et al., 1989). Subsequently, a variety of cellular stresses in addition to heat shock such as oxidative stress (ie, arsenite treatment), osmotic shock, UV irradiation as well as energy depletion through inhibiting glycolysis or mitochondrial functions were found to be capable of inducing stress granule formation in mammalian cells (Fig 1-3; Kedersha et al., 1999). Studies in yeast suggest that different stresses can lead to the assembly of stress granules with slightly different composition (Buchan et al., 2010). Arsenite treatment in particular was found to disrupt translational initiation through the phosphorylation of a particular serine residue S51 of translation initiation factor

eIF2a (Kedersha et al., 1999). Phosphorylation of this factor sequesters it away from priming the 48S preinitiation complex with the initiator $tRNA^{Met}$ (Kedersha et al., 2002). When S51 was mutated to alanine, arsenite treatment failed to induce SGs and a S51D mutant that mimics phosphorylated serine was able to induce SG assembly in the absence of stress (Kedersha et al., 1999). Furthermore, promoting mRNA dissociation from polysomes with puromycin can stimulate SG formation, while trapping mRNAs in polysomes using cyclohexamide or emetine leads to a reduced number of stress granules (Kedersha et al., 2000; Buchan et al., 2008). These key observations led researchers to believe that the function of stress granules in the cultured cells was to repress mRNAs from active translation (Kedersha et al., 2000; Buchan et al., 2008). Overexpression of self-aggregating proteins such as TIA1 and G3BP can nucleate SG assembly in the absence of stress and phospho-eIF2a (Gilks et al., 2004; Tourriere et al., 2003). TIA1 has a C-terminal glutamine-rich prion-like domain (Tian et al., 1991). Gilks et al. (2004) found that the last 97 amino acids from the C-terminal end are required for the recruitment of TIA1 into stress granules upon induction with arsenite and replacing it with the aggregation domain from yeast prion Sup35 can functionally rescue localization. This hints at the nature of interactions that may underlie stress granule assembly.

Stress granules have also been detected in the tissues of stressed animals. For example, in chickens treated with an antibiotic gentimycin that damages the

ear, stress granules appear in cochlear cells shortly before cell death (Mangiardi et al., 2004). In a rat model of stroke, stress granules accumulated in the cell bodies of cells in the dentate gyrus following 10 minutes of global brain ischemia, persisted 10 minutes into reperfusion, and fully disassembled after 90 minutes of reperfusion (Kayali et al., 2005). Stress granules can be induced in single-cell eukaryotes such as yeast and trypanosomes by glucose deprivation (Buchan et al., 2011; Cassola et al., 2007). More importantly, Cassola et al., 2007 found that trypanosomes isolated from the intestinal tract of insect hosts also contain stress granules and speculate that granule formation may protect non-translating transcripts during periods of starvation. These studies suggest that stress granules are not *in vitro* artifacts of cultured cell lines only induced by non-physiological stresses.

Processing bodies

The majority of mRNAs targeted for destruction are initiated by deadenylation of the 3' poly-A^+ tail and either followed by decapping of the modified guanosine at the 5' end and degraded by a 5'→3' exonuclease or directly degraded 3'→5' via an exosome associated complex (Parker and Sheth, 2007). Observations that mRNA decapping factors Dcp1 and Dcp2 and other proteins related to 5'→3' mRNA decay such as 5'→3' exonuclease Xrn1, Lsm proteins, mRNA deadenylase factor Ccr4, and DEAD-box helicase Rck/p54/Ddx6 colocalize to the

same cytoplasmic foci led to the discovery of processing bodies (P-bodies; Fig 1-4; van Dijk et al., 2002; Ingelfinger et al., 2002, Cougot et al., 2004; Sheth and Parker, 2003). Although these cytoplasmic bodies were different from stress granules because they were constitutively visible in the cell and because they did not co-localize with TIA1, a well-characterized marker for stress granules (Cougot et al., 2004), various cellular stresses that induce SG formation also increase P-body size and numbers (Sheth and Parker, 2003; Kedersha et al., 2005). P-bodies also contained components related to miRNA repression such as Ago1/2 and GW182, though this was not a conserved aspect in yeast (Sen et al., 2005; Liu et al., 2005; Eyastathiony et al., 2002, 2003). Using a reporter RNA that encoded luciferase with a let-7 miRNA binding site and 12x MS2 hairpins in the 3'UTR, Liu et al. (2005) demonstrated that mRNAs could be localized to Ago2-containing P-bodies in a miRNA or miRNA site dependent manner. In addition, RNA itself seems to be required for maintaining the structural integrity of partially purified P bodies *in vitro* as well as Ago2-containing P-bodies in permeabilized cells (Teixeira et al., 2005; Sen et al., 2005). Yeast strains lacking one or more P-body components were found to be defective in the global translational repression typically induced by glucose deprivation or amino acid starvation (Holmes et al., 2004; Coller and Parker, 2005), which seems to argue that P-bodies were essential to the general translational regulation machinery of the cell. However, when P-bodies are depleted through knockdown of

aggregation factors such as Lsm3p, GW182, or Ed3p, no significant detriment was observed on mRNA decay (Chu and Rana, 2006; Stoeckline et al., 2006; Decker et al., 2007). Parker and Sheth, 2007 speculate that the larger, visible aggregations may be more efficient, but smaller aggregates can still function. Finally, many P-body proteins have been characterized with "prion-like" domains (Decker et al., 2007; Reijns et al., 2008)

The C-terminal of Lsm4p, member of the Lsm1-7 complex involved in mRNA degradation, was first identified in a computational analysis of Q/N-rich proteins in the yeast proteome (Michelitsch and Weissman, 2000). Decker et al. (2007) found that the deletion of the C-terminal of Lsm4p abolishes formation of P-bodies under glucose deprivation in yeast strains lacking Edc3p, which also has a Q/N-rich polypeptide sequence located near the N-terminus (Reijns et al., 2008). These data suggest that aggregation mediated by prion-like domains may be involved in RNA granule assembly, which is a concept that will be explored in further detail in Chapter 3.

Components of RNA Granules

Working models for RNA granule function are inherently built around what we know about the function of its protein components. To date, most of the known components are related to RNA metabolism, which strongly suggests that the primary function of RNA granule function is for post-transcriptional regulation

(Anderson and Kedersha, 2006; Buchan and Parker, 2009; Voronina et al., 2011). Early attempts at identifying these components have mostly been through a candidate approach, although many germ granule factors were discovered through maternal effect screens in *Drosophila* which led to the discovery of a class of genes known as posterior group genes (Bashirullah et al., 1988). Females homozygous for mutations in posterior group genes gave rise to sterile progeny that lacked polar granules and germ line development (Bashirullah et al., 1988). Well-characterized examples from this group include *vasa*, *staufen*, and *oskar* which will each be discussed in this section under the appropriate subject heading (Hay et al., 1988; St Johnston et al., 1991; Lehmann and Nusslein-Volhard, 1986). Many lessons learned from germ granule studies can also be applied to our understanding of other types of granules, which we will explore further in the following discussion. Later studies of RNA granules applied less biased proteomic techniques to investigate this important problem, which still remains unresolved due to the intrinsically heterogeneous composition of RNA granules. The following functional classes of proteins are common to multiple types of granules and contribute to an understanding of unifying themes in RNA granule function.

Ribosomes/translation initiation factors

Translation initiation begins with the assembly of a diverse array of factors on a dissociated 40S ribosomal subunit (Fig 1-6). First, a large scaffolding factor eIF3 joins the 40S ribosomal subunit, followed by the ternary complex eIF2-GTP-$tRNA_i^{Met}$ which primes the 43S preinitiation complex with the initiator $tRNA^{Met}$ (Gingras et al., 1999). Next, the eIF4 protein complex recruits mRNAs to the 43S complex via 5' cap binding and interactions with PABP which binds the poly-A tail, forming the 48S initiation complex (Gingras et al., 1999). Once the 48S complex accesses a viable start codon, a GTPase-activating protein eIF5 promotes GTP hydrolysis by eIF2 which releases the initiation factors from the 40S subunit, allowing the 60S subunit to bind and processive ribosome scanning for start codons to commence (Gingras et al., 1999). All four types of RNA granules discussed here contain either full or partial ribosomes or translation initiation factors. This is not unexpected as one would assume the RNA localization provided by granule regulation necessarily requires translational repression of the transcript until the final destination for the protein product has been reached and/or the appropriate environmental stimuli are received.

Stress granules co-localize with most of the components of the 48S initiation complex including small ribosomal subunits, eIF3, eIF4E, and eIF4G, but not eIF2 (Kedersha et al., 2002). These immunofluorescence results were supported with the observation by electron microscopy that 18S rRNA but not 28S rRNA is also enriched in stress granules (Souquere et al., 2009). Since

phosphorylation of eIF2a robustly induces the formation of stress granules, it was hypothesized that stress granules essentially assembled around stalled 48S complexes that failed to initiate translation due to the lack of priming by initiator $tRNA^{Met}$ (Kedersha et al., 2002). Interestingly, although stress granules induced by sodium azide (NaN_3) in yeast co-localize with eIF3, eIF1A, eIF4A, and eIF4B, granules induced by glucose deprivation in yeast do not contain any translation initiation factors (Buchan et al., 2010) indicating that the type of stress may influence the composition of the granule. Furthermore, these yeast NaN_3-induced granules also contained eIF5B which is a GTPase involved in the assembly of the full ribosome downstream of the 48S initiation complex, inconsistent with the presented model of stress granule assembly in mammalian cells (Buchan et al., 2010; Gingras et al., 1999).

The eIF4 protein complex, which mediates recruitment of ribosomes to mRNAs, includes the following proteins: eIF4E, a 5' mRNA cap binding protein, eIF4A, a canonical DEAD box RNA helicase, and eIF4G, an adaptor molecule that binds to PABP and effectively circularizes the transcript with indirect attachments to both the cap and tail of the mRNA (Gingras et al., 1999). eIF4E specifically has been found to co-localize with Dcp1a, Me31B/DDX6, and Lsm1 in P-bodies, and its recruitment is mediated by the 4E-transporter (4E-T) protein (Ferrainolo et al., 2005; Andrei et al., 2005). 4E-T is a repressor of cap-dependent translation and knockdown of 4E-T prolongs the half-life of mRNA reporters

targeted for degradation (Ferrainolo et al., 2005). Knockdown of 4E-T causes a lack of accumulation of Dcp1a, Me31B/DDX6, and eIF4E to P-bodies (Ferrainolo et al., 2005) which suggests that 4E-T might be important for P-body formation.

By electron microscopy, neuronal transport granules have been shown to be enriched with ribosomal subunits, both large and small (Knowles et al., 1996; Elvira et al., 2006; Krichevsky and Kosik., 2001; Kanai et al., 2004). However, evidence that granules also contain translation initiation factors is insubstantial. Elvira et al., 2006 identified eIF4A in a partially purified neuronal granule preparation. Kanai et al., 2004 immunoprecipitated RNA granules associated with Kif5 and identified eIF2a, eIF2b, and eIF2g among the proteins. According to the mammalian stress granule model, when eIF2a is excluded from the preinitiation complex, translation stalls and stress granules form (Kedersha et al., 2002). One way to explain the scarce representation of translation initiation factors in neuronal granules may be that neuronal granules assemble around intact 80S ribosomes with mRNAs already locked and primed for translation, allowing for rapid jumpstart of translation elongation once the correct signal (ie, synaptic transmission) is received.

Finally, mitochondrial ribosomes have been reported to associate with the periphery of polar granules in *Drosophila* prior to pole cell formation (Amikura et al., 2001). Mitochondrial rRNAs have also been observed briefly on the surface of *Xenopus* germinal granules, also during a narrow window of developmental

time (Kashikawa et al. 2001). Interestingly, the microinjection of mitochondrial large rRNA rescues the germ line in UV-irradiated *Drosophila* embryos and can ectopically induce pole cell formation in the anterior embryo when co-injected with UV-irradiated germ plasm (Kobayashi and Okada, 1989). Because of the developmental timing of mitochondrial ribosome association with germ granules, these studies suggest that translation mediated by the mt ribosomes is possibly required for germ cell development and that the UV-sensitive component in the germ plasm is in fact ribosomal RNA and not poly-A^+ RNAs. Whether mitochondrial ribosomes are actually responsible for translation of maternal transcripts bound by germ granules, however, has not been proven. In general, not as many components have been identified for germ granules as for other types of granules, but more recent proteomic approaches have been able to achieve some new results. Thomson et al., 2008 immunoprecipitated two polar granule components *tudor* and *vasa* and used mass spectrometry to identify overlapping interacting proteins. They observed eIF4A, the RNA helicase component of the eIF4 complex, to be in common with both *tudor* and *vasa*-associated proteins and confirmed the localization of eIF4A to polar granules by electron microscopy (Thomson et al., 2008). Other members of the eIF4 protein complex have not yet been identified in polar granules, though an isoform of eIF4E is found in P granules and is required for spermatogenesis in *C. elegans* (Amiri et al., 2001). It is possible that the presence of these eIF4 proteins indicates the sequestration of

non-translating RNAs in germ granules for localized storage as is the case for stress granules.

Heterogeneous nuclear ribonucleoproteins (hnRNPs)

Heterogeneous nuclear ribonucleoproteins or hnRNPs are loosely defined as the set of proteins that associate with pre-mRNAs that are primary protein-coding transcripts produced by RNA polymerase II (Dreyfuss et al., 2002). As such, this nomenclature is practically meaningless -- more a reminder of how these proteins were historically discovered rather than descriptive of function. In addition to their nuclear functions that includes regulation of mRNA splicing, a subset of hnRNP proteins also accompany mRNAs out into the cytoplasm and participate in the localization and translational regulation of the RNAs (Dreyfuss et al., 2002). In polar granules, a *Drosophila* hnRNPA/B homolog called Hrp48 directly binds to *oskar* mRNA and is required for early *oskar* localization to the posterior pole of the oocyte (Yano et al., 2004; Huynh et al., 2004). Another hnRNP protein in *Drosophila* called Squid or Hrp40 is required for the localization of gurken mRNA that encodes a EGFR ligand specifically expressed in germ line cells (Kelley et al., 1993; Neuman-Silberberg and Schüpbach, 1993). 16 hnRNP proteins were identified by mass spectrometry in partially purified neuronal transport granules from developing rat brains as granule components (Elvira et al., 2006). Six of these proteins, hnRNPD, hnRNPA1, SYNCRIP, hnRNPA/B,

hnRNPA0, and hnRNPU were also independently identified in a study of neuronal granules that associate with conventional kinesin (Kanai et al., 2004). Kanai et al., 2004 additionally identified three other hnRNP proteins, FUS, EWS, and ALY, among their granule components. Of these 19 hnRNPs labelled as neuronal granule proteins, 5 have additionally been detected in stress granules: hnRNPA1, hnRNPK, SYNCRIP, TDP-43, and pathological versions of FUS (Guil et al., 2006; Fukuda et al., 2009; Quaresma et al., 2009; Colombrita et al., 2009; Bosco et al., 2010).

FUS is a particularly interesting example of an hnRNP protein found in RNA granules. Originally named TLS for *T*rans*L*ocated in *S*arcoma because it was first identified in a chromosomal translocation event in myxoid liposarcoma that fused the first 266 amino acids of FUS with the C-terminal end containing the DNA-binding domain of transcription factor CHOP, FUS was identified as hnRNP P2 by electrospray mass spectrometry (Crozat et al. 1993; Calvio et al. 1995). Like many hnRNPs, FUS appears to predominantly localize in the nucleus at steady state, but actually is capable of shuttling between the nucleus and cytoplasm to transport mRNA (Han et al. 2010; Zinszner et al. 1997). Mutations in the FUS protein were found in patients with an inherited form of neurodegenerative disease called amyotrophic lateral sclerosis or ALS (Kwiatkowski et al. 2009, Vance et al. 2009). Many of the mutations identified to date occur in the highly conserved C-terminal (amino acids 500-526) that encodes

a non-canonical nuclear localization signal (NLS) and in the long span of low complexity sequence at the N-terminus (Fig 1-7A; Lagier-Tourenne and Cleveland, 2009; Sun et al., 2011). Furthermore, in human post-mortem tissues, it appears that the FUS mutations caused mislocalization and cytoplasmic aggregation of the protein in motor neurons and glial cells (Fig 1-7B; Kwiatkowski et al. 2009, Vance et al. 2009). Wild type FUS can only be detected as nuclear in steady state, whereas ALS-linked nonsense mutations G515X and R495X, which result in FUS protein partially or fully missing the C-terminal NLS, are both robustly recruited to stress granules when stably transfected into HEK293 cells (Fig 1-7C; Bosco et al., 2010). In the brain, FUS localizes to neuronal dendrites in a microtubule-dependent manner, and transient activation of class I metabotropic glutamate receptor (mGluRs) by a small molecule agonist DHPG causes FUS to move into postsynaptic spines (Fujii et al. 2005) Since DHPG is known to stimulate translocation of mRNAs into dendritic spines (Job and Eberwine 2001), these data support the idea that FUS is a component of neuronal transport granules that deliver dendritic mRNAs for local protein synthesis (Kanai et al., 2004). Moreover, hippocampal neurons cultured from $FUS^{-/-}$ mice show abnormal spine morphology and fewer spine numbers (Fujii et al. 2005), indicating that FUS, and by extension functional RNA granules, are important to the development and/or maintenance of dendritic spines.

Involvement of hnRNP proteins that accompany mRNAs into cytoplasmic RNA granules suggest that complexes pre-formed in the nucleus are important for cytoplasmic localization and translational regulation of mRNAs.

DEAD-box RNA helicases

DEAD-box helicases have RNA-dependent ATPase activity and unwind secondary structures in RNA in an ATP-dependent manner *in vitro* (Fuller-Pace, 2006). Proteins in the DEAD-box family share 9 conserved motifs, one of which being the Asp-Glu-Ala-Asp or D-E-A-D (also known as Walker B) motif that the family is named for (Linder, 2006). This motif along with 3 other conserved motifs is required for ATP binding and hydrolysis (Linder, 2006). DEAD-box helicases are involved in diverse, RNA-related functions including pre-mRNA splicing, ribosome biogenesis, mRNA export, mRNA decay and translation initiation (Linder, 2006). The founding member of this family of proteins, translation initiation factor eIF4A, plays a key role in unwinding secondary structure in the 5'UTRs of mRNAs to make start codons accessible for ribosome scanning (Parsyan et al., 2011). The presence of RNA helicases in RNA granules is likely to facilitate RNP remodeling. Not only can the helicases unwind RNA and potentially dislodge bound proteins, but it might chaperone folding to promote formation of structures that cater to different sets of factors (Fuller-Pace, 2006).

Among the different types of granules, RNA helicases are best characterized in germ granules. Posterior group gene *vasa* was found to encode an RNA helicase with sequence similarity with eIF4A (Hay et al., 1988). *Vasa* is recruited by *oskar* to the posterior pole of the fly embryo and is required for the formation of germ granules (Ephrussi and Lehmann, 1992; Schupbach and Wieschauss, 1986). In addition, microinjected *oskar* mRNA is able to recruit *vasa* to the anterior compartment of the fly embryo and induce ectopic pole cell formation (Ephrussi and Lehmann, 1992). P granules in *C. elegans* also contain four *vasa* homologs called germline helicases (GLH1-4; Updike and Strome, 2010). Apart from their function as proper RNA helicases, it is interesting to note that GLH-1, GLH-2, and GLH-4 have phenylalanine-glycine (FG) repeats that are similar to a subset of nucleoporins that make up the nuclear pore complex (NPC) (Updike et al., 2011). FG repeats are regions of intrinsic disorder that coalesce in a meshwork of hydrophobic interactions involving phenylalanine contacts that can enter into a hydrogel-like phase *in vitro* and are thought to constitute the permeation barrier in the channel of the NPC (Denning et al., 2003; Frey et al., 2006; Frey and Görlich, 2007).

Various DEAD-box helicases have been observed to localize to stress granules and P-bodies (Buchan and Parker, 2009). DEAD-box proteins were well-represented in the 2006 proteomic study of neuronal granules by Elvira et al. who identified at least 7 representative members of this family. DDX3, also

called *belle* in *Drosophila*, is a DEAD-box helicase identified in all four classes of cytoplasmic RNA granules (Buchan and Parker, 2009; Updike and Strome, 2011; Elvira et al., 2006; Kanai et al., 2004). In yeast, the DDX3 homolog Ded1 interacts with eIF4G, an adaptor molecule in the eIF4 translation initiation complex that accumulates in stress granules (Hilliker et al., 2011). Overexpression of DDX3 in human cells inhibits cap-dependent translation, presumably through the disruption of the interaction between eIF4E and eIF4G (Shih et al., 2008). Like eIF4A, DDX3 helicase activity plays a key role in unwinding structural elements in the 5'UTR of mRNAs to facilitate processive ribosome scanning (Lai et al., 2008). Interestingly, DDX3 and several other helicases have been implicated in antiviral defense, potentially as sensors for viral nucleic acids (Schröder et al., 2011). Therefore, DDX3 appears to function as both a translational repressor in stalling eIF4-containing initiation complexes as well as translational activator in mediating ATP-dependent unwinding of long or structured 5' UTRs.

Germ granules contain homologs of DDX6/p54/RCK (called Me31B in flies, CGH-1 in worms), which is also a component of P-bodies as well as neuronal transport granules (Thomson et al., 2008; Updike and Strome, 2010; Buchan and Parker, 2009; Elvira et al., 2006). DDX6/p54/RCK was originally identified as a proto-oncogene and is disrupted in chromosomal translocations that cause various hematopoietic cancers (Akao et al., 1995). The *Xenopus* homolog

of DDX6, Xp54, interacts with 5' cap binding protein eIF4E and represses translation of transcripts in the oocyte (Minshall and Standart, 2004). FRET experiments indicate that DDX6 also interacts with eIF4E in P-bodies in human cells (Andrei et al., 2005). In yeast, the DDX6 homolog Dhh1p interacts with decapping and deadenylase factors and is required for efficient mRNA degradation (Coller et al., 2001). Furthermore, DDX6 has also been shown to interact with human Argonaute proteins in P-bodies and is required for Ago2 recruitment to P-bodies (Chu and Rana, 2006), thus it may also play a role in miRNA-mediated repression, a granule function that will discussed in more detail later in this chapter.

Sequence/structure-specific RNA-binding proteins

With the exception of 5'cap binding proteins, poly-A^+ tail binding proteins, and proteins like Y-box proteins that bind along the length of mRNAs in a sequence-independent manner, the majority of RNA-binding proteins target specific sequences or secondary structure elements in the 3'UTRs of mRNAs (Moore, 2005; Andressi and Riccio, 2009). One such protein is *staufen*, a double-stranded RNA-binding protein that is a translational regulator conserved from *Drosophila* to humans (St Johnston et al., 1991; Marión et al., 1999; Wickham et al., 1999). In *Drosophila*, *staufen* is a maternal factor required for establishing the anterior-posterior axis of the embryo (St Johnston et al., 1991). *Staufen* is required for the

localization of *bicoid* mRNA to the anterior pole and this event requires elements present in the 3'UTR of the bicoid message that features extensive predicted stem loop structures (Driever et al., 1990; Ferrandon et al., 1994). Posterior localization and translational repression of *oskar* mRNA is also dependent on *staufen*, but direct physical interaction has not yet been demonstrated (Kim-Ha et al., 1991; Breitwieser et al., 1996; Ferrandon et al., 1994). *Staufen* is also required for the asymmetrical localization of the *prospero* mRNA in mitotic *Drosophila* neuroblasts, which leads to the preferential segregation of the transcript into daughter cells (Broadus et al., 1998).

The mammalian homologs of Staufen (Staufen 1 and Staufen 2) are enriched in the somatodendritic domain of cultured hippocampal neurons and co-localizes with RNA-containing granules in distal dendrites (Kiebler et al., 1999). Furthermore, mammalian Staufen co-sediments with polysomes in sucrose gradients and associates with rough endoplasmic reticulum (ER) in human cancer cells (Marión et al., 1999; Wickham et al., 1999). Interestingly, immunoelectron microscopy shows Staufen to be enriched in tubular structures near synaptic contacts that resemble rough ER in morphology and protein composition (Kiebler et al., 1999; Pierce et al., 2000).

Xenopus Staufen protein is involved in localizing Vg1 RNA, which is one of the earliest asymmetrically localized RNAs to be identified, and associates with a kinesin motor protein SUK4 (Yoon and Mowry, 2004). Ferrandon et al., 1994

demonstrated that *bicoid* 3'UTR microinjected into fly embryos were incorporated in Staufen-containing particles that moved in a microtubule-dependent manner. TAP-tagged human Staufen protein also co-precipitates with tubulin and kinesins as well as other RNA-binding translation-associated factors DHX9/RNA helicase A, hnRNPU, PABP, FMRP, and ribosomal subunits (Villacé et al., 2004). Köhrmann et al., 1999 demonstrated that Staufen-containing neuronal granules originate in the cell bodies of neurons and move out into dendritic processes in a microtubule-dependent manner. Partially purified Staufen-containing RNA granules from rat brain were found to contain known dendritic mRNA Camk2a and a small noncoding RNA BC1 which is thought to translationally repress transcripts through interaction with another granule protein FMRP (Mallardo et al., 2003; Zalfa et al., 2003). Finally, mammalian Staufen is recruited to stress granules upon induction by arsenite treatment and modulates granule formation through its N-terminal polysome interacting domain (Thomas et al., 2008).

Messenger RNAs and noncoding RNAs

RNAs, in particular protein-encoding messenger RNAs, are the objects of spatial and translational regulation for all RNA granules discussed in this chapter, and thus constitute the fundamental unifying component of RNA granules (Anderson and Kedersha, 2006). A classic example of granule regulation of mRNAs is the

localization and translational regulation of *nanos* mRNA by polar granules in the Drosophila embryo, an essential developmental event for the patterning of abdominal structures (Ephrussi et al., 1991; Gavis et al., 1992). Posterior localization of *nanos* mRNA is dependent on multiple regions in its 3'UTR (Gavis et al., 1996). Unlike DNA, single stranded RNA has the capacity to adopt elaborate secondary structures, which are likely to be important for large and complex localization elements such as the *nanos* 3'UTR. Furthermore, *trans*-acting factors regulating RNAs through their 3'UTRs possibly do this because this region is allowed to fold without risking interference with translation. Establishment of the *nanos* gradient is in turn dependent on the localization of *oskar* mRNA to the posterior pole, which requires the first 242 nucleotides of the 3'UTR of *oskar* transcript (Ephrussi and Lehmann, 1992; Kim-Ha et al., 1993). Replacing this 3'UTR with the 3'UTR of anterior determinant *bicoid* causes *oskar* to mislocalize to the anterior pole and is sufficient to recruit germ granule components *nanos* and *vasa* and induce ectopic germ cell formation in the anterior embryo (Ephrussi and Lehmann, 1992).

Extensive work has been done to identify the RNAs regulated by neuronal transport granules. Early identification of dendritically targeted RNAs employed subcellular fractionation methods to prepare synaptodendrosomes that essentially pinched off synaptic material along with associated fragments of dendrites that included polyribosomes and mRNAs (Rao and Steward, 1991). By incubating

synaptodendrosomes with ^{35}S-labelled methionine, one could label newly synthesized proteins encoded by mRNAs localized to dendrites. RNAs identified in this manner include Camk2a, FMRP, and Arc (Sheetz et al., 2000; Weiler et al., 1997; Zalfa et al., 2003). Several dendritic mRNAs such as MAP2, Shank1, Camk2a, InsP3R1, Dendrin, and Arc have been able to be validated by *in situ* hybridization, the gold standard in RNA localization assays (Garner et al., 1988; Falley et al., 2009; Burgin et al., 1990; Furuichi et al., 1993; Herb et al., 1997; Link et al., 1995; Lyford et al., 1995). However, many candidate RNAs remain unverified by *in situ* hybridization likely due to low abundance. For example, FMRP cannot be detected by *in situ* in dendritic processes (Hinds et al., 1993; Valentine et al., 2000), but is repeatedly seen in studies using amplification techniques (e.g. Bagni et al., 2000; Zalfa et al., 2003). Miyashiro et al., 1994 used patch pipettes to sample the cytoplasmic contents of individual neurites of cultured hippocampal neurons and identified them through mRNA expression profiling. This approach results in a large number of candidate RNAs that have not or are unable to be verified by *in situ* hybridization. This raises the possibility that perhaps hundreds of mRNAs are subject to translational regulation by neuronal transport granules, but at levels undetectable by *in situ* hybridization.

An important aspect of studying granule regulated RNAs is identifying the *cis*-acting elements required for RNA localization. More often than not, these sequences reside in the 3' untranslated regions of the mRNAs, although there are

exceptions (Andressi and Riccio, 2009). For example, using a reporter construct containing the coding sequence of lacZ and 8 repeats of the MS2 hairpin loop, Rook, Kosik and colleagues demonstrated that the 3'UTR of the Camk2a mRNA was necessary and sufficient for dendritic localization (Rook et al., 2000). In addition, Miller, Mayford and colleagues generated a mutant mouse where the 3'UTR of Camk2a was mostly replaced by the 3'UTR of bovine growth hormone (Miller et al., 2002). By *in situ* hybridization, they showed that the Camk2a transcript with the mutated 3'UTR was exclusively localized to the cell body and by biochemical analysis, they showed that Camk2a protein levels in the postsynaptic density were reduced. 3'UTR reporter constructs have also been used to confirm 3'UTR elements in Shank1 and Arc mRNAs to be required for localization (Falley et al., 2009; Dynes and Steward, 2007).

Not many studies have addressed the sequence specificity of mRNA regulation by P-bodies other than that miRNAs might target certain messages to P-bodies for repression (Liu et al., 2005). Both electron microscopy and *in situ* hybridization experiments show that only a subset of poly-A^+ RNAs are recruited to stress granules upon induction, although few studies address the nature of these RNAs other than to speculate that RNAs encoding "housekeeping" genes are generally spared repression for the duration of the stress (Nover et al., 1989). Furthermore, as stress granules are hypothesized to assemble around stalled 48S preinitiation complexes (Kedersha et al., 2002), it is likely that the mechanism by

which some mRNAs escape repression is a generalized one and not sequence-specific.

Lastly, RNA granule function is also associated with regulation of small noncoding RNAs such as repeat-associated small interfering RNAs (rasiRNAs) that are proposed to protect the germline from selfish genetic elements, microRNAs that mediate sequence-specific translational repression in conjunction with the RISC complex, and BC1 which is a small noncoding transcript that mediates transcript silencing in neuronal granules (Vagin et al., 2006; Megosh et al., 2006; Zalfa et al., 2003). RasiRNAs require Ago proteins piwi and aubergine which are both constituents of *Drosophila* polar granules (Vagin et al., 2006; Megosh et al., 2006; Harris et al., 2001). Ago1 and Ago2 are highly concentrated in P-bodies which also contain other components relating to miRNA-mediated repression (Sen et al., 2005; Liu et al., 2005; Eyastathiony et al., 2002, 2003). Using a reporter RNA that encoded luciferase with a let-7 miRNA binding site and 12x MS2 hairpins in the 3'UTR, Liu et al., 2005 demonstrated that mRNAs could be targeted to P-bodies in a miRNA dependent manner. Suffice to say, small RNA studies in this field will likely yield rewarding findings in the future with interesting implications for RNA granule function.

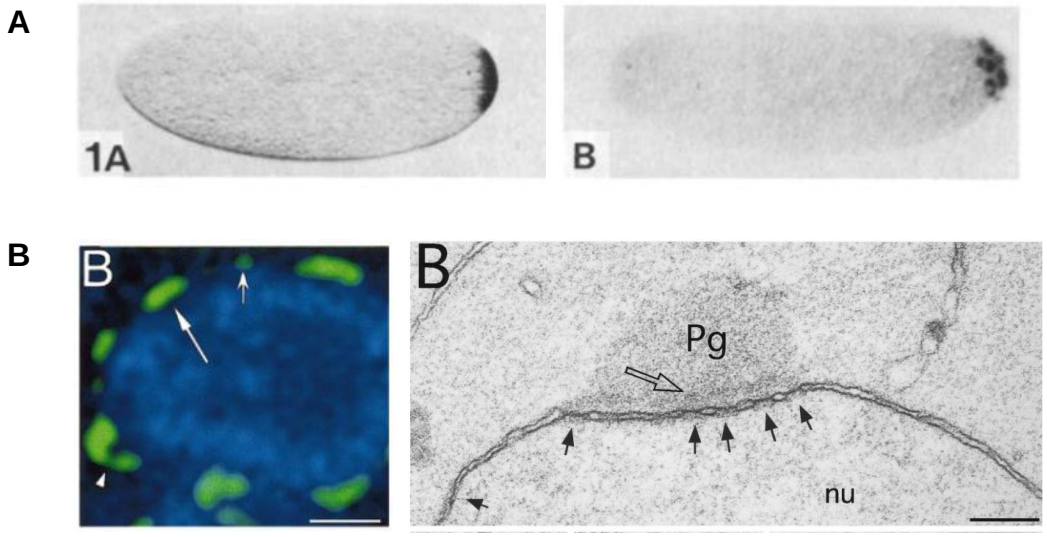


Figure 1-1. Germ granules in *Drosophila melanogaster and Caenorhabditis elegans* embryos.

(A) *Drosophila* embryos are stained with antiserum specific to Vasa, an ATP-dependent RNA helicase component of polar granules. Vasa preferentially segregates to the posterior of the early embryo (right) and partitions asymmetrically to the budding pole cells (left), precursor germline cells. Figure modified from Hay et al. (1988a).

(B) Germ granules in developing germ cells in adult *C. elegans* are visualized by GFP-PGL1 (left), a constitutive marker of P granules. P granule localization is distinctly perinuclear as indicated by the white arrows pointing to the green aggregates dotting the rim of the nucleus (blue). Immunostaining electron micrograph (right) shows P granules in close association with a cluster of five nuclear pore complexes (indicated by black arrows). Figure modified from Pitt et al. (2000).

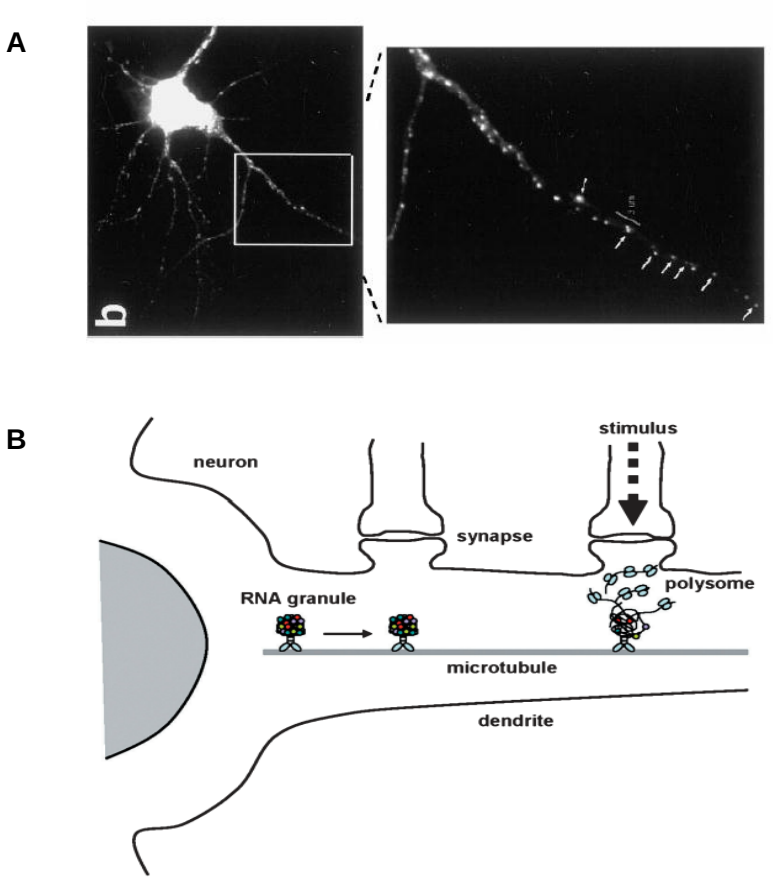


Figure 1-2. Neuronal transport granules and model of local dendritic translation.

(A) Neuronal transport granules visualized by the MS2-GFP system. The 3'UTR of Camk2a mRNA is tagged with MS2 RNA hairpins and co-transfected transiently into primary hippocampal neuronal culture with a construct that encodes MS2 RNA-binding protein fused to GFP. Fluorescence localizes to puncta in the neuronal processes, indicating that the Camk2a 3'UTR reporter is dendritically targeted in a manner similar to the endogenous transcript. Figure modified from Rook et al. (2000).

(B) Simplified model of localized translation at the synapse. Nascent RNAs are transcribed in the nucleus and packaged into translationally repressive granules that facilitate active transport along microtubules into dendritic processes. Once an activating signal is received, the granule releases silenced transcripts for translation. Figure from Miki et al. (2005).

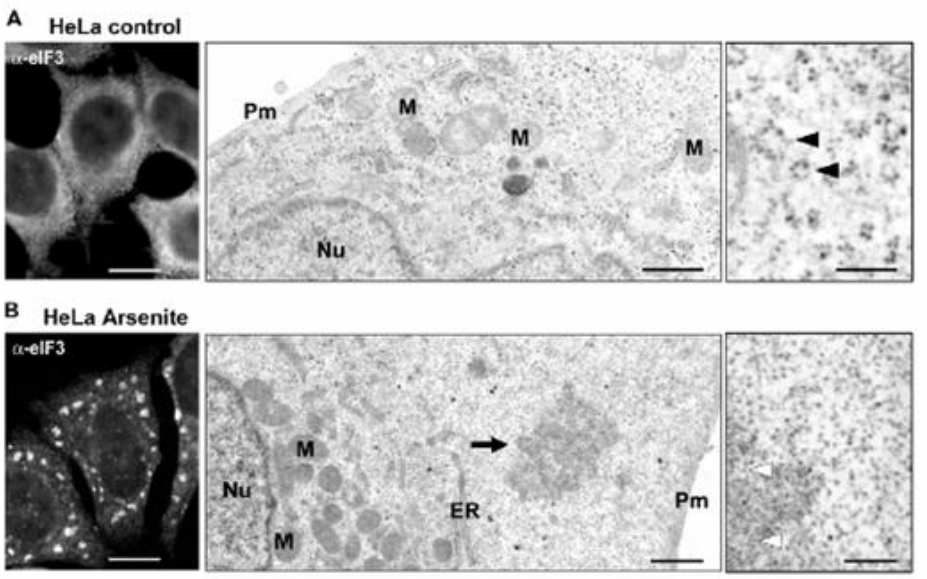


Figure 1-3. Stress granules induced by arsenite in HeLa cells.

(A) HeLa cells are stained with antiserum specific to eIF3, a translation initiation factor that is a component of the 48S preinitiation complex. In the unstressed cell, eIF3 staining is diffuse throughout the cytoplasm (left) and ribosomes can be seen by electron microscopy diffuse throughout the cytoplasm of the cell (right)

(B) HeLa cells are treated with arsenite to induce stress granule formation (left). Upon induction with arsenite, eIF3 is recruited to bright cytoplasmic foci that correspond to stress granules. Using electron microscopy, stress granules are observed as electron dense aggregates larger than mitochondria and enriched with ribosomes (right).

Figure from Souquere et al. (2009).

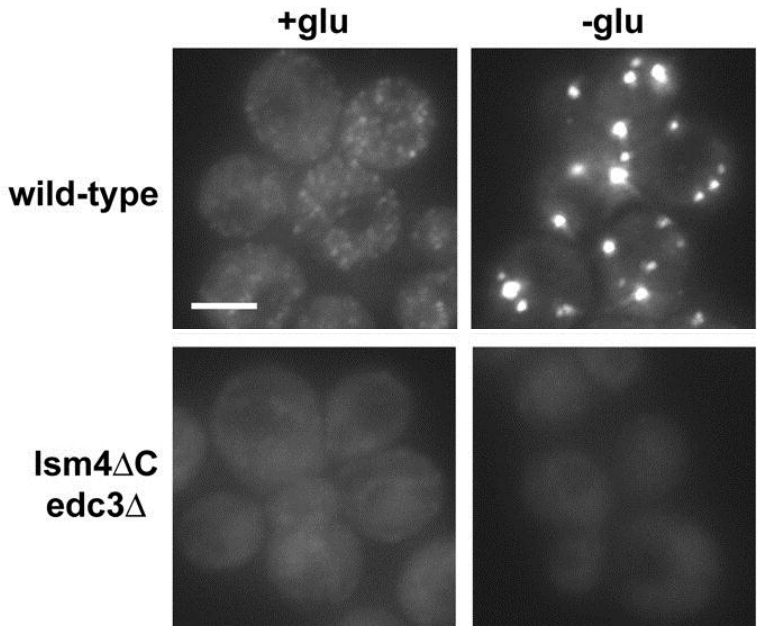


Figure 1-4. P-bodies are cytoplasmic foci that contain basic machinery for mRNA degradation.

P-bodies are visualized by Dcp2, an mRNA decapping factor, fused to GFP, in wild type yeast and lsm4ΔC edc3Δ yeast mutants. P-bodies can be detected in glucose replete conditions, but coalesce into larger and brighter foci upon glucose deprivation. Lsm4 and Edc3 harbor "prion-like" aggregation domains that are required for P-body assembly and mutants lacking Lsm4 and Edc3 are unable to form visible P-bodies even under glucose deprivation.

Figure modified from Decker et al. (2007).

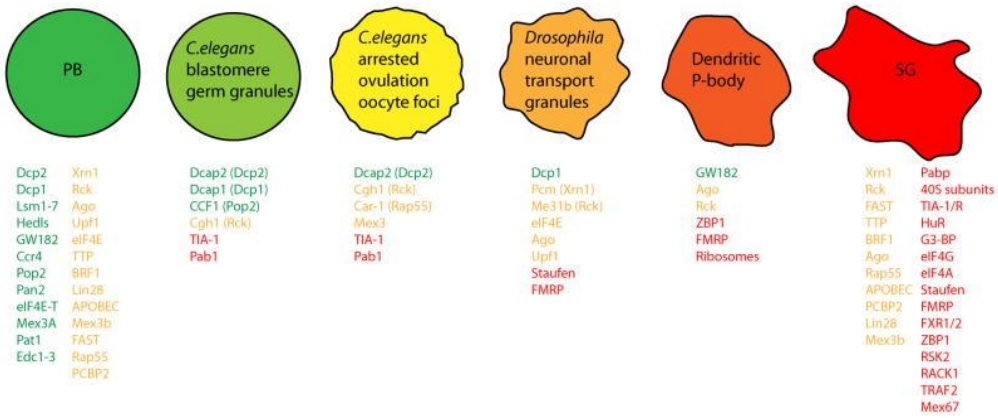


Figure 1-5. A continuum of RNA granules

Cytoplasmic RNA granules share many protein components. In this cartoon diagram from Buchan and Parker (2009), P-body proteins are color green (left), stress granule proteins are colored in red (right), and proteins common to more than one granule are color coded in yellow. Several of these components are also found in other classes of RNA granules including germ granules from *C. elegans* and neuronal transport granules from *Drosophila*. This figure illustrates both the similarities among various types of granules and the inherent heterogeneity of RNA granule composition.

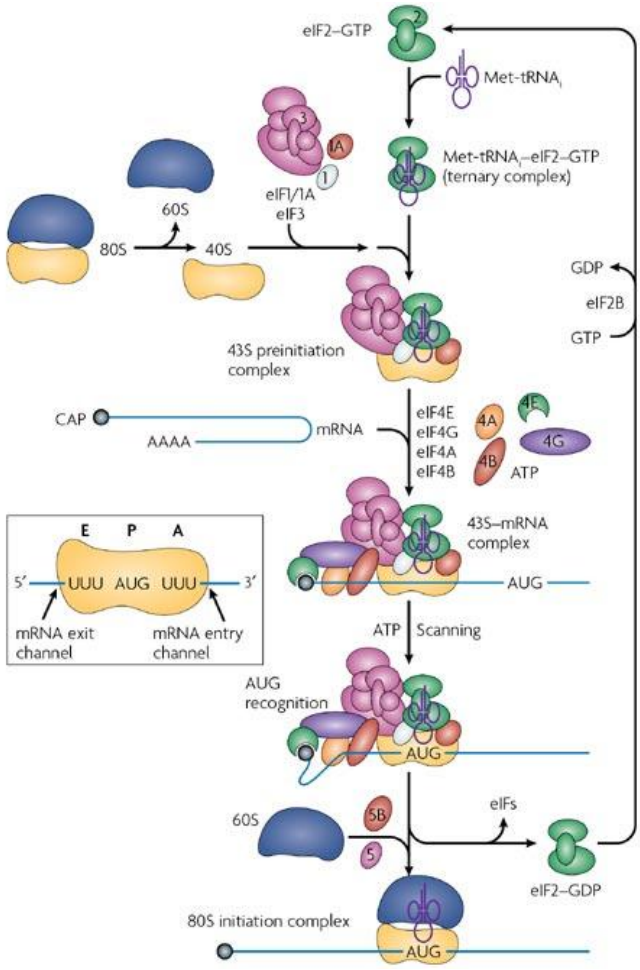


Figure 1-6. Cap-dependent translation initiation

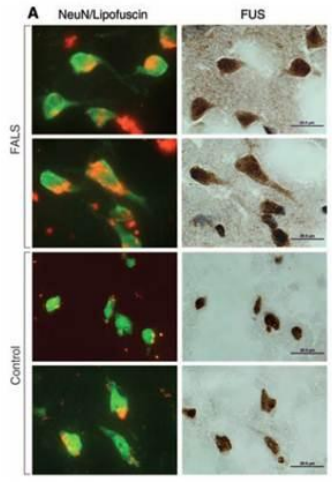
Transcription initiation begins with the disassembly of the 80S ribosome into 60S and 40S subunits. Scaffolding factors (eIF1 and eIF3) bind to the 40S small subunit and assemble with the eIF2-tRNA^{Met}-GTP ternary complex to form the 43S preinitiation complex. The eIF4 complex recruits mRNA to the preinitiation complex through the 5'cap and interactions with poly-A binding protein, forming the 48S preinitiation complex. Once the 48S complex accesses a viable start codon, a GTPase-activating protein eIF5 promotes GTP hydrolysis by eIF2 which releases the initiation factors from the 40S subunit, allowing the 60S subunit to bind and processive ribosome scanning for start codons to commence.

Figure from Fraser and Doudna (2007).

A


R522Q
R521C
R521H
G515C
R518K
R524S
R514S
H517Q
R521G
R524T
P525L
R495X
495- R G G R G G G D R G G F G P G K M D S R G E H R Q D R R E R P Y -526
QGSY-rich
Gly-rich
RRM
ZF
RG-rich

B


A
NeuN/Lipofuscin
FUS
FALS
Control

C

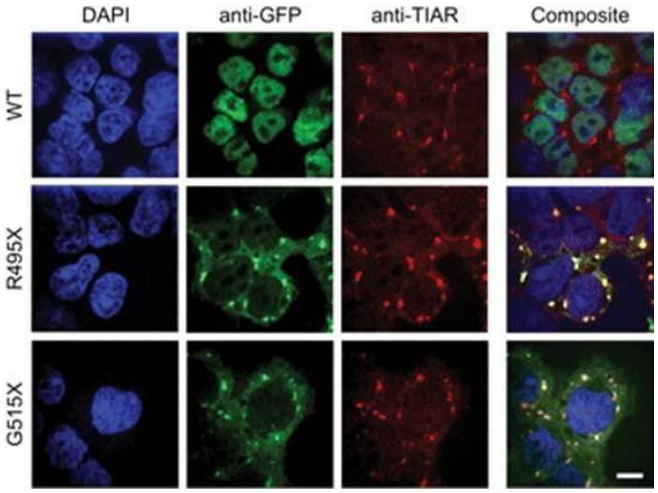

DAPI
anti-GFP
anti-TIAR
Composite
WT
R495X
G515X

Figure 1-7. Mutations in FUS cause amyotrophic lateral sclerosis (ALS)

(A) In this figure from Bosco et al. (2010), a subset of ALS mutations are represented on a schematic diagram of the FUS protein. These mutations are notably all either in the C-terminal nuclear localization signal (NLS) or, as in the case of R495X, causes a premature termination upstream of the NLS.

(B) Immunohistochemistry of postmortem tissue (right column) from an ALS patient shows that the FUS protein is mislocalized to the cytoplasm in diseased neurons whereas wild type FUS is localized predominantly to the nucleus. On the left column, lipofuscin is visualized in the same neurons (co-stained with NeuN as a marker for mature neurons). The ALS neurons exhibit increased levels of lipofuscin which is a product of fatty acid oxidation and may be an indication of mitochondrial damage. Figure from Kwiatkowski et al. (2009).

(C) GFP-tagged FUS protein constructs harboring either the R495X or G515X mutations are transiently transfected into tissue culture cells. Both mutations cause the deletion of the C-terminal NLS in the FUS protein. Cells are treated with arsenite to induce stress granule formation and stained with antiserum specific to GFP to visualize the transfected protein and co-stained with antiserum specific for TIAR as a stress granule marker. FUS proteins bearing the NLS mutations are recruited to stress granules upon induction with arsenite whereas the wild type FUS protein is predominantly nuclear in localization. Figure from Bosco et al. (2010).

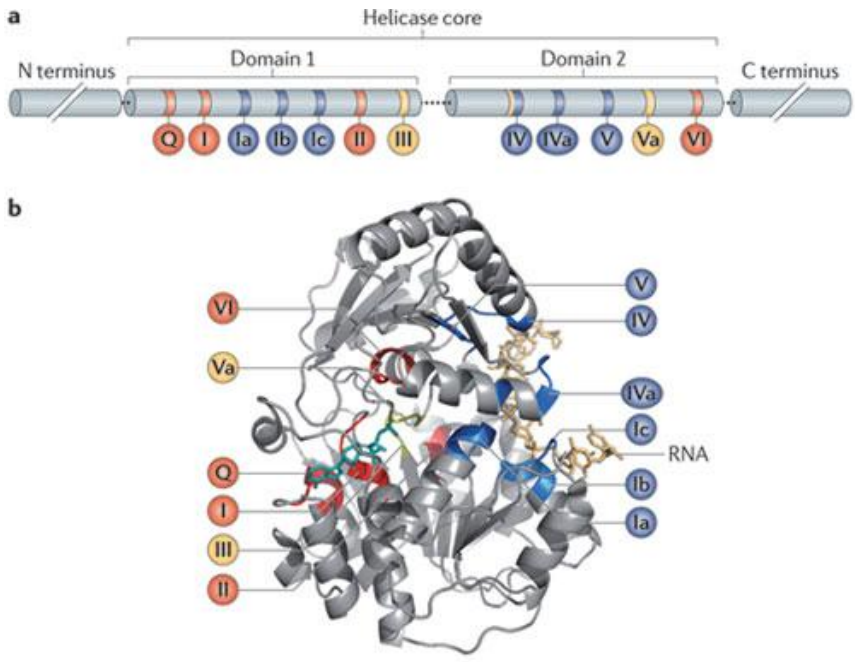


Figure 1-8. Domain architecture and structure of DEAD box helicases.

(A) Domain arrangement of conserved motifs in RNA helicases which are color coded according to primary functions. Red domains are required for ATP binding and hydrolysis, blue for RNA binding, and yellow for communication between the ATP and RNA binding sites. Helicase core of DEAD box helicases are composed of two recombinase A (RecA)-like domains (Domain 1 and Domain 2).

(B) Structure of *Drosophila* DEAD box helicase protein Vasa binding to cognate RNA (in beige). The conserved helicase motifs are color coded as in (A).

Figure modified from Linder and Jankowsky (2011).

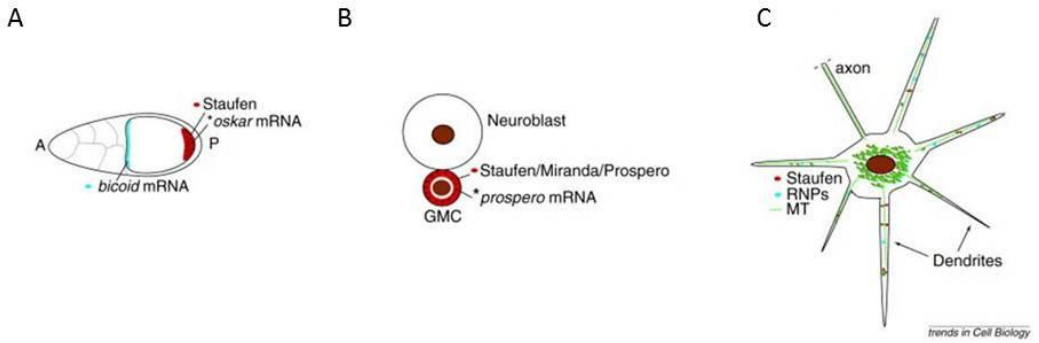


Figure 1-9. Staufen localization in *Drosophila* oocytes, neuroblasts, and mammalian neurons.

(A) In the *Drosophila* embryo, Staufen protein is required for the localization of *oskar* mRNA to the posterior compartment.

(B) During neuroblast division, Staufen protein is required for the preferential segregation of *prospero* mRNA to the GMG daughter cell.

(C) In mammalian neurons, Staufen protein is a component of neuronal granules that transport specialized RNAs in a microtubule-dependent manner into dendritic processes for local synthesis.

Figure modified from Roegiers and Jan (2000).

CHAPTER 2:

Cell-free Formation of RNA Granules

Introduction

Asymmetric localization of mRNAs is a mechanism cells use to restrict protein synthesis to specific subcellular compartments. This mechanism of RNA regulation is important for diverse biological processes and cellular states such as segregation of cell fate determinants, cell polarity, and synaptic plasticity. For example, uneven distribution of maternal RNAs in *Drosophila* oocytes allows for the asymmetric distribution of cytoplasmic determinants into germline precursor cells and the establishment of morphogen gradients required for axial patterning (Palacios and St. Johnston, 2001). Also, accumulation of β-actin mRNA in the leading edges of chick embryonic fibroblasts directs the site of active actin polymerization to drive cell migration (Kislauskis et al., 1993; Kislauskis et al., 1994). β-actin mRNA is also transported to the tips of growth cones in developing axons seeking synaptic connections (Zheng et al., 2001). Following treatment with antisense oligonucleotides directed towards the β-actin 3'UTR, neurons experienced reduced motility and, in some cases, retraction of growth cones protrusions (Zheng et al., 2001). Lastly, active transport of specialized RNAs to neuronal processes for localized translation in or near synapses is implicated in modulation of synaptic plasticity and establishment of synaptic

memories. One of the best characterized dendritically targeted mRNA in mammals is the transcript that encodes the alpha subunit of calcium/calmodulin-dependent protein kinase 2 or Camk2a (Burgin et al., 1990). In mice where the 3'UTR of Camk2a was altered to retain the transcript in the cell soma, there are significant reductions in late-phase long-term potentiation in hippocampal field recordings and deficits in hippocampal-dependent spatial memory tasks (Miller et al., 2002). Although exemplified in cells where RNAs may play highly specialized roles within discrete subcellular compartments, over 70% of genes in the developing Drosophila embryo were found to encode mRNAs with distinct spatial patterns (Lécuyer et al. 2007), suggesting that asymmetric mRNA localization is a much more widespread mechanism of translational control than previously imagined.

In many instances, RNA transcripts are packaged into ribonucleoparticle (RNP) aggregates called RNA granules for transport to specific sites in the cytoplasm and for translational regulation. RNA granules are large, electron-dense, non-membrane bound cellular structures rich in RNA that have been observed in diverse biological contexts (e.g. Mahowald, 1962; Knowles et al., 1996; Souquere et al., 2009). In the early 1900s, American zoologist Robert Hegner observed RNA granules enriched in the posterior pole of chrysomelid beetle eggs (Hegner, 1908; Hegner, 1911). Ablation of posterior pole plasm with a hot needle subsequently abolished germline development in the beetle embryo

(Hegner, 1908; Hegner, 1911). Granule-containing plasm was further established as germline determinants with transplantation experiments in the early 1970s that demonstrated rescue of germ cell formation in UV-irradiated fly embryos (Okada et al., 1974; Illmensee and Mahowald, 1974). Furthermore, purified poly-A^+ RNA isolated from early stage embryos was able to rescue germline development in UV-irradiated embryos (Okada and Kobayashi, 1987) suggesting that RNAs localized to the germ plasm are likely germ cell determinants.

Oskar mRNA is localized to the *Drosophila* germ granules and is required for nucleating the assembly of the germ plasm that contains both posterior and germ cell determinants (Ephrussi et al., 1991; Ephrussi and Lehmann, 1992). Localization of *oskar* is dependent on *cis*-acting elements residing in the 3' untranslated region of the transcript and the double-stranded RNA-binding activity of the Staufen protein (Kim-Ha et al., 1993; Micklem et al., 2000). Staufen is also required for the asymmetrical localization of the *prospero* mRNA in mitotic *Drosophila* neuroblasts, which leads to the preferential segregation of the transcript into daughter cells (Broadus et al., 1998). The mammalian homologs of Staufen are implicated in RNA transport in neurons (Kiebler et al., 1999; Köhrmann et al., 1999). Time lapse video microscopy shows Staufen-containing granules shuttling from the neuronal cell body into dendritic processes in a microtubule-dependent manner (Köhrmann et al., 1999). When depleted of Staufen protein either genetically or by RNAi, cultured hippocampal neurons

exhibit a reduction in the number of dendritic spines and abnormal spine morphology (Goetze et al., 2006; Vessey et al., 2008). Partially purified Staufen-containing neuronal granules were found to contain known dendritic RNAs Camk2a, a postsynaptic density component, and BC1, a small noncoding RNA that targets specific mRNA for translational repression (Mallardo et al., 2003; Zalfa et al., 2003).

It has been difficult, however, to study RNA granules *in vitro*. They are highly heterogeneous in composition, weakly associating aggregates, complicated to work with biochemically, and practically impossible to purify. The lack of a technique to easily biochemically isolate RNA granules for *in vitro* studies constitutes a major impediment in gaining a more detailed understanding of the mechanisms governing RNA granule function. Here we describe two *in vitro* models for studying RNA granule assembly. The first model was based on the observation that exposure of soluble, cytoplasmic extracts from a variety of cell and tissue types to a biotinylated isoxazole compound (isx) resulted in the selective and reversible aggregation of hundreds of RNA-binding proteins. In addition to RNA-binding domains, we found many of these proteins featured low complexity sequences that were both necessary and sufficient for precipitation by the biotinylated isoxazole. The second model was based on the observation that the N-terminal low complexity domain of RNA-binding protein FUS, a

representative precipitated protein, undergoes a concentration-dependent phase transition from a soluble state to a hydrogel-like state.

Cell-free formation of RNA granules: *trans*-acting factors

The predecessor of this biotinylated isoxazole, designated 5-aryl-isoxazole-3-carboxyamide, was identified in a high throughput chemical screen for activators of homeodomain transcription factor Nkx2.5, a hallmark of cardiac differentiation, in P19 embryonal carcinoma cells (Sadek et al. 2008). A subset of derivatives were subsequently found to harbor neurogenic rather than cardiogenic activity, inducing upregulation of neuronal genes such as neuron-specific class III beta-tubulin Tuj1 in both hippocampal neural stem cell line HCN as well as astrocytic cell line SS05 (Schneider et al. 2008; Zhang et al., 2011). The most potent of the neurogenic isoxazoles, N-cyclopropyl-5-(thiophen-2-yl)-isoxazole-3-carboxamide, was then biotinylated with a 3-carbon linker in order to identify molecular targets within the cell that mediated its neurogenic activity. The initial approach of Dr. Shanhai Xie, a postdoctoral fellow in the McKnight Lab, was to incubate the biotinylated compound with whole cell lysate and use affinity chromatography with streptavidin-coupled agarose beads to pull down any interacting proteins. Disappointingly, he recovered an overwhelming number of potential candidates which we assumed to be non-specific interactions (Fig 2-1A; lanes 5-7). Upon further investigation, it was revealed that these proteins were

precipitating from lysate with or without the addition of streptavidin beads. In other words, the biotinylated isoxazole compound alone mediated the protein precipitation. However, two surprising observations were in our favor. First, visual inspection of the bands in the supernatant and pellet fractions indicated that the precipitated proteins were not the most abundant proteins in the cell. Whereas the supernatant proteins (Fig 2-1A, lane 4) looked virtually identical to the lysate proteins (lane 1), the pellet proteins were distinctly different and therefore represented a minor fraction of the lysate proteins. Second, the precipitation reaction was notably reversible. When the precipitated proteins were resuspended in buffer and warmed up to 37°C, the proteins went back into solution (Fig 2-1B, lane 8). In order to further understand this precipitation activity of the biotinylated isoxazole, Shanhai used the biotinylated isoxazole to precipitate from lysates from four mouse tissues and cell lines (brain, testes, NIH 3T3 cells, and ES cells), excised the protein bands and submitted them for mass spectrometry identification.

Core isoxazole-precipitated proteins are RNA-related

Mass spectrometry analysis of proteins precipitated by the biotinylated isoxazole from four different mouse tissue and cell types revealed significant overlap. 162 proteins were found to be common to all datasets, which we took to constitute a core subset of precipitated proteins (Fig 2-2A). Upon analysis of predicted

protein functions, I found a substantial representation of RNA-binding proteins in the list of core precipitated proteins. 33% of the core proteins were ribosomal subunits (both large and small), 4% were select translation initiation factors belonging to the 48S preinitiation complex (eIF2, eIF3, and eIF4), 23% were hnRNPs (primarily involved in the processing of newly synthesized RNAs in the nucleus), 8% were DEAD-box containing RNA helicases, and 32% were other proteins with RRMs or KH RNA-binding domains (Fig 2-2B; Table 1). Several features of the core precipitation led me to believe that this list of proteins highly resembled components of RNA granule proteins that have been reported in the literature. First, there was a substantial representation of ribosomal subunits and translation initiation factors in our dataset. Stress granules are thought to assemble on translationally stalled 48S initiation complex that includes the 40S ribosomal subunit, eIF3, eIF2 that primes the complex with initiator $tRNA^{Met}$, and the eIF4 protein complex that recruits the ribosome to the mRNA through binding the 5' cap and indirectly the poly-A+ tail (Kedersha et al., 2002). Second, DEAD-box RNA helicases, which unwind RNA secondary structure in an ATP-dependent manner, are common features in all types of granules (Arkov and Ramos 2010). Third, hnRNP proteins that accompany mRNAs out into the cytoplasm have been implicated as *trans*-acting factors that regulate the localization, stability, and translation of specific RNAs (Dreyfuss et al., 2002). Lastly, I identified proteins implicated in neurodegenerative disorders and mental

retardation such as ataxin 2, TDP-43 and FMRP, which are documented components of stress granules and neuronal transport granules (Nonhoff et al., 2007; Liu-Yesucevitz et al., 2010; Barbee et al., 2006). Statistical analysis comparing the core isoxazole precipitated proteins (Table 1) with known RNA granule proteins compiled from literature references (Table 2) gave a p-value of < 0.0001 which strongly suggested that the resemblance was not random. We next confirmed by Western blot the presence of RNA granule proteins in the isoxazole precipitates (Fig 2-3). Figure 2-3A shows the isx-mediated precipitation of the FUS protein, which is an RNA-binding protein found mutated in ALS patients (Kwiatkowski et al. 2009, Vance et al. 2009), from four different mouse tissue and cell types. The bottom panel in Figure 2-3A confirms the precipitation of additional RNA-binding proteins EWS, ATXN2, and FXR1 from mouse brain lysate. Precipitation occurs in a dose dependent manner, and at 100 μM concentration, the compound is able to quantitatively deplete the proteins from the lysate. Moreover, isx-mediated precipitation is specific to a certain subset of proteins. Actin, which is a highly abundant protein in the cell, is retained in the supernatant even at the highest concentration of the compound tested (Fig 2-3B, lane 4). There also appears to be a range of sensitivity to the compound among the proteins tested. FUS is among the most responsive proteins to the compound, precipitating at lower concentrations than other proteins. Most strikingly, the precipitation of FUS and TIA1 is reversible by heat (Fig 2-3B, lane 8) and can be

re-precipitated upon cooling to 4°C (lane 9) which suggests that it is likely to be an ordered interaction and therefore distinct from pathological aggregation which is thought to be irreversible.

Low complexity sequence is necessary and sufficient for isoxazole-mediated precipitation

To determine the protein domains required for precipitation by the biotinylated isoxazole, I constructed eGFP-tagged truncation mutants of representative proteins precipitated by isx and expressed and purified them from bacteria. As shown in Figure 2-4A, the full-length TIA1 protein is robustly precipitated by isoxazole. When 92 amino acids are deleted from the the C-terminus of TIA1, the protein is retained exclusively in the supernatant and excluded from precipitation. This domain can also transfer its activity to recombinant eGFP protein which cannot be precipitated on its own in the background of whole cell lysate. Upon the addition of C-terminal sequence from TIA1, eGFP too can be recruited into the pellet. This domain has also been reported to be necessary for assembly into stress granules (Gilks et al., 2004). This 92 amino acid fragment of the TIA1 protein prominently features a stretch of low complexity sequence with an overrepresentation of glutamine, asparagine, tyrosine, and glycine and would be predicted to be structurally disordered (Dunker et al., 2001). Interestingly, this simple sequence domain can be functionally replaced by the aggregation domain

of yeast prion protein Sup35 to reconstitute stress granule assembly (Gilks et al. 2004). In the same manner, I also demonstrate that the low complexity sequence from a second RNA-binding protein, hnRNP A1, was necessary and sufficient for precipitation by isoxazole (Fig 2-4B). Furthermore, low complexity sequences in several other isoxazole-precipitated proteins such as FUS, Rbm3, Cirbp, and yeast Sup35 were also found to be sufficient for precipitation (Fig 2-4C).

With help from Dr. Jimin Pei at UT Southwestern, we quantified the fraction of low complexity domain containing proteins among the lists of isoxazole precipitated proteins and RNA granule proteins curated from the literature. The SEG program is commonly used to mask repetitive sequence from BLAST alignments to prevent false sequence similarity (Wooten and Federhen, 1993). We set the definition of having low complexity sequence to be the longest length of continuous low complexity sequence greater than 35 amino acids (as determined by SEG). About 40% of isoxazole precipitated proteins and 35% of known granule proteins have low complexity sequence. Enrichment of proteins with low complexity sequence in both lists was statistically significant when compared to the average percentage in the mouse proteome (around 10%) with a p-value of $>2.2x10^{-16}$ (Fig 2-4D). This encouraged me to believe that I was on the right track in terms of finding a link between the precipitation activity of biotinylated isoxazole and the mechanism of RNA granule formation which apparently both involve certain low complexity sequences.

Low complexity region of RNA-binding protein FUS forms hydrogel

While working with concentrated purified protein of the first 214 amino acids of the FUS protein, it was noted that the protein adopted a gel-like state when stored at low temperatures (Fig 2-5A). The LC domain of FUS is characterized by repetitive tyrosine residues interspersed with more hydrophilic residues such as glycines and serines (Fig 2-8A, wild type) and is reminiscent of the FG-repeat domains of certain nucleoporins that also adopt a gel-like state when concentrated and chilled and are hypothesized to constitute a molecular permeability barrier of the nuclear pore (Frey et al. 2006, Frey and Görlich 2007). Using a fluorescence-based gel binding assay, Frey et al. (2006) demonstrate that large macromolecules such as mCherry protein are incapable of penetrating the FG- hydrogel, whereas nuclear transport proteins containing FG-repeats accumulate in the gel and can carry associated cargo proteins with them. Inspired by the FG-hydrogel work, Dr. Masato Kato, assistant professor in the McKnight Lab, designed a hydrogel-based experimental system to investigate domains required for interaction with the FUS gel. For this assay, he exposed preformed mCherry-FUS gel droplets to eGFP-tagged soluble proteins of interest in a fluidic chamber and evaluated the degree of co-localization between the eGFP and mCherry signals (Fig 2-5B). Purified eGFP protein was not retained by the gel and instead remained diffuse throughout the chamber even after days of incubation, whereas eGFP-FUS214 displayed

strong binding to the FUS gel (Fig 2-5C). Furthermore, he observed gel retention of the low complexity sequences of several other RNA-binding proteins (hnRNPA2, TDP-43, Cirbp, CPEB2, FMRP) to varying degrees, demonstrating the sufficiency of these domains for retention by the FUS gel (Fig 2-6). In order to demonstrate that low complexity sequence was also necessary for retention by the FUS gel, I purified recombinant protein from bacteria using the same TIA1 and hnRNPA1 mutants that were used previously in isoxazole precipitation assays (Fig 2-4A, B) and incubated them with the FUS hydrogel. The full-length versions of TIA1 and hnRNPA1are quite readily retained by the FUS gel as are their LC domains alone, but deletion of these sequences causes the protein to remain diffuse throughout the fluidic chamber, similar to eGFP alone (Fig 2-7). Finally, the FUS gel also exhibits strong heterotypic interaction with the aggregation domain of the yeast prion protein Sup35 (Fig 2-6), a characteristic it shares with the nucleoporin FG-hydrogel that is suggestive of amyloid-like interactions (Ader et al. 2010).

Tyrosine residues in FUS low complexity region critical for gel retention and stress granule recruitment

Due to the similarities between the imperfect YG and YS repeats that make up the FUS gelation domain (Fig 2-8A, wild type) and the nucleoporin FG repeats, we wondered whether the FUS gel was also cross-linked by hydrophobic and/or

aromatic interactions between tyrosine residues in the low complexity sequence (Frey et al. 2006). In order to test this hypothesis, we mutated clusters of tyrosine residues in the FUS low complexity domain to serine, and assessed the performance of the mutants in the FUS gel retention assay. Fig 2-8A shows the exact positions of the mutations made in each construct. As shown previously, the eGFP-FUS214 exhibits strong homotypic interactions with the FUS gel, which sometimes manifests as a sharp halo accumulating in the periphery of the mCherry-FUS gel. The S1 mutant harbors 5 tyrosine-to-serine substitutions and is still retained in the gel albeit with decreased avidity since the signal diffused inwards (Fig 2-8B). The S2 mutant with its 9 tyrosine replacements clearly shows weaker gel retention than S1 or wild type constructs. The S3 and allS mutants, which have 15 and all 27 mutated tyrosines respectively, show no accumulation in the hydrogel and instead appear diffuse throughout the fluidic chamber. To demonstrate the biological relevance of these mutations, I constructed Flag-tagged full-length FUS harboring the same tyrosine mutations. Since full-length FUS protein shows predominantly nuclear localization even under conditions that induce stress granules (Fig 2-8D), I deleted the nuclear localization signal in the last 32 amino acids of the C-terminus. This permits robust accumulation of the FUS protein in stress granules upon induction by arsenite (Bosco et al., 2010). These plasmids were transiently transfected into U2OS cells and metabolic stress was induced in these cells with exposure to 0.5 mM sodium arsenite for 1 hour.

Immunostaining for the Flag epitope revealed the induction of large cytoplasmic foci for the wild type FusΔ32 (Fig 2-8C) which co-localizes with stress granule marker TIA1 (Fig 2-8D). The S1 mutant displays similar staining patterns to the WT, while the S2 mutant exhibits smaller but still clearly evident perinuclear aggregation. The S3 and allS, on the other hand, are weakly punctate and mostly diffuse in the cytoplasm. As a control, we verified that cells not transfected with the allS mutants were still able to assemble TIA1-positive granules under stress (Fig 2-8D), thereby excluding the possibility that we failed to induce granule formation.

Cell-free formation of RNA granules: *cis*-acting factors

If the precipitation activity only requires a low complexity peptide sequence, how are proteins completely lacking simple sequence recruited to the pellet fraction? Because these proteins are predominantly RNA-binding proteins, I wondered if RNA was required for the recruitment of proteins to the pellet. To address this question, I employed a method of quantitative mass spectrometry utilizing stable isotope labeled amino acids in cell culture, or SILAC to chart changes in precipitation upon RNA degradation by RNase treatment. For this experiment, two populations of cells were fed with either heavy or light isotope containing lysine and passaged seven times to fully incorporate the labeled amino acids (Fig 2-9A). I treated the heavy lysate with RNase A but added RNase inhibitor to the

light lysate to keep RNAs intact. Both lysates were then subjected to precipitation with biotinylated isoxazole. Figure 2-9B shows the proteins precipitated from the lysates, with and without RNase treatment, and ethidium bromide staining corresponding to the level of intact RNA in the sample. Next, I mixed the heavy samples with the light samples 1:1 and gel purified the mixed samples. With help from Dr. Hamid Mizraei and the UTSW mass spectrometry core, we successfully identified 81 proteins that required intact RNA to precipitate; in other words, these proteins were under-represented in the pellet after RNase treatment. Figure 2-9C shows five representative proteins that do not change in precipitation irrespective of treatment and five proteins that exhibit reduced precipitation as a function of RNA degradation. 62% of the proteins that were underrepresented in the pellet were identified as ribosomal. Furthermore, Seg analysis revealed a statistically significant enrichment of LCS-containing proteins among the proteins overrepresented or non-changing in the precipitant following RNase treatment when compared to control (Fig 2-9D). A straightforward interpretation of this data is that active ribosomes assembled onto translating mRNAs would be indirectly recruited to the isoxazole pellet if LC sequence containing RNA-binding proteins were also bound to the same mRNAs. Following degradation of RNAs in the lysate, ribosomes are no longer in association with LCS-containing RNA-binding proteins, and thus are not recruited to the isoxazole pellet (Fig 2-9E).

Isoxazole-mediated precipitation of RNA granule mRNAs

At this point, we suspected that the precipitation activity exhibited by the biotinylated isoxazole was assembling RNA granules *in vitro* or at least ribonucleoparticles that highly resembled them. We have so far established that the isoxazole precipitant was significantly enriched with RNA granule proteins and that the same low complexity domains required for precipitation were also required for hydrogel retention. The next question we asked was: does the biotinylated isoxazole precipitate granule-regulated RNAs? Any RNAs resident in the pellet would necessarily have been recruited by proteins since the compound was found incapable of precipitating purified total RNA (Fig 2-10D), and since isoxazole pulls down granule proteins, it would be straightforward to expect the precipitation of granule mRNAs. We purified total RNA and enriched for poly-A^+ RNA from both the pellet and supernatant fractions following isoxazole-mediated precipitation from both mouse brain lysate and U2OS cell lysate (Fig 2-10A). Using RNA-Seq, we identified mRNAs both enriched and excluded from the isoxazole precipitant from these two sources. Among the mouse brain pellet RNAs, we observed an enrichment of well-characterized known dendritic mRNAs (Fig 2-10B, top) that included transcripts encoding postsynaptic density proteins such as Shank1/3, Arc, and PSD95 (Böckers et al., 2004; Link et al., 1995; Lyford et al., 1995; Muddashetty et al., 2007);

neurotransmitter receptors subunits like GluR1 and GluR2 (Grooms et al., 2006); voltage-gated potassium channel Kv1.1 (Wang et al., 1994); Camk2a, a kinase required for hippocampal long term potentiation (Burgin et al., 1990); microtubule-stabilizing protein MAP2 (Caceres et al., 1983). Five representative mRNAs from pellet enriched and supernatant retained categories were subsequently verified by qPCR (Fig 2-10C, top). We then cross-referenced the mRNAs sequenced from the mouse brain precipitant with the U2OS dataset. Not all of the mRNAs sequenced from the mouse brain experiment were found to be expressed in the human cancer cell line, which can be attributed to the neuronal cells being specialized for unique tasks. Among the overlapping genes however, we found that 95% of mRNAs enriched in the mouse brain pellet were also enriched in the U2OS pellet and 80% of the mRNAs enriched in the mouse brain supernatant were also enriched in the U2OS supernatant. This is not entirely unexpected since we observe similar proteins precipitating from both tissue types, therefore we should expect similar RNAs to be precipitating as well.

In depth analysis of isoxazole precipitated mRNAs from U2OS cells revealed that the 3' untranslated regions (UTRs) of pellet-enriched mRNAs were significantly longer than those of supernatant-enriched mRNAs (Fig 2-11A). Regulation through 3'UTRs are involved in many aspects of mRNA metabolism such as nuclear export, cytoplasmic localization, translation efficiency, and mRNA stability (Andreassi and Riccio 2009). Furthermore, with few exceptions,

cis-acting elements controlling asymmetric mRNA localization are located exclusively in the 3'UTR (Andreassi and Riccio 2009). Thus, we find it compelling that pellet mRNAs have longer 3'UTRs which provides real estate for regulation by RNA granule proteins. One of the RNA binding proteins present in the isx precipitates derived from human, mouse and fruit fly samples is the Pumilio protein. Flies encode a single Pumilio protein, and mammals encode two paralogous versions of this RNA binding protein (Pum1 and Pum2). Pumilio proteins bind to the nanos response element (NRE) in the 3'UTR of mRNAs via their PUF domains (Murata and Wharton, 1995). This domain, located towards the C-termini of Pumilio proteins, is composed of eight tandem, imperfect repeats of 36 amino acids that self-organize to form the RNA binding domain (Zamore et al., 1997; Zhang et al., 1997). Pumilio proteins also contain LC sequences composed of poly-glutamine and poly-alanine segments located on the N-terminal side of the RNA binding domain. Mouse strains carrying inactivating mutations in both alleles of the Pum2 gene have been reported to suffer deficits in dendrite maturation and morphology, consistent with the idea that Pumilio proteins play an important role in the dendritic transport of selected mRNAs (Vessey et al., 2010). Having observed that the fly Pumilio protein was present in the isx precipitate of *Drosophila* S2 cells, and that both Pum1 and Pum2 were precipitated in lysates prepared from human U2OS cells, mouse brain and testis tissue, and mouse ES cells, we performed a computational search for the canonical NRE sequence, 5'-

UGUANAUA-3', in the 3'UTR's of mouse brain mRNA samples that were either precipitated by the isx compound or left in the supernatant fraction. mRNAs harboring the canonical NRE sequence were found at a roughly 10X enhanced frequency in the isoxazole precipitated mRNAs compared with those left in the supernatant in both mouse brain and U2OS derived datasets (Fig 2-11B). The same enrichment was found when the computational search was restricted to the 3'UTRs of the precipitated RNAs compared with supernatant RNAs, thereby indicating that most of the NREs found in the pellet mRNAs are located in their 3'UTRs.

Gene ontology (GO) analysis of the two groups of mRNAs uncovered different biological processes enriched in each fraction (Fig 2-12). GO terms enriched in the pellet include cell adhesion, small GTPase mediated signal transduction, and transcription factors, while the supernatant was enriched in terms describing proteins involved in mitochondrial electron transport, nucleotide metabolism, and translational elongation, among others. We speculate that pellet mRNAs encode proteins that benefit from asymmetric mRNA localization which is a fast and efficient way to respond to extracellular cues whereas supernatant mRNAs encode proteins that are more "housekeeping" in nature. Enrichment of GO categories among the pellet and supernatant RNAs provides further evidence that the distribution of RNA is non-random and provides a biological rationale that dictates what RNAs are precipitated by biotinylated isoxazole.

Gel retention of RNA is co-dependent on both RNA-binding domain and low complexity address

Using the fluorescence microscopy-based gel retention assay, we demonstrated in a proof of principle experiment that RNAs too can be retained by mCherry-FUS gel if recruited by RNA-binding proteins that have both the low complexity address for interaction with the gel and an RNA-binding domain to bring with it RNA. We constructed an artificial RNA-binding protein tagged with CFP composed of the FUS LCS and the RNA-binding domain of bacteriophage capsid protein MS2 (Fig 2-13A, left). The MS2 RNA-binding domain binds a well-characterized 19-nt hairpin sequence with high affinity (Fig 2-13A, right). This strong interaction is the basis of a popular technique commonly used to study the *cis*-acting elements that target mRNAs to RNA granules (Keryer-Bibens et al., 2008). We synthesized a mutant derivative of this hairpin loop and covalently attached a fluorophore for visualization of RNA localization. As a control, we showed that recombinant CFP protein does not accumulate in the hydrogel (Fig 2-13C). CFP-FUS214 is retained by the gel but lacks the RNA-binding domain to carry in the hairpin RNA. By gel shift, we observe that CFP-MS2 is capable of binding to the synthetic RNA (Fig 2-13B), but this construct lacks the LCS address required for interaction with the FUS gel. Only CFP-FUS214-MS2 is able to recruit the hairpin RNA to co-localize with the FUS gel (Fig 2-13C). This

provides a model for how we think RNA granules may assemble using the FUS gel as a nucleating scaffold for LCS interactions. Simply put, proteins that are recruited into the *in vitro* granule minimally require two components: RNA-binding domains to bind cognate RNAs and low complexity sequence to transition into a reversible aggregate state.

We have demonstrated that in its gel form, FUS behaves as a nucleating scaffold for the recruitment of other granule proteins through the interactions of low complexity domains (Fig 2-6, 2-7). We have also showed that RNA can be trapped by the hydrogel if it is accompanied by the appropriate RNA-binding protein with low complexity sequence to access the gel (Fig 2-13). With that in mind, we next identified endogenous RNAs retained by the mCherry-FUS gel by incubating FUS gel droplets in whole cell lysate. The FUS gel is composed of only the first 214 amino acids of the N-terminal low complexity sequence and thus does not contain RNA-binding domains of its own. Therefore, any RNA that is retained in the gel must necessarily be carried into the gel by other RNA-binding proteins that interact with the FUS LCS. When compared to the mRNAs precipitated by biotinylated isoxazole, we find a striking concordance between the datasets human U2OS cells (correlation = 0.86) and mouse brain (correlation = 0.79) which are values as good as biological replicates (p-value < 2.0×10^{-16}; Fig 2-14A). Figure 2-14B shows qPCR validation of five representative RNAs that are trapped within the hydrogel and five RNAs that are retained in the lysate.

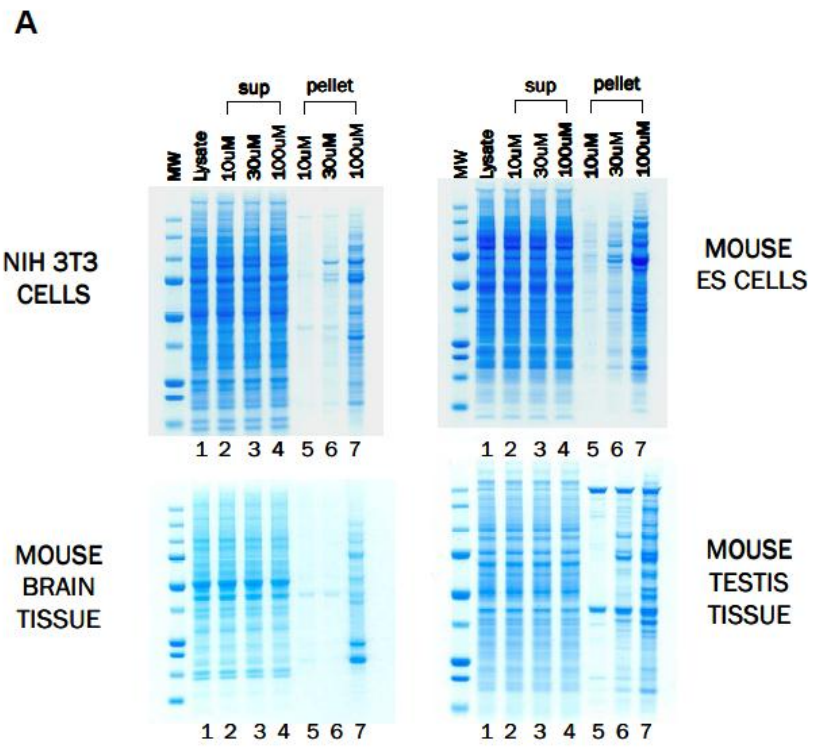
Using FUS gel as a scaffold for RNA granule formation, we determined the mRNA content of our gel-assemble *in vitro* granules to be in agreement with our isoxazole-assembled granules.

Precipitation of mRNAs in shRNA-mediated knockdown of RNA-binding proteins FUS and EWS

Because FUS plays very important roles both in the context of dysregulation of RNA metabolism in ALS pathology as well as normal biological function in local dendritic translation (Kwiatkowski et al., 2009; Vance et al., 2009; Fujii et al., 2005; Fujii and Takumi, 2005), we were especially interested in identifying target mRNAs of the protein that mediate its downstream functions. To address this question, we used biotinylated isoxazole to precipitate RNPs from U2OS cells that were treated with lentivirus expressing short hairpin RNA (shRNA) to knockdown FUS and its highly related paralog EWS. We then purified poly-A^+ RNAs from both the supernatant and the pellet fractions following isoxazole precipitation and identified RNAs enriched in either fraction by RNA-Seq. Knockdown efficiency was confirmed through both qPCR and Western blot (Fig 2-15A). Curiously, reducing the levels of FUS in the cell caused a slight upregulation of EWS levels on both mRNA and protein levels, and a lesser compensation was witnessed for the reverse situation. FUS/EWS double knockdown was able to decrease mRNA levels of both proteins by more than 5-

fold. Isoxazole-mediated precipitation from knockdown lysates would likely show no overt deficiency in assembled proteins besides the reduced presence of the shRNA-targeted protein considering the fact that many other granule proteins have their own low complexity domains and do not require FUS for precipitation. A subset of precipitated mRNAs were indeed found to be underrepresented in the isoxazole pellet following shRNA-mediated knockdown of FUS alone, EWS alone, or FUS and EWS together (Fig 2-15B). RNAs exhibiting reduced precipitation after single FUS or EWS knockdown show a high degree of overlap with the double knockdown, which is to be expected since the double knockdown is essentially a biological duplicate with the addition of a second lentivirus. 280 RNAs exhibited at least 2-fold or greater underrepresentation in the pellet (compared to control) following double FUS/EWS knockdown. 38 of 47 RNAs that were underrepresented following single FUS knockdown overlapped with the 280 RNAs of the double knockdown. 18 of 21 RNAs that were underrepresented as a function of single EWS knockdown overlapped with the 280 RNAs of the double knockdown. The statistical probability of the overlap exhibited in each case is $p < 2.2 \times 10^{-16}$. Interestingly, although FUS and EWS single knockdowns do not seem to significantly overlap with one another, double FUS/EWS knockdown affects a much wider range of mRNAs, indicating that the two paralogs likely share a significant subset of target RNAs. At the time of this writing, Hoell, Tuschul and colleagues published a crosslinking

immunoprecipitation-based approach to identifying RNA targets of FUS and paralogs EWS and TAF15 (Hoell et al., 2011). Significantly, 30 out of 38 of our predicted FUS targets (underrepresented in both the single FUS and the double FUS/EWS knockdowns) are found in their CLIP data (p-value = 9.982×10^{-12}), which supports the validity of our subtractive approach involving isoxazole precipitation.


A
sup
pellet
MW
Lysate
10uM
30uM
100uM
10uM
30uM
100uM
NIH 3T3 CELLS
MOUSE ES CELLS
MOUSE BRAIN TISSUE
MOUSE TESTIS TISSUE
1 2 3 4 5 6 7

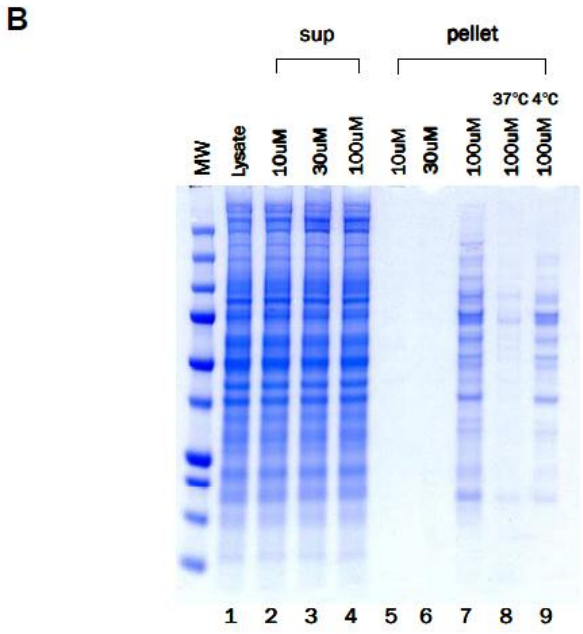
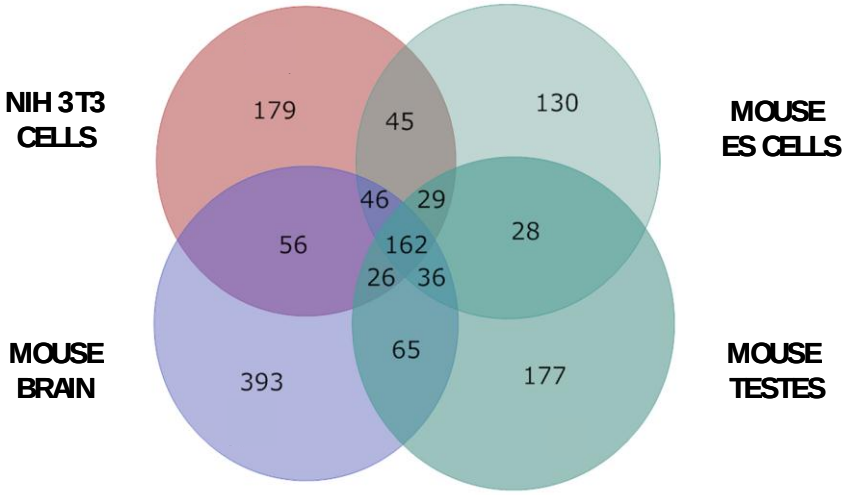

B
sup
pellet
37°C 4°C
MW
Lysate
10uM
30uM
100uM
10uM
30uM
100uM
100uM
100uM
1 2 3 4 5 6 7 8 9

Figure 2-1. Selective precipitation of proteins by the biotinylated isoxazole
(A) Biotinylated isoxazole-mediated precipitation of proteins from mouse 3T3 cells, ES cells, brain tissue, and testis tissue. Lysates were incubated with 10, 30, or 100 μM isx, pelleted by centrifugation, resuspended and resolved by SDS-PAGE and Coomassie staining. Precipitation increased in a dose-dependent manner and appeared to aggregate a minor subset of proteins coming from the cell or tissue lysates. The data for this figure was generated by Dr. Shanhai Xie.

(B) Isx precipitant, when resuspended in fresh buffer and warmed to 37°C, is solubilized and can be re-precipitated upon repeated exposure to the 100uM level of the isx chemical. The data for this figure was generated by Dr. Shanhai Xie.

A


NIH 3T3 CELLS
MOUSE ES CELLS
MOUSE BRAIN
MOUSE TESTES
179
45
130
46
29
56
162
28
26
36
65
393
177

B

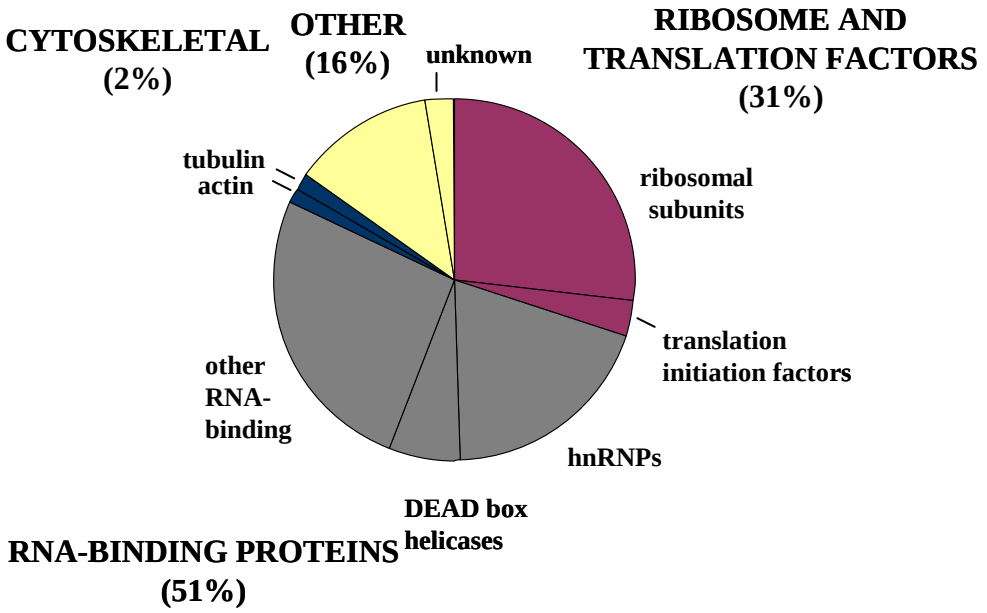

CYTOSKELETAL (2%)
OTHER (16%)
unknown
RIBOSOME AND TRANSLATION FACTORS (31%)
tubulin
actin
ribosomal subunits
translation initiation factors
other RNA-binding
hnRNPs
DEAD box helicases
RNA-BINDING PROTEINS (51%)

Figure 2-2. Mass spectrometry analysis of isx precipitates reveals 162 proteins precipitated in all four tissue and cell types

(A) Isx precipitated proteins were recovered from SDS gels and subjected to shotgun mass spectrometry. Between 300-500 polypeptides were identified as being precipitated by the b-isox chemical from lysates prepared from mouse brain and testis tissue, and mouse embryonic stem cells (ES) and NIH 3T3 cells. As visualized in the Venn diagram, proteins precipitated from the four lysates revealed significant overlap. 162 proteins were found in common in all four datasets, and were defined as a "core" subset of precipitated proteins highlighted for subsequent analyses.

(B) Breakdown of core isoxazole precipitated proteins by functional domains. 39 out of the 162 are ribosomal subunits, both large and small. Translation initiation factors present in the core set are eIF2, eIF3, and eIF4 which are components of the 43S pre-initiation complex that includes the 40S small ribosomal subunit and mRNA bound by cap-binding and poly-A binding factors. Other RNA-binding proteins predominantly contain RRM, KH, and DEAD-box motifs.

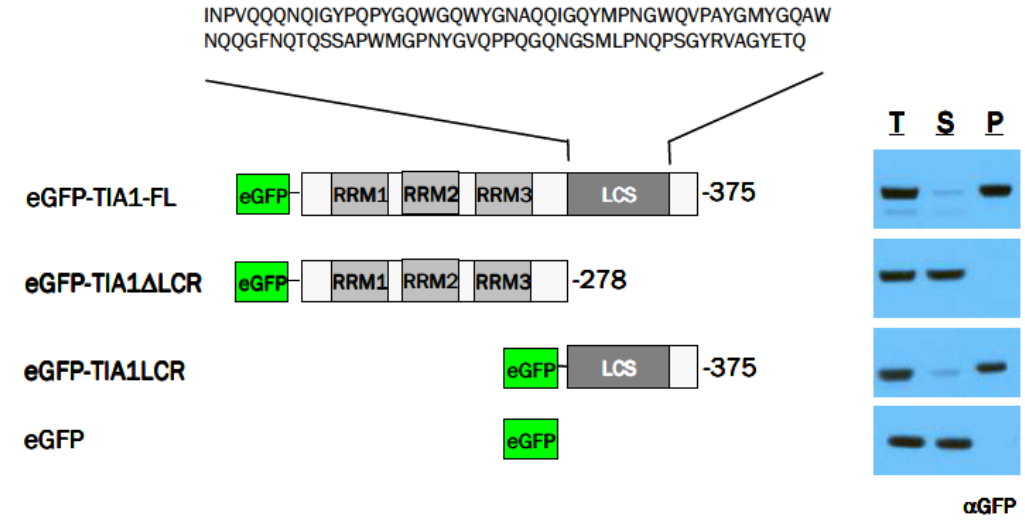

A
αFUS
sup
pellet
Lysate
10uM
30uM
100uM
10uM
30uM
100uM
BRAIN
TESTIS
ES
3T3
MOUSE BRAIN
sup
pellet
Lysate
10uM
30uM
100uM
10uM
30uM
100uM
αFUS
αEWS
αATXN2
αFXR1
B
sup
pellet
37°C
4°C
Lysate
10uM
30uM
100uM
10uM
30uM
100uM
100uM
100uM
αFUS
αTIA1
αActin
1 2 3 4 5 6 7 8 9

Figure 2-3. Biotinylated isoxazole precipitates proteins implicated in neurodegenerative diseases

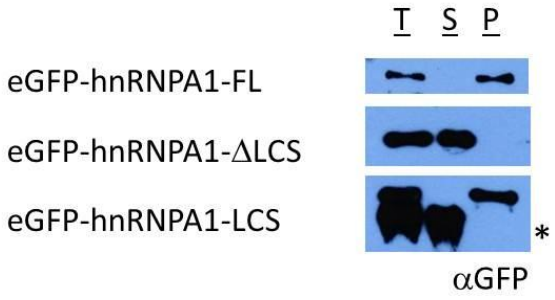
(A) Western blot assays showing behavior of FUS, ataxin 2, EWS and FXR1 proteins in response to exposure to the isx chemical. FUS was partially precipitated from mouse ES and NIH 3T3 cell lysates upon exposure to 30 μM of the isx compound. Quantitative precipitation of FUS was observed in all four lysates at the 100 μM compound concentration. EWS was also quantitatively precipitated from the mouse brain lysate at the 100 μM level of the isx chemical. For ataxin 2 and FXR, nearly complete precipitation was observed at the 100 μM level of the isx chemical. The data for this figure was generated by Dr. Shanhai Xie.

(B) Western blot assays showing that isx-mediated precipitation of FUS and TIA1 proteins is reversible. Mouse brain lysate was exposed to 100 μM levels of the isx chemical. Precipitated materials were retrieved by centrifugation, resolved by SDS PAGE and subjected to western blotting to identify the FUS and TIA1 polypeptides (lane 7). After resuspension in fresh buffer and warming to 37°C, both FUS and TIA1 became soluble (lane 8). Upon re-exposure to fresh isx at 100 μM and incubation on ice, both proteins were re-precipitated (lane 9). The data for this figure was generated by Dr. Shanhai Xie.

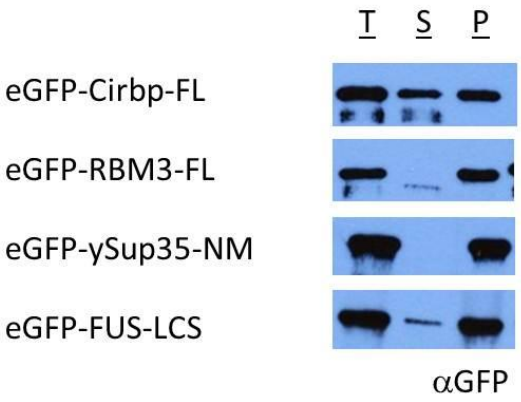
A


INPVQQQNQIGYPQPYGQWGQWYGNAQQIGQYMPNGWQVPAYGMYGQAW
NQQGFNQTQSSAPWMGPNYGVQPPQGQNGSMLPNQPSGYRVAGYETQ
eGFP-TIA1-FL
eGFP
RRM1
RRM2
RRM3
LCS
-375
eGFP-TIA1ΔLCR
-278
eGFP-TIA1LCR
eGFP
T S P
αGFP

B


T S P
eGFP-hnRNPA1-FL
eGFP-hnRNPA1-ΔLCS
eGFP-hnRNPA1-LCS
*
αGFP

C


T S P
eGFP-Cirbp-FL
eGFP-RBM3-FL
eGFP-ySup35-NM
eGFP-FUS-LCS
αGFP

D

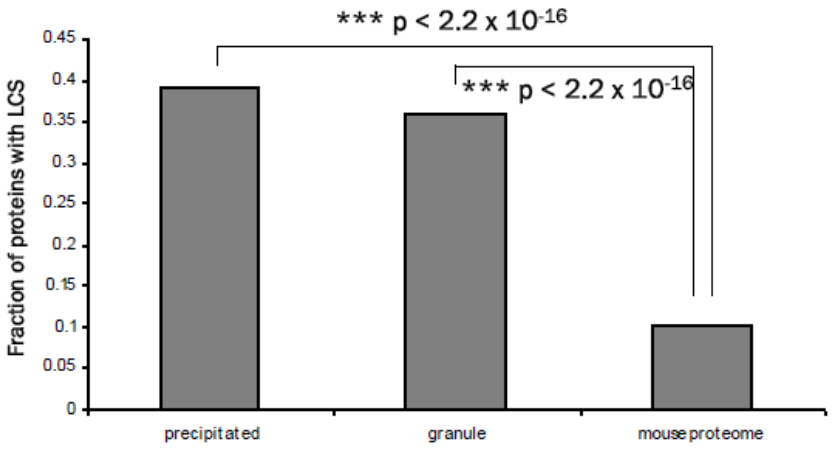

*** p < 2.2 x 10-16
*** p < 2.2 x 10-16
Fraction of proteins with LCS
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
precipitated
granule
mouse proteome

Figure 2-4. Low complexity sequence of TIA1 is necessary and sufficient for isoxazole-mediated precipitation

(A) Recombinant GFP-tagged proteins were purified from bacteria and incubated in lysis buffer with 100uM isx chemical. Lysate, supernatant and precipitant fractions were recovered and resolved by SDS PAGE. Western blot assays were carried out to probe for the recombinant proteins using antiserum specific to GFP. Full-length TIA1 protein was quantitatively depleted from the supernatant and precipitated into the pellet fraction. Deletion of 97 amino acids from the C-terminus of TIA1 protein, corresponding to the low complexity (LC) domain of the protein, abolished isx-mediated precipitation. Deletion of 278 amino acids from the N-terminus of the TIA1 protein, corresponding to the three RRM RNA binding domains of the protein, did not affect isx mediated precipitation. The isolated GFP domain was not precipitated by the 100uM level of the isx chemical.

(B) Western blot assays using antibody to GFP were carried to probe for GFP-tagged hnRNPA1 constructs precipitated by isx. Deletion of 135 amino acids from the C-terminus of the hnRNPA1 protein, corresponding to the low complexity (LC) domain of the protein, abolishes isx-mediated precipitation. Deletion of 185 amino acids from the N-terminus of the protein did not affect precipitation. The LC domain of hnRNPA1 is prone to degradation, likely from the C-terminus. These degradation products do not precipitate when incubated with isx (see asterisk).

(C) Western blot assays using antibody to GFP were carried to probe for GFP-tagged recombinant precipitated by isx. Full length Cirbp and RBM3 are precipitated by isx, as well as the prion domain of yeast Sup35 protein and the first 214 amino acids of the FUS protein that comprises the low complexity domain.

(D) Proteins were considered to harbor low complexity domains if they contained a stretch of low complexity sequence at least 35 amino acids in length as predicted by the SEG program. The fraction of proteins containing LC domains is greater in precipitated proteins and literature-derived granule proteins than the average for the human proteome. *** p-value < 0.0001 by unpaired two-tailed t-test, error is standard deviation from the mean.

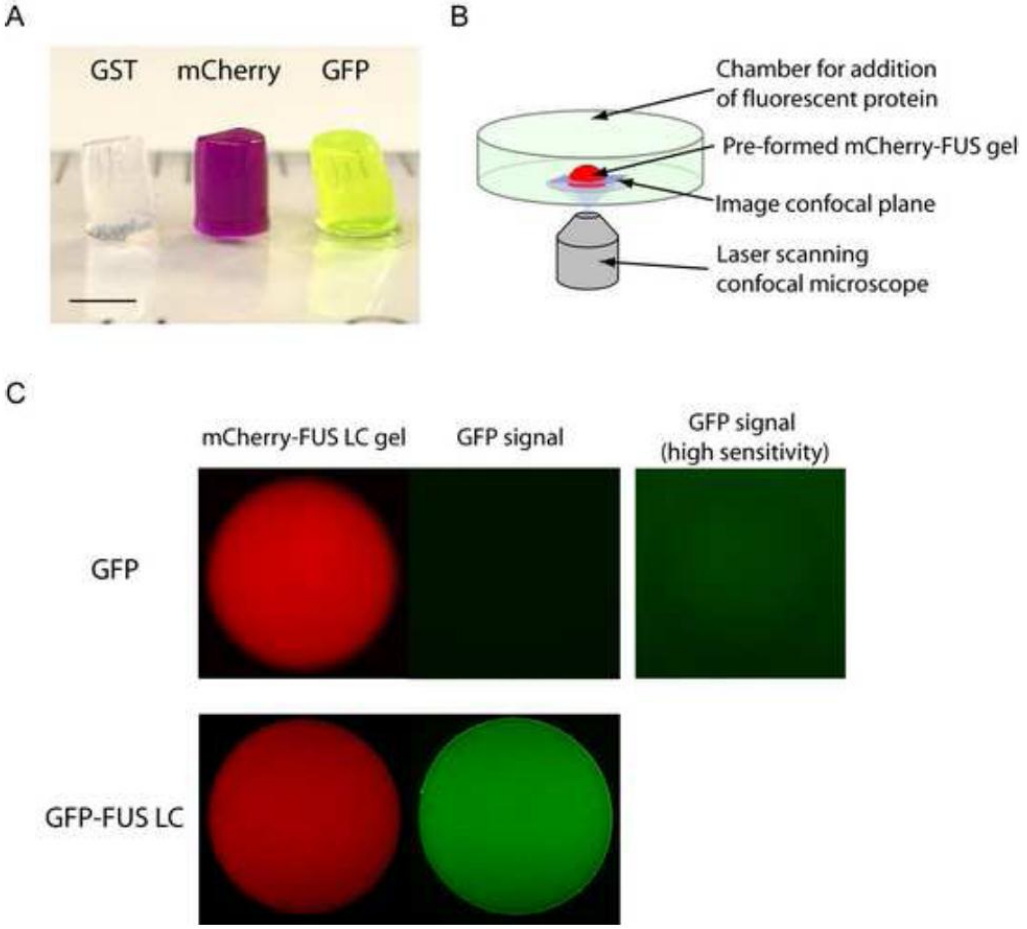

A
GST
mCherry
GFP
B
Chamber for addition
of fluorescent protein
Pre-formed mCherry-FUS gel
Image confocal plane
Laser scanning
confocal microscope
C
mCherry-FUS LC gel
GFP signal
GFP signal
(high sensitivity)
GFP
GFP-FUS LC

Figure 2-5. Hydrogels form upon concentration of the N-terminal, low complexity (LC) domain of FUS.

(A) Hydrogels formed following expression, purification proteins linking the N-terminal LC domain of FUS linked to three different protein tags GST, mCherry, and GFP. Concentrated solutions of the fusion proteins were filled in silicon tubes and incubated at 4 ˚C for 3 days. The pre-formed gels were then squeezed out and photographed. The hydrogels retained the cylindrical shape of the silicon tube and exhibit the respective colors of the tags (GST: clear, mCherry: red, GST: green). Scale bar at the lower left corner corresponds to 3mm.

(B) Schematic representation of experimental setup for fluorescent microscopic hydrogel binding assays. A hydrogel droplet of the mCherry:FUS LC domain was pre-formed on a glass-bottomed microscope dish. After the gel was exposed to 1 μM GFP-fusion proteins for two days, a section of the hydrogel was scanned with both the mCherry and GFP excitation wavelengths by confocal microscopy.

(C) Fluorescent images of mCherry:FUS LC hydrogel droplets exposed to GFP (top) and GFP:FUS LC domain (bottom). Left panels show mCherry signals (colored in red) from a section of a hydrogel of the mCherry-FUS LC domain. Right panels show GFP signals (colored in green) from the same section of the mCherry-FUS hydrogel. In GFP alone case (top), the GFP signal was also scanned at a higher detector sensitivity to emphasize the fact that no signal difference could be detected between inside and outside of the hydrogel. The data for this figure was collected by Dr. Masato Kato.

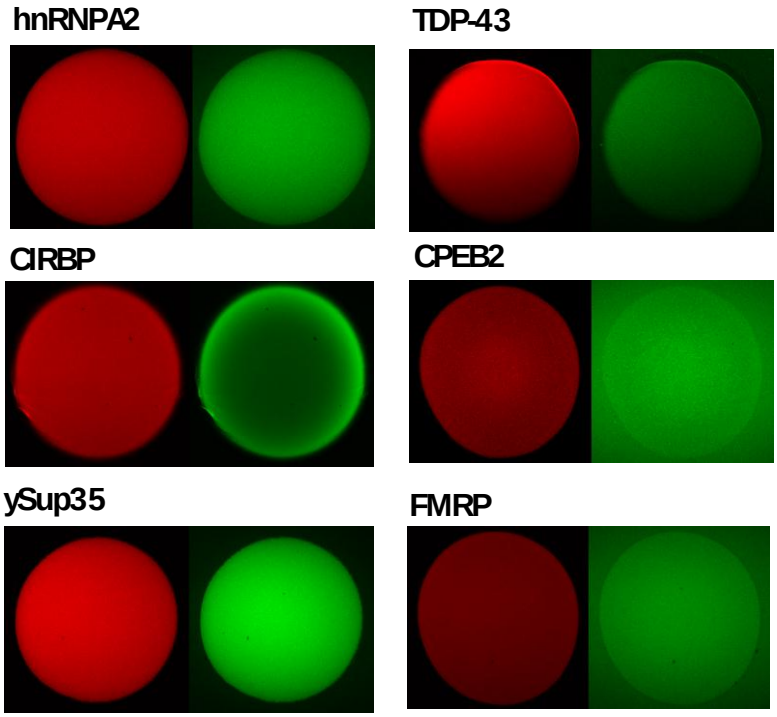


Figure 2-6. Low complexity polypeptides of RNA granule proteins are also retained by mCherry-FUS hydrogel.

Fluorescent images of mCherry:FUS LC hydrogel droplets exposed to GFP-tagged low complexity domains from several other isoxazole precipitated proteins (hnRNP A2, FMRP, CPEB2, TDP-43), full length Cirbp, and the NM fragment of yeast prion protein Sup35. These recombinant proteins show hetertypic trapping by the hydrogel to varying degrees, with Cirbp showing the strongest retention signals reminiscent of homotypic FUS hydrogel interactions and CPEB2 signals barely above background. The data for this figure was collected by Dr. Masato Kato

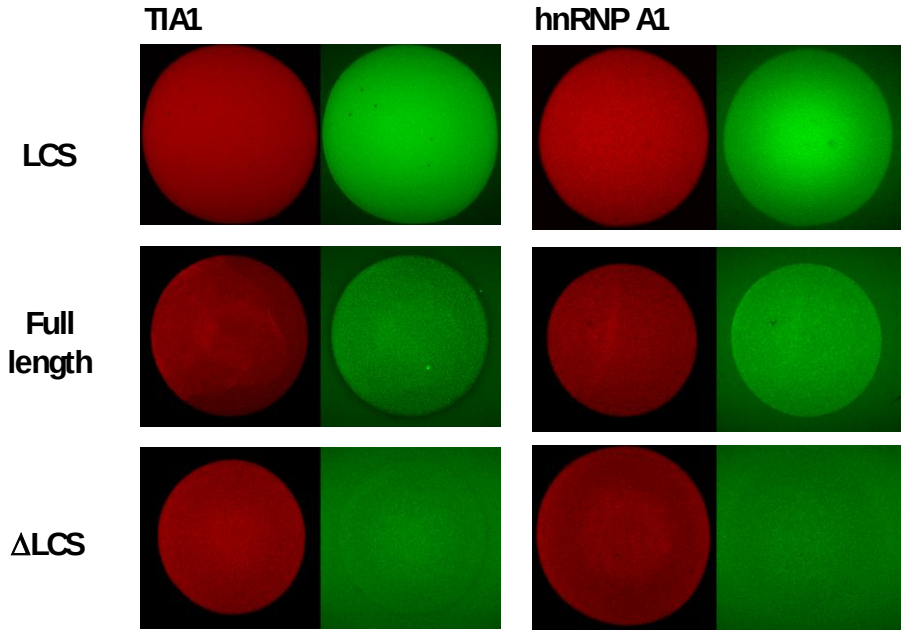


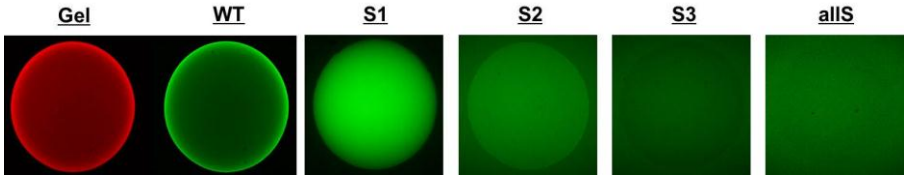
Figure 2-7. Hydrogel retention of TIA1 and hnRNPA1 require LCS

As with isoxazole-mediated precipitation (see Fig 2-4A and B), the low complexity domains of TIA1 and hnRNP A1 are also necessary for hydrogel retention. Deletion of this domain leads to diffuse fluorescence signal throughout the gel chamber whereas deletion of the N-terminal portion of the protein containing the RNA-binding domains does not prevent trapping of the protein.

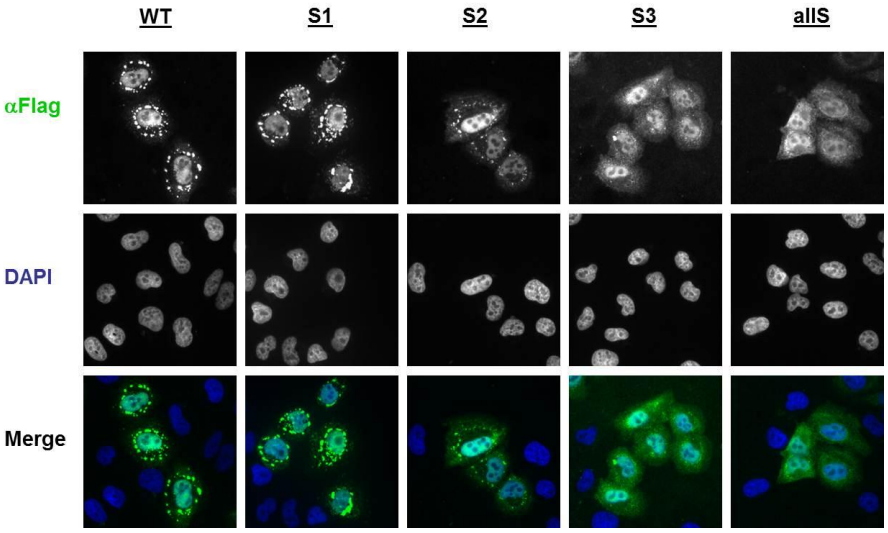
A


Wild type
MASNDYTQQATQSYGAYPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGY 50
GQSSYSSYGQSQNTGYGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGY 100
GQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQQPSYGGQQQSYGQQQSYN 150
PPQGYGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYGNQDQS 100
GGGGSGGYGQQDRG
S1
MASNDYTQQATQSYGASPTQPGQGYSQQSSQPYGQQSYSGYSQSTDTSGY 50
GQSSYSSYGQSQNTGYGTQSTPQGSGSTGGSGSSQSSQSSYGQQSSYPGY 100
GQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQQPSYGGQQQSSGQQQSSN 150
PPQGYGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYGNQDQS 100
GGGGSGGYGQQDRG
S2
MASNDYTQQATQSSGASPTQPGQGYSQQSSQPSGQQSSSGSSQSTDTSGY 50
GQSSYSSYGQSQNTGYGTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGY 100
GQQPAPSSTSGSYGSSSQSSSYGQPQSGSSSQQPSYGGQQQSSGQQQSSN 150
PPQGSGQQNQYNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYGNQDQS 100
GGGGSGGYGQQDRG
S3
MASNDYTQQATQSSGASPTQPGQGYSQQSSQPSGQQSSSGSSQSTDTSGY 50
GQSSSSSSGQSQNTGYGTQSTPQGYGSTGGYGSSQSSQSSSGQQSSSPGS 100
GQQPAPSSTSGSYGSSSQSSSYGQPQSGSSSQQPSYGGQQQSSGQQQSSN 150
PPQGSGQQNQSNSSSGGGGGGGGGGNYGQDQSSMSSGGGSGGGYGNQDQS 100
GGGGSGGYGQQDRG
allS
MASNDSTQQATQSSGASPTQPGQGSSQQSSQPSGQQSSSGSSQSTDTSGS 50
GQSSSSSSGQSQNTGSGTQSTPQGSGSTGGSGSSQSSQSSSGQQSSSPGS 100
GQQPAPSSTSGSSGSSSQSSSSGQPQSGSSSQQPSSGGQQQSSGQQQSSN 150
PPQGSGQQNQSNSSSGGGGGGGGGGNSGQDQSSMSSGGGSGGGSGNQDQS 100
GGGGSGGSGQQDRG

B


Gel
WT
S1
S2
S3
allS

C


WT
S1
S2
S3
allS
αFlag
DAPI
Merge

D

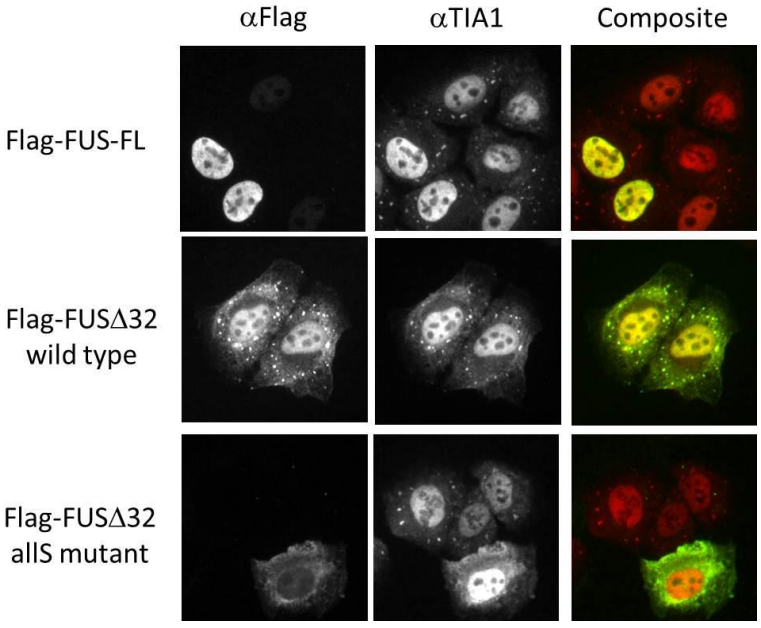

αFlag
αTIA1
Composite
Flag-FUS-FL
Flag-FUSΔ32
wild type
Flag-FUSΔ32
allS mutant

Figure 2-8. Mutation of tyrosine residues within the FUS LC domain correlatively affect hydrogel retention and stress granule association

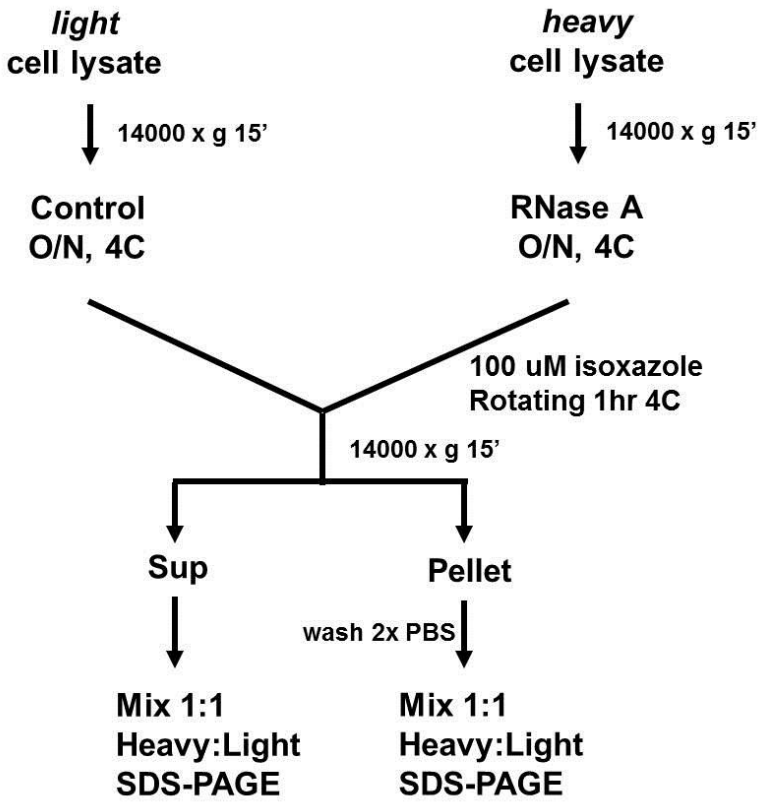
(A) The positions of tyrosine-to-serine mutations. Out of 27 tyrosine residues in the FUS LC domain, the indicated number of were randomly selected and replaced with serine.

(B) Impaired hydrogel binding of FUS LC domains bearing tyrosine-to-serine mutations. The resulting GFP:FUS tyrosine mutants were tested for hydrogel binding and compared with the wild type LC domain of FUS. The data for this figure was collected by Dr. Masato Kato.

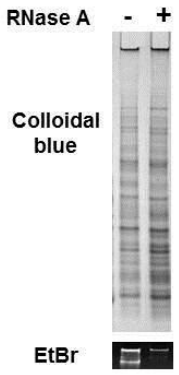
(C) Mammalian expression vectors were prepared wherein the FUS was fused to a Flag epitope tag at its N-terminus and the C-terminal nuclear localization sequence of FUS was deleted. In addition to a starting construct containing an intact, N-terminal low complexity (LC) domain, versions carrying the S1, S2, S3 and allS mutations were also prepared. Human U2OS cells were transfected with each expression vector, incubated for 36 hr, then exposed to 0.5mM arsenite to induce metabolic stress. Cells were then fixed and stained with antiserum specific to the Flag epitope tag. Stress granule staining was observed for the cells transfected with the expression vectors encoding the FUS variant containing an intact LC domain, as well as cells transfected with the S1 mutant. Marginal stress granule staining was observed in cells transfected with the S2 mutant, but no stress granule staining was observed for cells transfected with either the S3 or allS mutants.

(D) Transient transfection in U2OS cells of Flag-tagged constructs of full length FUS, ΔNLS FUS, and ΔNLS FUS with all 27 tyrosines in the N-terminal low complexity sequence mutated to serines. Cells were treated with 0.5 mM arsenite for 1 hr to induce stress granule formation and stained with antibody to the Flag epitope tag and co-stained with antibody to TIA1 protein as a marker of stress granules. Flag-FUS-FL appears predominantly nuclear and does not co-localize with TIA1 in cytoplasmic foci. Flag-FUSΔ32 with wild type sequence appears in both nuclear and cytoplasmic compartments and is readily recruited to cytoplasmic granules that co-stain with TIA1. Flag-FUSΔ32 with tyrosine to serine mutations, on the other hand, fails to be recruited into stress granules.

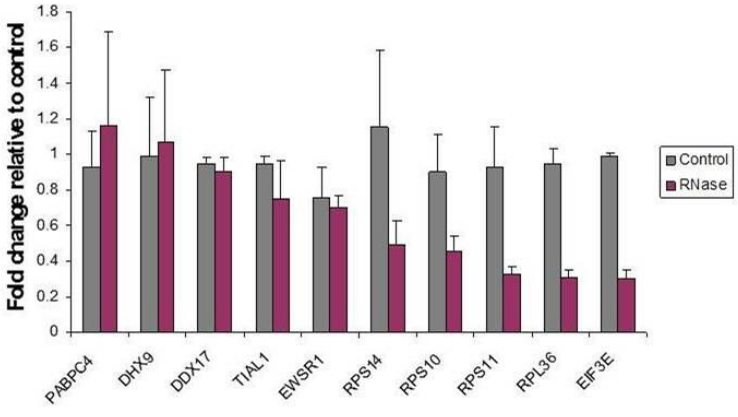
A


light
cell lysate
heavy
cell lysate
14000 x g 15'
14000 x g 15'
Control
O/N, 4C
RNase A
O/N, 4C
100 uM isoxazole
Rotating 1hr 4C
14000 x g 15'
Sup
Pellet
wash 2x PBS
Mix 1:1
Heavy:Light
SDS-PAGE
Mix 1:1
Heavy:Light
SDS-PAGE

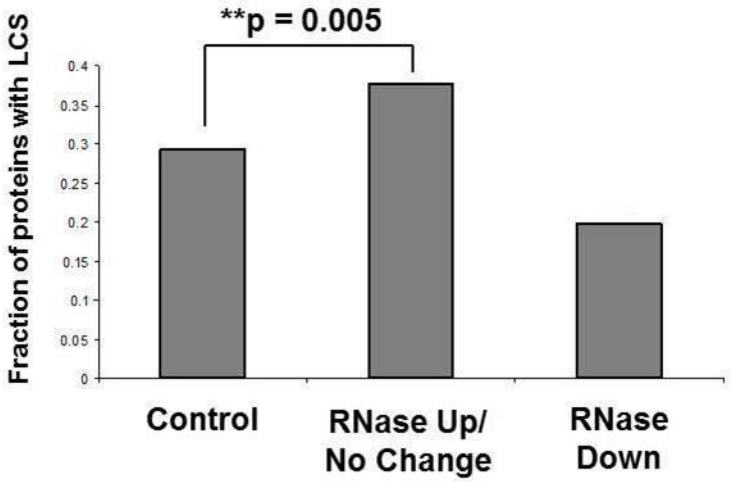
B


RNase A
-
+
Colloidal
blue
EtBr

C


Fold change relative to control
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
PABPC4
DHX9
DDX17
TIAL1
EWSR1
RPS14
RPS10
RPS11
RPL36
EIF3E
Control
RNase

D


**p = 0.005
Fraction of proteins with LCS
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Control
RNase Up/ No Change
RNase Down

E

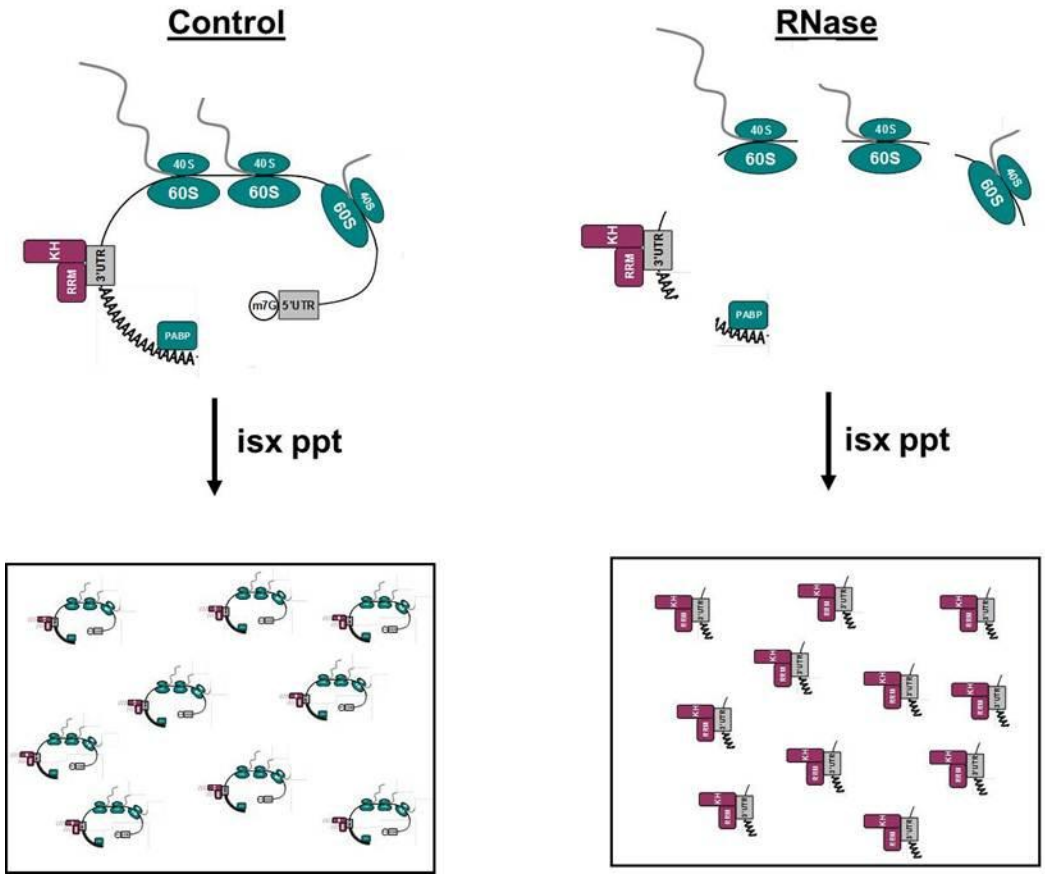

Control
RNase
40S
60S
KH
RRM
3'UTR
5'UTR
m7G
PABP
isx ppt
isx ppt

Figure 2-9. Isoxazole-mediated precipitation is partially sensitive to RNase.

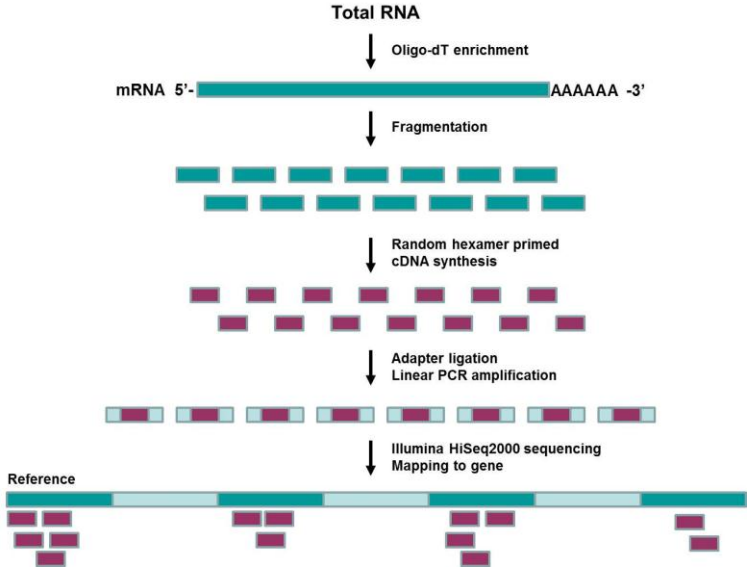
(A) Lysate is made from cells that are grown in the presence of either heavy isotope-containing lysine or normal or "light" lysine. The heavy lysate is treated with 100 μg/mL RNase A overnight at 4°C while RNase inhibitor is added to the light lysate to keep RNAs intact. Treated lysates are then subjected to precipitation with 100uM biotinylated isoxazole. The precipitated mixture is pelleted, washed, and resuspended in SDS sample. For both supernatant and pellet, heavy samples are mixed 1:1 with light samples for gel purification. The entire lane was subsequently excised from the gel submitted for SILAC mass spectrometry.

(B) Protein-RNA mixtures precipitated by biotinylated isoxazole were analyzed by SDS-PAGE. Top panel shows colloidal blue staining of proteins precipitated from these lysates by isoxazole with and without RNase treatment. Bottom panel shows ethidium bromide staining of high molecular weight RNA.

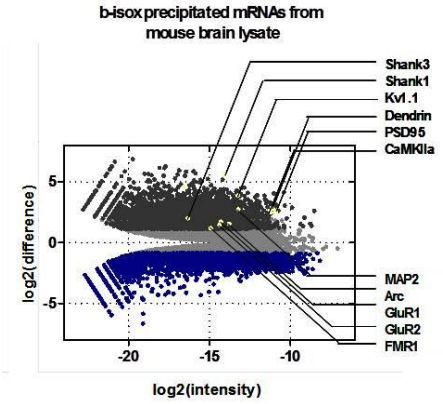
(C) SILAC mass spectrometry reveals the decreased presence of a subset of isx precipitated proteins following RNase treatment. Five representative proteins that precipitate independent of RNA are shown here alongside five representative proteins whose precipitation is sensitive to RNase. Error bars are standard deviation of triplicate experiments.

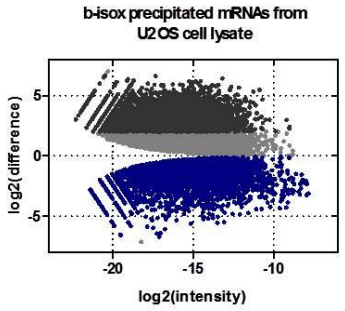
(D) The SEG program was used to determine whether the proteins that exhibit increased precipitation or no change versus reduced precipitation as a function of RNA degradation have low complexity domains. Proteins with LC domains are enriched among the proteins that upregulated precipitation or that do not change following RNase treatment when compared to control samples that were not treated with RNase. Proteins containing LC are not enriched domains in the reduced precipitation category.

A

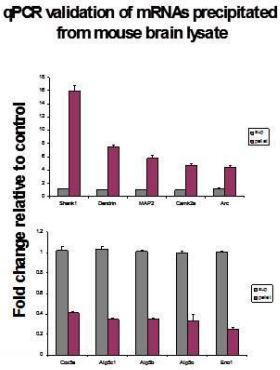

Total RNA
Oligo-dT enrichment
mRNA 5'-
AAAAAA -3'
Fragmentation
Random hexamer primed
cDNA synthesis
Adapter ligation
Linear PCR amplification
Illumina HiSeq2000 sequencing
Mapping to gene
Reference

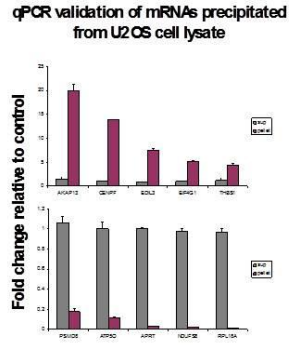
B


b-isox precipitated mRNAs from
mouse brain lysate
Shank3
Shank1
Kv1.1
Dendrin
PSD95
CaMKIIa
MAP2
Arc
GluR1
GluR2
FMR1
log2(difference)
log2(intensity)


b-isox precipitated mRNAs from
U2OS cell lysate
log2(difference)
log2(intensity)

C


qPCR validation of mRNAs precipitated
from mouse brain lysate
Fold change relative to control


qPCR validation of mRNAs precipitated
from U2OS cell lysate
Fold change relative to control

D

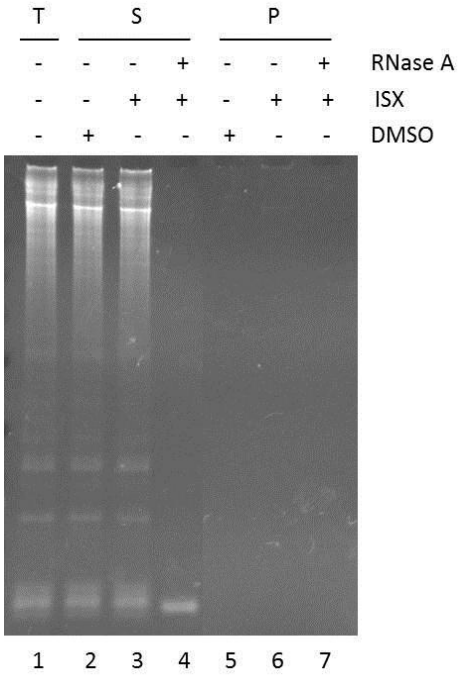


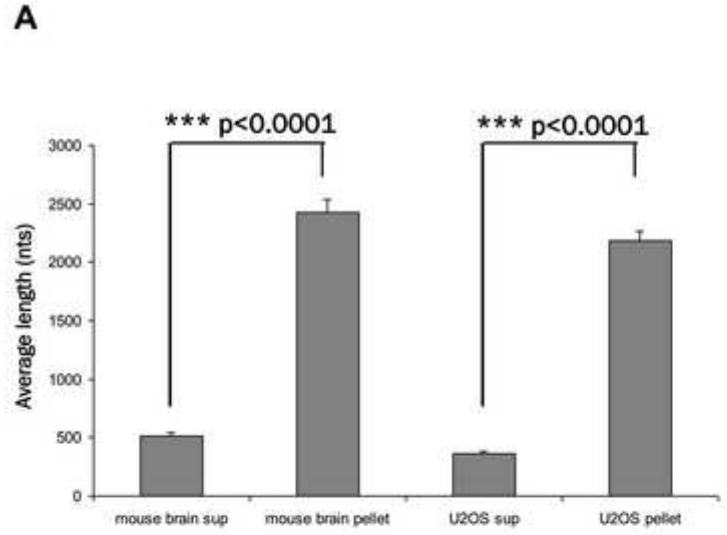
Figure 2-10. Isoxazole-mediated precipitation of mRNAs.

(A) Schematic cartoon of RNA-Seq workflow. Total RNA is purified from both the isoxazole precipitant (pellet) and remaining lysate (supernatant) and enriched for poly-A+ RNAs. RNA is then fragmented to a size range amenable to sequencing with the Illumina HiSeq platform (150-250 nts). cDNA is generated from these RNA fragments and adaptors are ligated onto both ends in order to provide sequence to prime amplification. Reads from Illumina sequencing are mapped to the reference genome for subsequent data analysis.

(B) RNA-Seq plots for mRNAs precipitated by isx from mouse brain lysate (top panel) and human U2OS cell lysate (bottom panel). Relative abundance of the transcript in the cell is plotted on the X-axis in log2 scale and fold enrichment of the transcript in the isx precipitant (the pellet to supernatant ratio) is plotted on the Y-axis in log2 scale. Shaded in dark grey are mRNAs more than 2-fold enriched in the pellet and shaded in blue are mRNAs that are 2-fold or greater excluded from the pellet. Eleven known dendritic RNAs are highlighted in the mouse brain dataset.

(C) qPCR validation for 5 representative mRNAs enriched in pellet and 5 representative RNA retained in supernatant precipitated from mouse brain lysate (top panel) and U2OS cell lysate (bottom panel). Data normalized to internal control are plotted as fold change relative to supernatant. Error bars are standard deviation of triplicate experiments.

(D) Purified total RNA was exposed to either biotinylated isoxazole or DMSO (control) under the same conditions as isoxazole precipitation from whole cell lysate and resolved by SDS-PAGE and stained with ethidium bromide (EtBr). RNase A was added to the reaction in lane 4 and 7 to show that the EtBr staining corresponded to RNA. Neither DMSO nor biotinylated isoxazole precipitate purified RNA (lanes 5, 6) as all of the RNA is retained in the supernatants (lanes 2, 3) which looks identical to total RNA input (lane 1).


A
*** p<0.0001
*** p<0.0001
Average length (nts)
3000
2500
2000
1500
1000
500
0
mouse brain sup
mouse brain pellet
U2OS sup
U2OS pellet

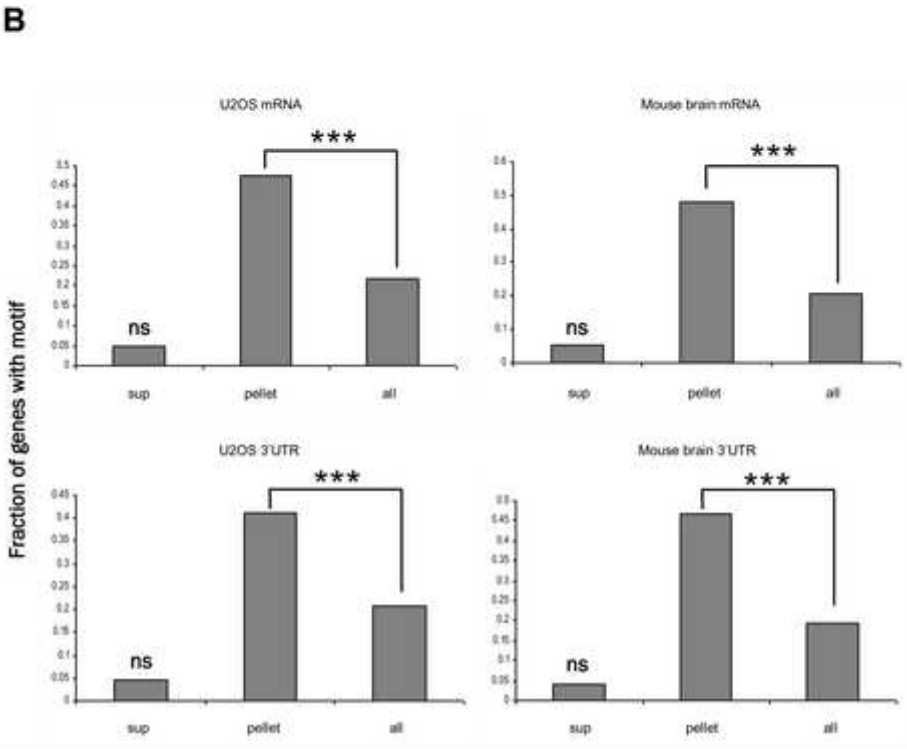

B
Fraction of genes with motif
U2OS mRNA
Mouse brain mRNA
U2OS 3'UTR
Mouse brain 3'UTR
***
ns
sup
pellet
all

Figure 2-11. Analysis of 3'UTR lengths in supernatant and pellet RNAs

(A) Pellet-enriched 3'UTRs from both mouse brain and U2OS RNA-Seq datasets are longer than supernatant-retained 3'UTRs. The average 3'UTR length in nucleotides is compared here. Error is standard deviation from the mean and *** p-value < 0.0001 by Mann-Whitney test.

(B) Computational analysis of the consensus *pumilio* binding motif 5'-UGUANAUA-3' in supernatant- and pellet-enriched mRNAs (top panels) and analysis of motifs in 3'UTRs only (bottom panels) in both mouse brain and U2OS datasets. Fraction of RNAs with *pumilio* motif is significantly enriched in pellet mRNAs and 3'UTRs compared to supernatant mRNAs and 3'UTRs in both U2OS and mouse brain RNA-Seq datasets. *** p-value < 2.2x10-16 by Fisher's Exact test.

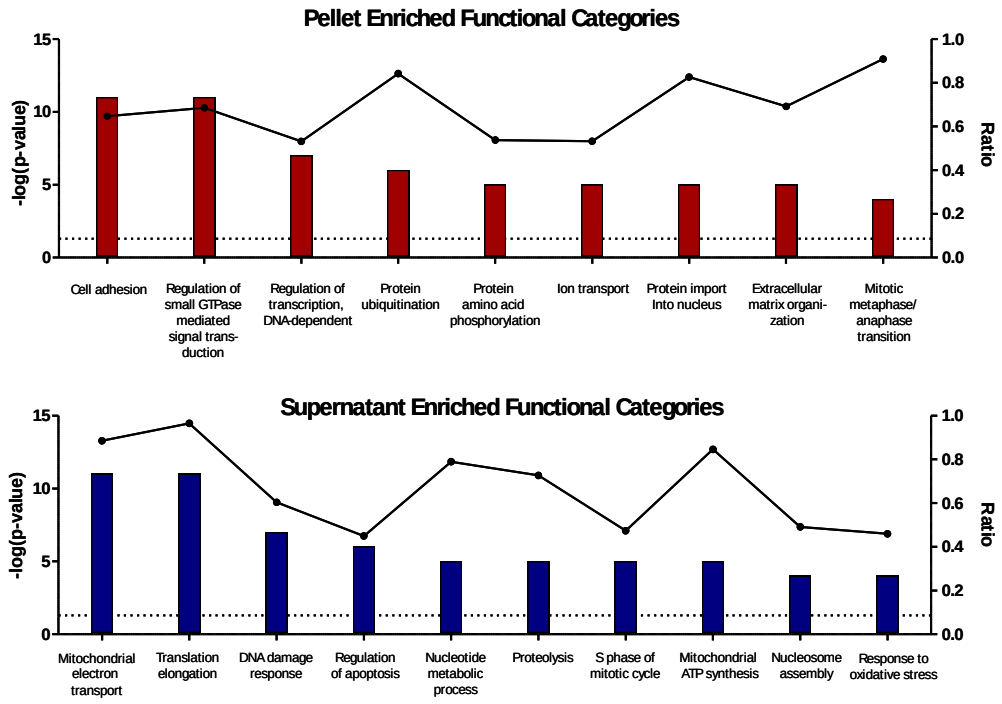
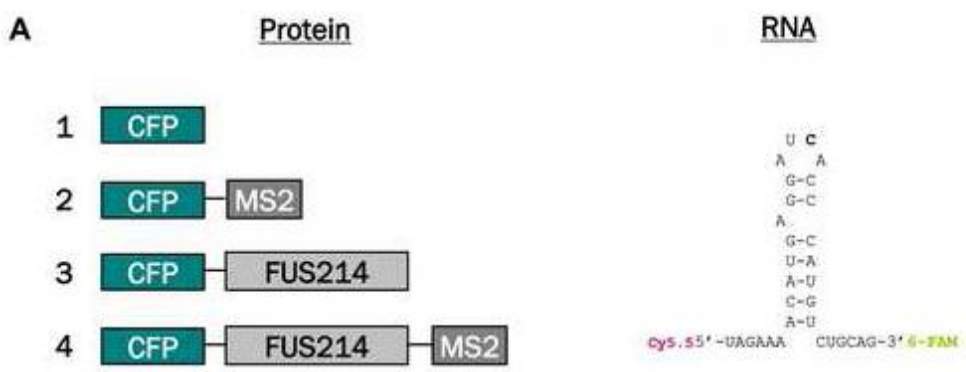
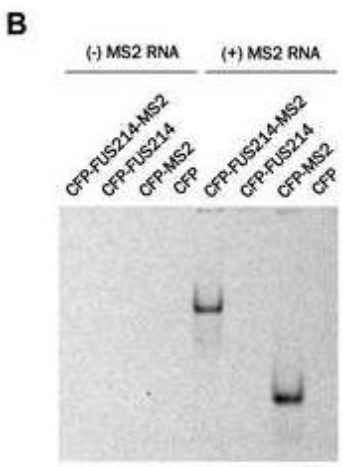


Figure 2-12. Gene ontology analysis of pellet enriched and supernatant retained mRNAs

Hypergeometric distribution analysis was used to determine overrepresented GO categories in pellet and supernatant RNAs. Bar graph (axis on the left) shows the –log of the p-value (higher is more significant) and the line graph (axis on the right) denotes the ratio of enrichment in these categories. Dotted line indicates the threshold of significance ($p = 0.05$).


A
Protein
RNA
1 CFP
2 CFP MS2
3 CFP FUS214
4 CFP FUS214 MS2
U C
A A
G-C
G-C
A
G-C
U-A
A-U
C-G
A-U
Cy5.5 5'-UAGAAA CUGCAG-3' 6-FAM


B
(-) MS2 RNA
(+) MS2 RNA
CFP-FUS214-MS2
CFP-FUS214
CFP-MS2
CFP
CFP-FUS214-MS2
CFP-FUS214
CFP-MS2
CFP

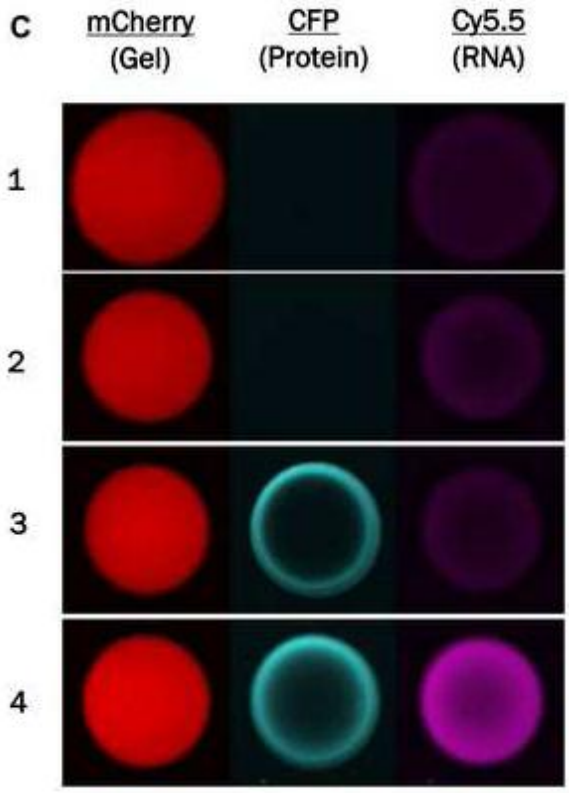

C
mCherry (Gel)
CFP (Protein)
Cy5.5 (RNA)
1
2
3
4

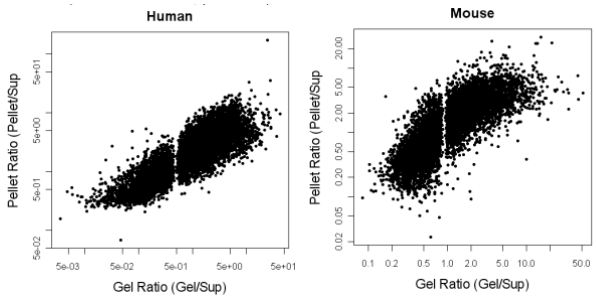
Figure 2-13. Gel retention of RNA is co-dependent on both RNA-binding domain and LCS velcro

(A) CFP fusion protein constructs used for gel retention experiment are recombinant His-tagged proteins purified from bacteria. Hairpin RNA was synthesized covalently linked with either a 3' 6-FAM (used for gel shift) or a 5' Cy5.5 (used for gel retention).

(B) Gel shift demonstrates binding of 6-FAM labeled MS2 RNA hairpin to fusion proteins with MS2 RNA binding domain. First four lanes act as control lanes showing that the Cy2 channel does not detect CFP fluorescence and that the unbound 6-FAM labeled RNA has been run off the gel. Only the constructs with MS2 domains (CFP-MS2 and CFP-MS2-FUS214) interact with the hairpin RNA.

(C) Gel retention assay performed with both CFP-tagged protein and Cy5.5-labeled hairpin RNA. MS2 RNA is only recruited into the hydrogel when incubated with CFP-FUS214-MS2. Lacking low complexity sequence, CFP and CFP-MS2 do not accumulate in mCherry-FUS gel. CFP-FUS214 is retained by the gel, but lacks the MS2 RNA binding domain required for interaction with the MS2 RNA.

A


Human
Mouse
Pellet Ratio (Pellet/Sup
Gel Ratio (Gel/Sup)
5e-02
5e-01
5e+00
5e+01
5e-03
0.02
0.05
0.20
0.50
2.00
5.00
20.00
0.1
0.2
0.5
1.0
2.0
5.0
10.0
50.0

B

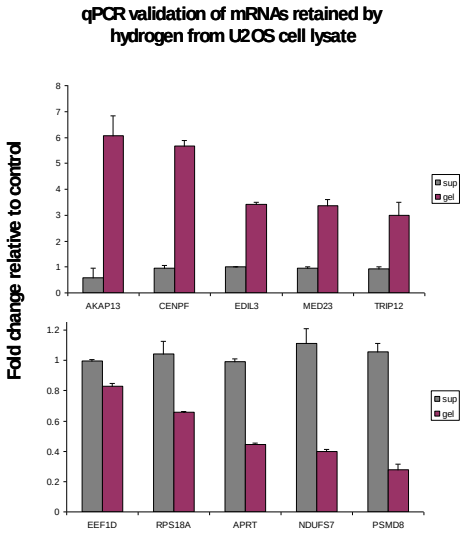

qPCR validation of mRNAs retained by hydrogen from U2OS cell lysate
Fold change relative to control
sup
gel
AKAP13
CENPF
EDIL3
MED23
TRIP12
EEF1D
RPS18A
APRT
NDUFS7
PSMD8

Figure 2-14. Gel retention of mRNAs from cell lysate correlates with isoxazole-precipitated mRNAs

(A) RNA-Seq data from gel retained RNAs were compared to RNA-Seq data from isoxazole precipitated RNAs with gel to supernatant ratio graphed on x-axis and isx pellet to supernatant ratio graphed on the y-axis for every RNA with an RPKM value > 1. High degree of concordance was found between these two datasets: U2OS cells (correlation = 0.86, p-value < 2 x 10-16) and mouse brain (correlation = 0.79, p-value < 2 x 10-16).

(B) qPCR validation for 5 representative mRNAs enriched in hydrogel and 5 representative RNA retained in supernatant precipitated from U2OS cell lysate (bottom panel). Data is plotted as fold change relative to supernatant. Error bars are standard deviation of triplicate experiments.

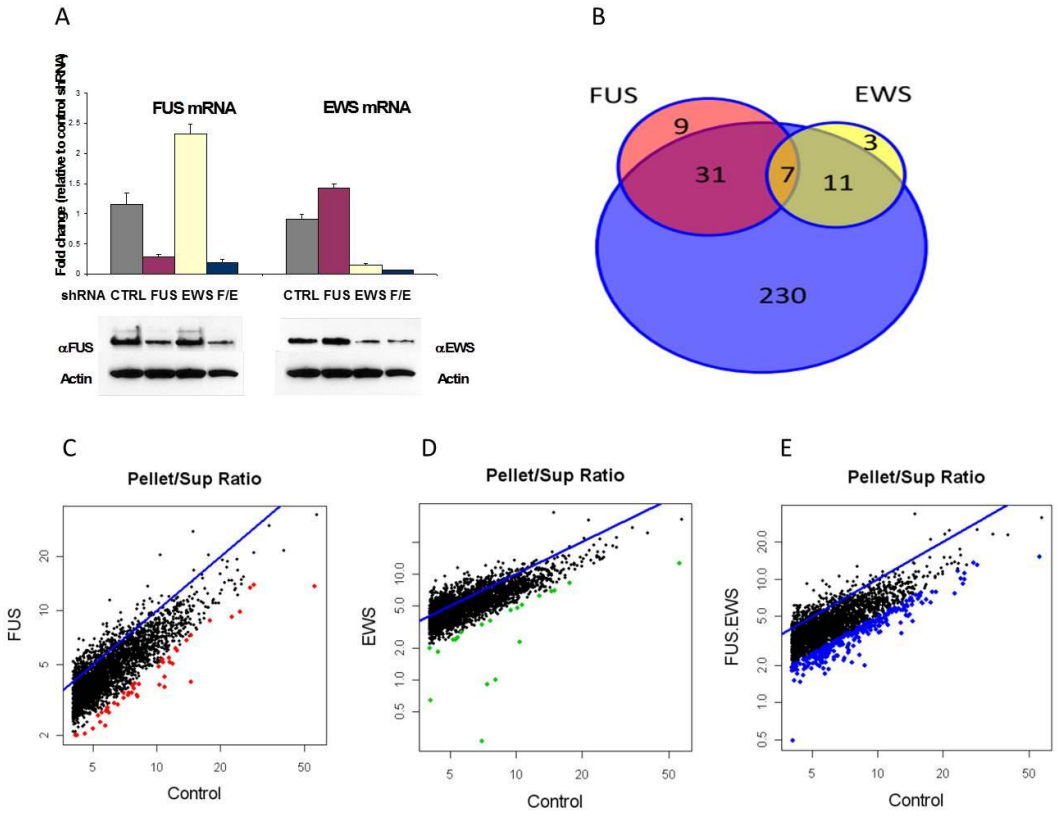

A
FUS mRNA
EWS mRNA
Fold change (relative to control shRNA)
shRNA CTRL FUS EWS F/E
CTRL FUS EWS F/E
αFUS
Actin
αEWS
Actin
B
FUS
EWS
9
31
7
3
11
230
C
Pellet/Sup Ratio
FUS
Control
D
Pellet/Sup Ratio
EWS
Control
E
Pellet/Sup Ratio
FUS.EWS
Control

Figure 2-15. Precipitation of mRNAs following knockdown of FUS, EWS or both in U2OS cells.

(A) Efficiency of shRNA-mediated knockdown of FUS and EWS in U2OS cells was confirmed by both qPCR and Western blot analysis. U2OS cells were infected with lentivirus constructs expressing shRNA specific to FUS, EWS, or both. FUS only knockdown cells exhibit a mRNA reduction of roughly 70% relative to cells infected with an inert virus. EWS only knockdown cells exhibit an mRNA reduction of roughly 85% relative to control cells. Under double knockdown conditions, mRNA levels of FUS and EWS were reduced to 80% and 90% respectively as compared with control.

(B) Overlap of RNAs underrepresented in isx precipitant as a function of FUS, EWS or FUS/EWS knockdown. 280 RNAs exhibited at least 2-fold under representation (as compared to control) following double FUS/EWS knockdown. 38 of 47 RNAs that were under-represented at least 2-fold following single FUS knockdown overlapped with the 280 RNAs of the double knockdown. 18 of 21 RNAs that were under-represented at least 2-fold following single EWS knockdown overlapped with the 280 RNAs of the double knockdown. The statistical probability of the degree of overlap exhibited in each case is $< 2.2 \times 10^{-16}$ calculated using the Chi-square test.

(C) RNA-Seq was performed on both RNAs precipitated by isx from knockdown cell lysate and RNAs left in the supernatant. Scatterplot representing the RNA-Seq data from the single FUS knockdown with control pellet/sup ratio graphed on the X-axis and FUS knockdown pellet/sup ratio on the Y-axis. The blue line indicates that the ratio between knockdown and control is 1 or no change. Data points highlighted in red indicate RNAs that exhibit at least a 2-fold reduction in precipitation as a function of FUS knockdown.

(D) RNA-Seq data from the single EWS knockdown with control pellet/sup ratio graphed on the X-axis and EWS knockdown pellet/sup ratio on the Y-axis. Data points highlighted in green indicate RNAs that exhibited at least a 2-fold reduction in precipitation as a function of EWS knockdown.

(D) RNA-Seq data from the double FUS/EWS knockdown with control pellet/sup ratio graphed on the X-axis and FUS/EWS knockdown pellet/sup ratio on the Y-axis. Data points highlighted in blue indicate RNAs that exhibit at least a 2-fold reduction in precipitation as a function of double FUS/EWS knockdown.

Core Precipitated Proteins

Actb	hnRNPM	Ubap2l
Actg1	hnRNPR	Upf1
Ap2a2	hnRNPU	Ybx1
Ap2b1	hnRNPUl	Ythdf1
Atxn2	Hsp8	Ythdf2
Atxn2l	Hspa2	Ythdf3
Bat2d	ILF2	Zfr
Calm2	Khdrbs1	
Caprin1	Khsrp	
CW17R	Leng8	
Dazap1	Lsm12	
Ddx1	Matr3	
Ddx3x	mCBP	
Ddx5	mot-1	
Ddx6	Ncl	
Dhx9	Npm1	
Ddx17	Nufip2	
eIF2b	Nup214	
eIF3d	Pabpc1	
eIF3h	Pabpc4	
eIF3i	Pcbp1	
eIF4g3	Pcbp2	
EWS	Pcid2	
Fam98a	Polr2b	
FMR1	Purb	
Fubp1	Rbm3	
Fubp3	Rbms1	
FUS	Rbmxrt	
FXR1	Serbp1	
FXR2	Sfrs1	
G3bp1	Sfrs5	
G3bp2	Sfrs6	
GNB2L1	Sfrs7	
Hist1h2ai	Smap2	
Hist1h2bp	Snrnp70	
hnRNPAB	Srp20	
hnRNPA0	Ssb	
hnRNPA1	Stau1	
hnRNPA2	Syncrip	
hnRNPA3	TAF15	
hnRNPC	TDP-43	
hnRNPD	TIA1	
hnRNPDl	TIAl1	
hnRNPF	Tnrc6b	
hnRNPG	Tnrc6c	
hnRNPH1	Tra2b	
hnRNPH2	Tuba1b	
hnRNPK	Tubb5	
hnRNPL	Ubap2	

RNA Granule Proteins

Ago2	FXR2	Pan2/3	ZBP1
AKAP350	G3bp1	Pat1	
Apobec3g	G3bp2	Pcbp2	
Atxn2	Gbp2	PKP1/3	
Bat1a	Ge-1/Hedls	PMR1	
BRF1	GNB2L1	Pop2/Caf1	
Calreticulin	Grb7	Prkra	
Caprin1	GW182	Prohibitin 2	
CCAR1	hMex3A	PRTB	
Ccr4	hMex3B	Psd3	
Celf2	hnRNPA0	Pspc1	
Cirbp	hnRNPA1	Pum1	
Cpeb	hnRNPA3	Pum2	
CUG-BP1	hnRNPAB	Pura	
Cyfip2	hnRNPC	Purb	
Dcp1	hnRNPD	Rap55	
Dcp2	hnRNPH1	Rbm3	
Ddx1	hnRNPK	Rbm42	
Ddx3x	hnRNPL	RHAU	
Ddx5	hnRNPM	Roquin	
Ddx6	hnRNPR	Rpl3	
Dhx9	hnRNPU	Rpm2	
DIC1	Hrp1	RSK2	
DIS1	Hsp27	Rtcd1	
Eap1	Hspa8	Serbp1	
Ebs1	Htt	Sfpq	
Edc1-2	Igf2bp1	SGNP	
Edc3	ILF2	Smaug1	
eIF1a2	ILF3	SMN	
eIF2a	Importin-8	Stau1	
eIF2c2	IP5K	Syncrip	
eIF2s2	KHC/KLC	TAF15	
eIF2s3x	Khdrbs1	TDP-43	
eIF3	Khsrp	TDRD3	
eIF4a1	Lin28	Thoc4	
eIF4a2	LINE1	TIA1	
eIF4g1	Lsm1	TIAl1	
Elavl1	Matr3	Tnrc6b	
Elavl2	MBNL1	TRAF2	
Elavl3	MEX67	Trim2	
Elavl4	MLN51	Trim3	
eRF1	Musashi	TTP	
eRF3	Ncl	Upf1	
EWS	Nono	Upf2	
FAK	Npm1	Upf3	
FAST	Nrp1	Vtx1	
FMR1	NXF7	Xrn1	
FUS	p97/NAT1	Ybx1	
FXR1	Pabpc1	Ygr250c	

Table 1. Core precipitated proteins. 162 core proteins precipitated by biotinylated isoxazole from four different mouse cell and tissue types. Ribosomal subunits and mass spectrometry contaminants (human keratin protein) have been removed from the list as well as any unknown proteins. Highlighted in yellow are the proteins that overlap with Table 2.

Table 2. RNA granule proteins. Known RNA granule proteins referenced from literature. List currated from Kanai et al., 2004; Elvira et al., 2006; Buchan and Parker, 2009. Highlighted in yellow are the proteins that overlap with Table 1.

CHAPTER 3:

Discussion and Future Directions

Proposed mechanism of precipitating activity by biotinylated isoxazole

We do not fully understand the capability of the biotinylated isoxazole to precipitate proteins from cellular lysate. This behavior does not seem to be related to the cardiogenic or neurogenic activities of its predecessor small molecules (Sadek et al., 2008; Schneider et al., 2008) as these parental compounds fail to reproduce the precipitation. Because we observed that the isoxazole itself will precipitate under the same concentrations and conditions required for precipitation of RNA granule proteins, we hypothesize that the small molecule may be assembling into some sort of nucleating aggregate, amenable for the conversion of RNA granule proteins into the insoluble state. Moreover, the isoxazole itself will return to solution when warmed from 4°C to 37°C (Fig 2-1B, 2-2B), which results in the release of the proteins back to solubility as well.

Low complexity sequences are important for RNA granule formation

In the previous chapter, I showed that the sequence determinants for both isoxazole-meditated precipitation and hydrogel retention are domains of low complexity. Because both protein and RNA components seem to match that of known RNA granule components, we consider these to constitute models of *in*

vitro RNA granule assembly. Thus, the question remains, are low complexity sequences important in a biological context? I demonstrated that [G/S]Y[G/S] repeats in the FUS protein are essential for the protein to be recruited into stress granules upon induction with arsenite. Furthermore, the hydrophobic contacts supplied by these tyrosine motifs resemble those of phenylalanine-glycine (FG) repeats that are found in a subset of nucleoporins that make up the nuclear pore complex (Fig 3-2). FG repeats are thought to be intrinsically disordered polypeptide chains (Denning et al., 2003) and as such they have been difficult to crystallize and examine function from a structural standpoint. FG nucleoporins are thought to constitute the permeability barrier of the nuclear pore channel, effectively gating the nucleus against passive diffusion of macromolecules above a certain size (Wälde and Kehenbach, 2010). Like the tyrosine repeats in the FUS low complexity domain, FG repeats can also form a three-dimensional hydrogel above a certain concentration (Frey et al., 2006). Frey and Görlich further demonstrated that the FG hydrogel reproduces some of the behaviors of FG nucleoporins in their role in selective gating of the nucleus (Frey and Görlich, 2007; Frey and Görlich, 2009).

Interestingly, *C. elegans* P granules are observed to tightly associate with clusters of NPCs during a window of developmental time in the adult gonads (Pitt et al., 2000). In addition, 14 out of 20 *C. elegans* nucleoporins were found to be required for the perinuclear distribution of P granules (Updike and Strome, 2009),

suggesting that the RNA granules physically interacted with proteins of the nuclear pore complex to maintain its localization. Furthermore, the *C. elegans* homolog of Nup98, which is a FG nucleoporin known to shuttle on and off the NPC, was found to be in P granules and even required for granule formation (Voronina and Seydoux, 2010). Nup98 was also detected among the proteins precipitated by isoxazole from mammalian tissues, an interesting connection between FG repeat containing proteins and isoxazole-mediated precipitation. Finally, and most intriguingly, P granule components themselves, helicases called GLH-1, GLH-2, and GLH-4, contain FG repeats remniscent of nucleoporins (Updike et al., 2011). Updike, Strome and colleagues further demonstrated that P granules *in vivo* exhibited similar characteristics of size exclusion as the FG hydrogel using fluorescently-labelled dextrans of various sizes (Updike et al., 2011). Moreover, Brangwynne, Hyman and colleagues demonstrated that P granules will behave in a liquid-like fashion, dripping and coalescing off of nuclei, when a shearing force is applied across cells (Brangwynne et al., 2009). This unique observation provides evidence that RNA granules formation is analogous to liquid-gas phase transition. These studies not only suggest that the FG hydrogel has a biological role in NPC gating, but also that RNA granules may use the same multivalent hydrophobic contacts for association with the NPC and mitigating the dynamics of granule assembly.

Other domains of low complexity have been previously shown to be required for granule formation. Several granule proteins have been described to have "prion-like" domains such as TIA1, Lsm4p, and Edc3p (Gilks et al., 2004; Decker et al., 2007; Reijns et al., 2008). The amino acid composition of a prion-like protein is typically to have overrepresentation of glutamines and asparagines (Kramer et al., 2009). One protein of particular interest is the cytoplasmic polyadenylation element binding protein (CPEB) that promotes adenylation of mRNA tails and, by extension, enhances the translation efficiency of the transcript (ref). The *Aplysia* CPEB protein also has a prion-like domain located in the first 160 residues of the protein, 48% of which is composed of glutamines (Si et al., 2003b). Of the 4 mammalian homologs of CPEB, the characteristically glutamine-rich LCS of ApCPEB is conserved in CPEB2 and CPEB3. Deletion of this polypeptide fragment in *Aplysia* reduces the number of ApCPEB-positive puncta present in neuronal processes (Si et al., 2010). Consistent with our model of low complexity sequences being required for RNA granule assembly, my interpretation of this data is that ApCPEB localizes to RNA granules in these neuronal processes and that deletion of the prion-like domain abrogates this recruitment into granules. Depletion of this protein in neurons leading to destabilization of synaptic facility (Si et al., 2003a) would be a consequence of dysfunctional transport granules not delivering RNAs for localized translation. An alternate hypothesis would be that the biological function of this prion-like

domain is to form permanent aggregates that in turn encode synaptic memories. However, due to the intrinsically dynamic nature of RNA granule assembly and disassembly, it seems unlikely that the basis of these interactions is strictly prion-like, classically defined as irreversible and highly resistant to protease and detergent treatment in *in vitro* assays (Kramer et al., 2009) but certainly, there is the possibility that dysfunction of RNA granules as in diseases such as ALS might lead to pathological prion-like aggregations.

Finally, an area in our studies that we have not explored fully is the varied composition of what we generalize as low complexity sequences. In this chapter, I have discussed two types of sequences: FG repeats and Q/N-rich "prion-like" sequences. It is not difficult to imagine that amino acid repeats of different composition might have different binding partners and that further classification of these types of sequences may shed light on interesting biological interactions.

Dynamic aggregation of low complexity polypeptides is a widespread mechanism regulating cellular metabolism

In several independent studies, depletion of microscopically visible RNA granules were shown to have little or no effect on what was presumed to be the cellular function of the granules – mRNA decay and translational repression for P-bodies and germ cell specification for *C. elegans* P granules (Chu and Rana, 2006; Stoeckline et al., 2006; Decker et al., 2007; Gallo et al., 2010). One could

interpret these studies as saying that RNA granules have no real function, that the individual protein components are competent to carry out their tasks without forming large aggregates, or that smaller aggregates are still present despite not being able to visualize them under the microscope. By extension, if granules that we cannot see are still functional, does this mean that other small aggregates are also present in the cell? Transactivation domains of transcription factors are thought to promote transcription through the recruitment of various factors to the proximity of the targeted gene (Tsai and Nussinov, 2011). It is important here to note that although transcription factors have infamously simple sequences, they are not found in RNA granules and are not precipitated by isoxazole or retained by hydrogel, already hinting at the specificity of LCS interactions as I alluded to earlier. Could it be possible that a transactivation domain here is acting under the same principles of dynamic aggregation principles as RNA granule components, except that instead of recruiting RNA-binding proteins, it seeks out proteins with compatible low complexity polypeptides? Furthermore, it has been documented that chromosomal translocations involving the low complexity domains of FUS and EWS, as well as the FG repeat domains of Nup98 and Nup214 and DNA-binding domains of various transcription factors cause cancers in human patients (Janknecht, 2005; Xu and Powers, 2009). Perhaps, in this aberrant setting, these LCS domains are recruiting transcription loci into the wrong kinds of granules and somehow, this translates into LCSs behaving like super-transactivation

domains which transform these fusion proteins into oncogenes. It does seem to corroborate the idea that these proteins are operating in a modular fashion. In other words, formerly attached to an RNA-binding domain, the LCS would be targetting RNAs specified by the RRM or KH domains into translationally repressed particles. In these oncogenic fusion proteins, however, the RNA-binding domains have been swapped out for zinc fingers and leucine zippers that recognize double-stranded DNA, thus changing the cargo carried into the aggregation.

Isoxazole precipitated RNAs and hydrogel retained RNAs are predicted to be regulated by RNA granules

Depletion of specific RNA granule proteins causes abnormalities in the structure of dendritic spines (e.g. Vessey et al., 2010; Comery et al., 1997), possibly due to the mislocalization of RNAs transported by these proteins. Dendritically targeted RNAs are important for building the postsynaptic density such as Camk2a, Arc, AMPA receptor subunits, and PSD95, which are all RNAs enriched in the isoxazole pellet and trapped by the FUS hydrogel (Burgin et al., 1990; Link et al., 1995; Lyford et al., 1995; Böckers et al., 2004; Grooms et al., 2006; Muddashetty et al., 2007). Furthermore, mislocalization of Camk2a and Arc transcripts has been shown to impair long-term potentiation, which is a cellular model for memory and learning (Miller et al., 2002; Plath et al., 2006). Thus, our data

seems to be consistent with the current model of localized translation at the synapse.

An unexpected category found to be enriched in the isoxazole precipitant were RNAs encoding transcription factors (Fig 2-12). Surprisingly, this observation is not entirely without precedence. Using radiolabeled antisense RNA from isolated dendrites, Crino, Eberwine and colleagues were able to detect the RNA encoding the cAMP response element binding proteins or CREB and confirm its dendritic localization by *in situ* hybridization (Crino et al., 1998). The mRNA for this transcription factor is also captured in our *in vitro* RNA granule preparations. In both *Aplysia* and *Drosophila*, CREB has been implicated in establishment of long-term memories (Dash et al,, 1990; Yin et al., 1994). It has long been established that long-term memory consolidation paradigms in these organisms requires new protein synthesis (Davis et al., 1984), which supports the hypothesis that dendritic translation of CREB mRNA might be important for memory formation. Other dendritically localized mRNAs encoding transcription factor that have been reported include zfr-268/EGR-1, neuralized, and engrailed (Crino et al., 1998; Timmusk et al., 2002; Di Nardo et al., 2007), all three of which were found among isoxazole precipitated RNAs. A compelling model called “dendritic imprinting” proposed by Eberwine and colleagues is the idea that transcription factors might be dendritically translated in response to synaptic activity, possibly modified or tagged with post-translational modifications, and

then delivered via retrograde transport to the cell nucleus for function (Eberwine et al., 2001). Our results would suggest that more transcription factors might be subjected to dendritic imprinting than previously thought.

RNAs encoding cell adhesion proteins, regulators of small GTPase signaling, and the extracellular matrix were among the top gene ontology categories overrepresented in isoxazole precipitated RNAs (Fig 2-12). One biological process that potentially represents a cross section of these three is the establishment of polarity in migrating cells (Kole et al., 2005). For example, cell adhesion factors are important for attachment of the cell's protruding edge to the substrate in order to gain traction for movement (Huttenlocher et al., 1995). Also, the Rho family of small GTPases has long been established as critical regulators of cytoskeletal organization at the leading and trailing edge of migrating cells that is essential for motility (Nobes and Hall, 1999). From our data, it would seem to suggest that these proteins, which are asymmetrically partitioned in the cell to facilitate directional movement, may also be encoded by asymmetrically localized RNAs. In fact, there is evidence that mRNAs are translocated to cellular protrusions in the process of cytoskeletal remodeling. As mentioned previously, β-actin mRNA is found localized to the leading edge of migrating fibroblasts as well as growth cone protrusions (Kislauskis et al., 1993; Zheng et al., 2001). In addition, mRNAs encoding all seven subunits of the Arp2/3 complex, which is required for nucleating actin polymerization, were found localized to leading edge

protrusions (Mingle et al., 2005). By expression microarray analysis, Mili et al. (2008) identified 50 other transcripts that were enriched in the pseudopodial protrusions of migrating fibroblasts and encoded proteins involved in cytoskeletal rearrangement, vesicle trafficking, microtubule-based transport, and RNA metabolism. Furthermore, ribosomes and translation factors have been documented to traffic to fibroblast protrusions upon stimulation with migratory cues (Willett et al., 2010), lending support to the hypothesis that these RNAs for actin and motility factors might be locally translated. Mislocalization of β-actin using antisense oligos directed towards the 3'UTR of the transcript impairs directional motility, as does treatment with translation inhibitor puromycin (Kislauskis et al., 1997), which further supports this interpretation for these studies.

A subtractive method for identifying candidate targets of RNA binding proteins

Using RNA-Seq to identify mRNAs that exhibit decreased precipitation after shRNA knockdown of precipitated proteins constitutes a general method that can be used to define networks of interactions between RNA binding proteins and mRNAs. Conventional techniques employed to identify candidate target RNAs rely on purification of RNAs that immunoprecipitate with RNA-binding proteins of interest and chemical or UV crosslinking to strengthen the contacts between

protein and RNA and prevent inappropriate binding upon cell lysis that creates a mixing of the contents between sub-cellular compartments. Though these methods have been able to yield testable results, especially if the researcher has specific targets in mind that can be easily queried by qPCR, they are typically complicated and labor-intensive, requiring many steps in library preparation and data validation. Our method requires only that one achieves an efficient knockdown of the protein under investigation, which is still non-trivial but has been majorly simplified with the advent of catalogs of RNAi oligos and shRNA viruses available commercially. There is no background caused by non-specific binding of antibodies used for pull-downs and no complicated fractionation steps that can potentially degrade one's samples. One caveat to our subtractive method, however, is the probable redundancy among RNA-binding proteins targeting the same mRNAs. This is a possibility that we tried to circumvent by knocking down the FUS paralog EWS which has highly conserved RNA-binding domains and is also observed as a component of RNA granules (Kanai et al., 2004). With our knockdown method, we are likely to identify only those RNAs that especially dependent on FUS for their localization to granules.

Thomas Tuschul and colleagues published a report on RNA targets of FUS and paralogs EWS and TAF15 using a technique called photoactivatable ribonucleoside-enchanged cross-linking and immunoprecipitation or PAR-CLIP (Hoell et al., 2011). For this method, cells were fed a photoreactive uridine

analog 4-thiouridine which was allowed to incorporate into nascent RNAs and then irradiated with UV light at 365 nm to induce crosslinking (Hafner et al., 2010). One interesting and useful consequence of using 4-thiouridine is the analog exhibits a thymidine to cytosine conversion upon UV irradiation, which effectively maps the site of crosslinking. 30 out of 38 of predicted FUS target RNAs identified in our study were also captured in the Tuschul study, indicating a statistically significant overlap between the two studies (p-value = 9.982×10^{-12}). Another interesting aspect of this study is the use of a Flag-HA-tagged mutant construct of FUS to recapitulate the pathological form of the protein found in ALS patients. In the diseased ALS brain, FUS protein is found in cytoplasmic aggregates (Kwiatkowski et al., 2009; Vance et al., 2009) which would lead to the speculation that sequestration of the factor would lead to loss of regulation on RNA targets or perhaps the acquisition of new, pathological targets. Hoell et al., 2011 show that potentially both may be occurring, as the pathological FUS immunoprecipitates some RNA targets more frequently and others less frequently than wild type FUS. Therefore, these 30 candidate mRNAs that overlap between the two studies may be of particular interest to understanding the dysregulation of RNA metabolism at the heart of ALS pathology.

Finally, to complement my identification of candidate RNAs of FUS binding, I would like to computationally search for *cis*-acting elements that might mediate this interaction. In order to facilitate this motif search, it is essential to be

able to highly enrich the analysis with RNAs that actually interact with FUS, and which I believe our technique has accomplished. Complicating this study is the frequently observation that *cis*-acting elements in single-stranded RNAs are not always simple primary sequence as in transcription factor DNA binding sites, but often adopt elaborate secondary structures like *bicoid* mRNA (Ferrandon et al., 1994). Hoell, Tuschul and colleagues identified a hairpin loop in several of their candidate FUS targets that have three positions at the base of the loop in common (Hoell et al., 2011), which would have been challenging to identify on a primary sequence level.

Future directions

Confirmation of our RNA localization predictions by microscopic visualization of would be one area of our studies that would greatly benefit from further study. However, the majority of localized RNAs recently coming to light appear to be under the detection sensitivity of traditional *in situ* hybridization. Also, *in situ* hybridization requires fixation of the cell and cannot give information about the dynamics of RNA localization. One technique that can be used for real-time detection is the MS2-GFP system where a GFP-tagged MS2 RNA-binding domain of phage coat protein is co-expressed with the RNA of interest, tagged with the MS2 RNA which is a 19-nt hairpin loop that binds with high affinity (Zenklusen et al., 2007). The disadvantage of this technique is that it involves

indirect visualization of the GFP fluorescence that presumable is tethered to the RNA of interest by virtue of the RNA-binding activity of the MS2 domain. Furthermore, overexpression of the MS2-GFP fusion protein easily aggregates in the cell, therefore appearing as false positives in granule assays (Robert Singer, personal communication). Also, the MS2-GFP contains its own trafficking elements so that the introduction of multiple copies of this element to the targeted RNA may in turn affect RNA trafficking (Tyagi, 2009). Another real-time strategy for monitoring RNA localization involves the introduction of fluorescently-labeled antisense oligos directly into the living cells, either by microinjection or RNA transfection (Okabe et al., 2011; Tyagi and Kramer, 1996). These methods require the efficient delivery of exogenous reagents into the cell, which could be difficult in certain cell lines. Furthermore, synthesizing fluorescently-labeled RNA oligos may prove cost ineffective when it comes to the validation of hundreds of RNAs. A third option requires the fusion of a 98-nt RNA aptamer to the RNA of interest which complexes and activates 3,5-dimethoxy-4-hydroxybenzylidene imidazolinone (DMHBI) a small molecule derived from the GFP fluorophore (Paige et al., 2011). In this report, Paige et al. tag the 5S rRNA with the aptamer, aptly designated “spinach” because of its green fluorescence, and deliver this fusion gene into cultured cells by transient transfection. One can easily imagine creating stable cell lines where the spinach aptamer is knocked into the 3’ regions of targeted transcripts so that the

processing of the RNA remains under the control of endogenous elements. By this method, we would be able to optimize the levels of expression and consistently monitor the targeted RNAs without concern over delivery of reagents into the cell. However, whether low abundance RNAs can be detected by any of these techniques remains an unanswered question.

Another area of interest is the potential discovery of noncoding RNAs that act in either a regulatory or structural capacity within the RNA granule. Because we specifically enrich for poly-A^+ RNAs in our RNA-Seq analyses, we do not capture, for example, small noncoding RNAs like microRNAs. One could easily imagine miRNAs targeting specific dendritic mRNAs to transport granules and sequestering the transcripts in a silenced complex until translation is initiated. Accordingly, Ago2-containing puncta has been observed in neuronal processes, along with P-body components Dcp1a, RCK/p54, and GW182 (Cougot et al., 2008). These structures were found to co-localize with zipcode binding protein 1 (ZBP1) which localizes β-actin transcript to growth cones and lamellipodia, and GFP-FMRP, both well-studied transport granule proteins (Cougot et al., 2008; Kislauskis et al., 1993; Zheng et al., 2001). Using laser capture to isolate hippocampal neurites and cell bodies, Kye, Kosik, and colleagues identified 75 out 187 rat miRNAs queried by multiplexed RT-PCR to be 2-fold or greater enriched in the neurites over soma (Kye et al., 2009) and several miRNAs have been localized in a candidate-based approach (e.g. Schratt et al., 2006). In several

studies, suppression of dendritically localized miRNAs with 2'-O-methyl modified antisense oligos is shown to affect spine morphology (Schratt et al., 2006; Siegel et al., 2009). I propose to use the cell-free RNA granule preparations that we can achieve with biotinylated isoxazole and hydrogel retention to comprehensively identify miRNA constituents of RNA granules in an unbiased fashion. Assuming that the mRNA regulated by these miRNAs co-resides in the same granules, it should then be a relatively straightforward computational analysis to identify potential targets of repression.

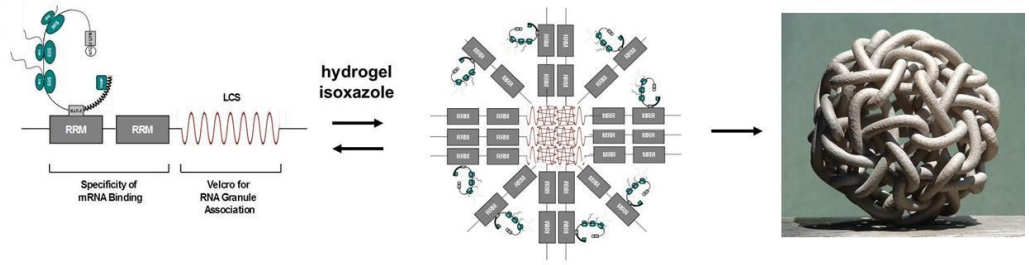


Figure 3-1. A Working Hypothesis for Cell-free RNA Granule Assembly

A simplistic explanation for *in vitro* RNA granule assembly by the biotinylated isoxazole and hydrogel models. In the first panel (left) is a cartoon representation of a typical RNA-binding protein in solution. It has RNA-binding domains that interact with its cognate RNAs (here depicted as RRM domains) and a low complexity sequence that allows it to reversibly aggregate into RNA granules. The biotinylated isoxazole compound and hydrogel retention somehow facilitate this transition from soluble protein to reversible aggregate and both require low complexity sequence to do so. In making the transition to the granule-like state, RNA-binding proteins subject their cognate RNAs to granule regulation. In the third panel (right) is a representation of an irreversible aggregate that the same RNA-binding proteins can be trapped in. These include prions, amyloid plaques, neuronal inclusions, and other pathological aggregates. The image in the third panel is a picture of a sculpture entitled “DodecaPentafoil Tangle” by Carlo H. Séquin (April 2005).

Figure 3-2. Side-by-side comparison of [G/S]Y[G/S] repeats in FUS low complexity domain with FG repeat domain of yeast nucleoporins Nsp1

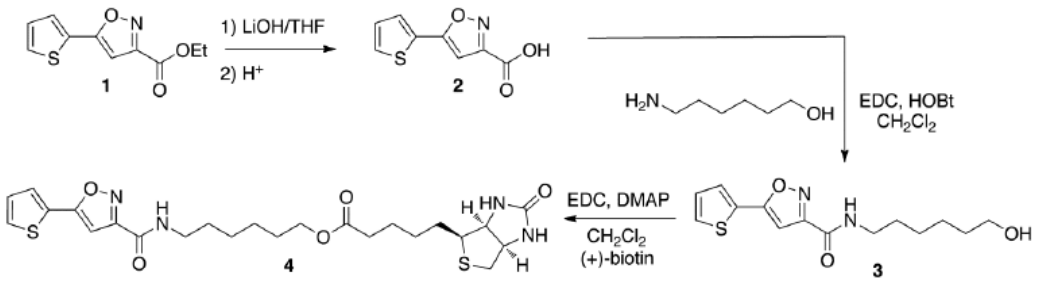
Phenylalanine repeats in the gelation domain of nucleoporin Nsp1 are responsible for the ability of the domain to form hydrogels (Figure adapted from Frey et al. (2006). They are typically surrounded by serines and glycines (right column). Repeats in the gelation domain of RNA-binding protein FUS feature tyrosines instead of phenylalanines and are also typically surrounded by serines and glycines (left column).

CHAPTER 4:

Materials and Methods

Chemical synthesis of biotinylated isoxazole

To a solution of ethyl 5-(thiophen-2-yl)-isoxazole-3-carboxylate (**1**) (3.8g) in THF (20mL) was added an aqueous solution of LiOH (34mL of 1M). The reaction was heated to 60°C for 4 hours and then cooled to room temperature. The reaction was extracted twice with toluene and the remaining aqueous layer acidified with citric acid to pH 3. The aqueous layer was extracted with EtOAc and then concentrated to provide the crude acid product (**2**) as a white solid (2.06g, 62%). This product was combined with EDC (2.08g), HOBt (1.47g), and 6-amino-1-hexanol (2.26g) in dichloromethane (150ml). The reaction was stirred at room temperature for 3.5 hours, concentrated and then redissolved in EtOAc (25mL). This solution was washed with water, brine and dried over anhydrous MgSO4. Concentration provided the crude product that was purified by recrystallization from methyl *tert*-butyl ether to give the 1.09g (35% yield) of the desired alcohol product (**3**).



A portion of the alcohol product (**3**) (30mg) was stirred for 3 days with (+)-biotin (57mg), EDC (51mg) and DMAP (6mg) in dichloromethane (2mL). After concentration to dryness, the crude reaction was partitioned between EtOAc (5mL) and water (2mL). The resulting precipitate that formed was collected by vacuum filtration and then purified by silica gel column chromatography using 10% methanol in dichloromethane as eluent. Concentration of the appropriate fractions provided the desired product that was further purified to >98% purity by recrystallization from EtOAc to yield 14mg (26% yield) of the biotinylated isoxazole (**4**). The product was fully characterized by 1H NMR and LC/MS (electrospray ionization).

Biotinylated isoxazole-mediated precipitation

Biotinylated isoxazole was resuspended in DMSO as 100x stock for use at 10, 30, or 100 uM final concentration. Cells or tissue samples were homogenized into a lysis buffer containing 20 mM Tris buffer with 150 mM NaCl, 5 mM MgCl2, 20 mM β-mercaptoethanol (BME), 0.5% NP-40, 10% glycerol, 1:300 mammalian protease inhibitor (Sigma, St. Louis, MO), 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 1:100 RNasin ribonuclease inhibitor (Promega, Madison, WI, USA), and 2 mM vanadyl ribonucleoside complex (NEB, Ipswich, MA, USA). Lysates for precipitation were pre-cleared with a 14K spin for 15 minutes. Biotinylated

isoxazole was added at 1:100 and rotated gently at 4°C for 1 hour. The incubated reaction was then spun at 14K for 15 minutes to pellet th**e** precipitant. The pellet was washed twice in lysis buffer and resuspended in 2x SDS loading buffer for protein analysis. For identification by mass spectrometry, proteins were precipitated from cell lysates and the resuspended pellets were resolved on 4-12% Tris-Glycine gradient gels (Invitrogen, USA). The entire lane was then excised into 10 or more fractions for mass spectrometry analysis.

Mass spectrometric analysis

Protein bands on the SDS-PAGE gel were de-stained and in-gel digested with sequencing grade trypsin (10 ng/μL trypsin, 50 mM ammonium bicarbonate, pH 8.0) overnight at 37°C. Peptides were extracted with 5% formic acid/50% acetonitrile and 0.1% formic acid/75% acetonitrile sequentially and then concentrated to ~ 20 μl. The extracted peptides were separated by an analytical capillary column (50 μm x 10 cm) packed with 5 μm spherical C18 reversed phase material (YMC, Kyoyo, Japan). An Agilent 1100 series binary pumps system (Agilent Technologies, Santa Clara, CA, USA) was used to generate the following HPLC gradient: 0-5% B in 5 min, 5-40% B in 70 min, 40-100% B in 10 min (A = 0.2 M acetic acid in water, B = 0.2 M acetic acid /70% acetonitrile). The eluted peptides were sprayed into a LTQ mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) equipped with a nano-ESI ion source. The

mass spectrometer was operated in data-dependent mode with one MS scan followed by five MS/MS scans for each cycle. Database searches were performed on an in-house Mascot server (Matrix Science Ltd., London, UK) against IPI (International Protein Index) protein database. Methionine oxidation was set as variable modification.

Computation analysis of low complexity sequences and [GS]Y[GS] motifs

The mouse and human proteomes were downloaded from the NCBI refseq database. The fruit fly proteome was downloaded from the NCBI genome database. For each protein-coding gene, only one protein isoform (the longest one) was selected, giving rise to 19,752, 23,096 and 13,543 proteins for the human, mouse and fly proteomes respectively. Low complexity sequences were identified using the SEG program with default parameter settings (Wootton and Federhen, 1996). The length of the longest continuous low complexity segment was calculated for each protein. Proteins with such lengths no less than 35 were considered to have a long, continuous stretch of low complexity sequence. The cutoff value of 35 was selected by examining the distribution of such lengths among the human proteome. 2,345, 2,195 and 2,239 proteins were determined to have long, continuous low complexity regions for the human, mouse and fly proteomes respectively. The [GS]Y[GS] tripeptide motif was counted for each protein in the human, mouse and fly proteomes. Proteins with four or more

copies of this motif were considered to be enriched. 189, 192 and 150 proteins were determined to be enriched with the [GS]Y[GS] motif in the human, mouse and fly proteomes. The same methods used to score for both low complexity sequences and [GS]Y[GS] tripeptides in the human, mouse and fly proteomes were employed to interrogate protein samples precipitated by the b-isox compound both before and after RNase treatment, and to score for these domains in literature cited RNA granule proteins. Fisher's Exact test using "R" software (R Development Core Team, 2010) was used to determine if any set was enriched for proteins containing LC domains and [GS]Y[GS] motifs as compared to the statistics of the proteome.

Stress granule recruitment assay

A 10-cm confluent dish of U2OS cells was transfected with 10 μg of plasmid with 10 uL of Lipofectamine 2000 (Invitrogen, USA) in Opti-MEM for 6 hours for each construct. Cells were split 24 hours after transfection on to glass coverslips in 24-well plates. Stress granules were induced with 0.5mM sodium arsenite for 1 hour 36 hours after transfection. Coverslips were fixed with 4% paraformaldehyde for 10 minutes, permeabilized in 1% Triton for 15 minutes, blocked in PBTA (5% BSA, 0.2% Tween in PBS) for 30 minutes, and incubated with 1:1000 αFlag antibody (Sigma, USA) in PBTA overnight at 4C. Primary

antibody was developed with 1:1000 αmouse conjugated to AlexaFluor 488 (Invitrogen, USA) for 1 hour at room temperature. Coverslips were then mounted on to glass microscope slides with Vectashield and DAPI (Jackson, Bar Harbor, Maine, USA). Flag-FUSΔ32 plasmids were constructed by amplifying the region encoding FUS214 with Y->S mutations from bacterial expression vectors with the Flag epitope encoded in the forward primer and fusing it to the truncated C-terminal fragment lacking the last 32 amino acids by bridging PCR. The Flag-FUSΔ32 PCR product was then cloned into the multiple cloning sites of the pcDNA3.1(+) vector for mammalian expression (Invitrogen).

Protein expression and purification

DNA fragments encoding the LC domains of human FUS and other RNA-binding proteins were amplified by PCR using a cDNA library made from human U2OS cells as a template. DNA fragments encoding MS2 were amplified by PCR from the pMS2-GFP plasmid (Addgene plasmid 27121). The DNA fragments were inserted into the multiple cloning sites of the pHis, pHis-GST, pHis-mCherry, pHis-GFP, or pHis-CFP parallel vectors (Sheffield et al., 1999). The latter three plasmids were made by cloning the mCherry, GFP or CFP coding DNA fragments at the NcoI site of the pHis-parallel vector. The sequences of the resulting vectors were confirmed by DNA sequencing. All proteins were over-expressed in *E. Coli* BL21 (DE3) cells with 0.5 mM IPTG at 16°C for overnight.

Harvested cells were lysed with 0.4mg/mL lysozyme in a lysis buffer containing 50mM Tris-HCl pH7.5, 500mM NaCl, 20mM β-mercaptoethanol (BME), 1% Triton X-100 and protein inhibitor cocktail (Roche, USA) for 30 min on ice, and then sonicated. The cell lysate was centrifuged at 35,000 RPM for 1 hour. The supernatant was mixed for 30 min at 4oC with either Ni-NTA resin (Qiagen, USA) for the His-tagged proteins, or glutathion resin for the GST-tagged proteins (GE Healthcare, USA). The Ni-NTA resin was packed in a glass column and washed with a buffer containing 20mM Tris-HCl pH7.5, 500mM NaCl, 20mM imidazole, 20mM BME, and 0.1mM phenylmethylsulfonyl fluoride (PMSF). The bound proteins were eluted from the resin with a buffer containing 20mM Tris-HCl pH7.5, 500mM NaCl, 200mM imidazole, 20mM BME, and 0.1mM PMSF. The glutathione resin was washed with a buffer containing 20mM Tris-HCl pH7.5, 500mM NaCl, 20mM BME, and 0.1mM PMSF. The bound proteins were eluted with a buffer containing 20mM Tris-HCl pH7.5, 500mM NaCl, 20mM BME, 0.1mM PMSF, and 10mM glutathione. EDTA was added to a final concentration of 0.5mM to the eluted protein solutions. The purified proteins were concentrated with Amicon Ultra centrifugal filters (Millipore, USA), and then glycerol was added to the final concentration of 50%. The protein solutions were kept at -20 ˚C. The purity of the purified proteins was confirmed by SDS-PAGE, and the concentrations were determined by absorbance at UV280.

Formation of FUS hydrogel

The purified FUS LC domain linked to three different tags GST, mCherry, and GFP were dialyzed against a gelation buffer containing 20mM Tris-HCl pH7.5, 200mM NaCl, 20mM BME, 0.5mM EDTA and 0.1mM PMSF for overnight. The protein solutions were concentrated to roughly 60mg/ml. After concentration, the protein solutions (~50 μl) were filled into 2-cm silicon tubes (3.6 mm diameter) followed by sealing the tube ends with parafilm. The tubes were incubated at 4˚C for 2 days. The formed hydrogels were carefully squeezed out from the tubes for photography. Similar methods were employed to form mCherry:hnRNPA2 and GFP:hnRNPA2 hydrogels. For hydrogel binding assays, the dialyzed mCherry:FUS and mCherry:hnRNPA2 protein solutions (~300 μl) were sonicated 10 seconds at a 12% power level on a Fisher Scientific Sonic Dismembrater Model 500. After centrifugation, a 0.5 μl droplet of the supernatant was deposited onto a glass-bottomed microscope dish (MatTek, MA, USA). The dish was sealed with Parafilm and incubated overnight at room temperature.

Hydrogel binding assays

Glycerol stocks of the purified GFP-fusion LC domains were diluted in 1mL of the gelation buffer at the final concentration of 1 μM of the protein. This solution was poured into the hydrogel dish so as to soak the hydrogel droplets in the GFP solution. The hydrogel dish was incubated at 4˚C for up to a week. For the MS2

gel retention assay, 2uM CFP-tagged constructs were incubated with 1uM Cy5.5 hairpin RNA in gel shift buffer for 30 minutes on ice. The protein-RNA mixture was then added to the mCherry-FUS214 gel droplet and allowed to equilibrate at least overnight at 4oC. Horizontal sections of the soaked hydrogel droplets were scanned with both the mCherry, GFP, and CFP excitation wavelengths on Zeiss LSM510 or Leica TCS SP5 confocal microscopes. The montage images were made by the program ImageJ (NIH, USA).

SILAC for mass spectrometry

Two separate populations of U2OS cells were maintained and fed either heavy isotope containing lysine or normal lysine (Thermo Pierce, Rockford, IL, USA). Cells were passaged seven times to fully incorporate the labeled amino acids. Cells were then lysed with a buffer containing the following: 20mM Tris buffer with 150mM NaCl, 5mM MgCl2, 20mM β-mercaptoethanol (BME), 0.5% NP-40, 10% glycerol, 1:300 protease inhibitor (Sigma, St. Louis, MO), 0.1mM phenylmethylsulfonyl fluoride (PMSF), 1:1000 RNasin ribonuclease inhibitor (Promega, Madison, WI, USA), and 2mM vanadyl ribonucleoside complex (NEB, Ipswich, MA, USA). Heavy lysate was either treated with 100ug/mL RNase A (Sigma) overnight at 4oC or left untreated as was light lysate. The samples were exposed to 100uM of the b-isox chemical and rotated at 4oC for 1 hour. The pellets were then collected, washed, and resuspended in 2X SDS loading buffer.

Heavy and light samples were mixed 1:1 and gel purified by SDS-PAGE. The full lane of the gel was excised and submitted for mass spec identification.

Computational analysis of low complexity sequences and *pumilio* motif

Low complexity sequences were identified using the SEG program with default parameter settings (Wootton and Federhen, 1996). The length of the longest continuous low complexity segment was calculated for each protein, and proteins with lengths of at least 35 amino acids were considered to have a low complexity domain, selected as the cutoff value as described in Kato et al., 2012. Fisher's Exact test using the "R" software (R Development Core Team, 2010) was used to determine whether any set of proteins was enriched with polypeptides containing LC domains compared with the statistics of the mouse and human proteomes. mRNA and 3'UTR sequences of b-isox pellet- and supernatant enriched RNAs were analyzed for the *pumilio* binding motif 5'-UGUANAUA-3'. Fisher's Exact test using the "R" software was used to determine whether any set of RNAs was enriched with RNAs containing at least one *pumilio* motif compared with the statistics of the mouse and human transcriptomes.

RNA-Seq

100 μM b-isox compound was used to precipitated RNAs from mouse brain and U2OS cell lysates as described in . Brains were harvested from 12-14 week old

male mice that were sacrificed according to the guidelines of the Institutional Animal Care and Use Committee (IACUC). Total RNA was purified from both the pellet and supernatant fractions from the isoxazole precipitation experiments using RNA-Stat 60 (Tel-test, Friendswood, TX, USA), two phenol-chloroform extractions, and precipitation in 100% isopropanol with 70% ethanol wash. For preparing cDNA libraries for high throughput sequencing, we used the Illumina mRNA-Seq kit as per manufacturers' instruction (Illumina, San Diego, CA, USA). Briefly, poly-A+ RNA was isolated using magnetic oligo-dT beads, then fragmented using buffer containing divalent cation. Invitrogen reverse transcriptase SuperScript II was used for first strand synthesis, and followed manufacture protocol for second strand synthesis, end repair, adenylation, and adaptor ligation. Ligation reactions were either purified using magnetic beads or gel purified with 2% MetaPhore agarose (Lonza, Basel, Switzerland) to isolate fragments in the 300 base pair range. Linear amplification was performed for 15 cycles to generate final cDNA libraries for RNA-Seq. Library quality was checked with the Agilent 2100 BioAnalyzer (Agilent, Santa Clara, CA, USA) and quantified for DNA concentration with Quant-iT PicoGreen reagent in a 94-well plate format (Promega). 10 μl of 10 nM sample was sequenced on the Illumina HiSeq2000 platform.

Analysis of the isoxazole precipitation RNA-Seq experiment from mouse brain and U2OS cell lysates was carried out using the CLC Bio Genomic

Workbench (CLC Bio, Aarhus, Denmark). Sequencing reads that passed the quality control filters were aligned with the appropriate genome and annotated with reference gene names and exon/intron junctions. Genes with reads that had p-values < 0.05 were removed from the annotated data and assembled in pairwise comparisons to analyze the fold enrichment of mRNAs in the b-isox pellet versus the supernatant. Genes were annotated with gene ontology categories using the CLC Bio software, and hypergeometric distribution was used to determine whether GO categories were overrepresented in RNAs enriched at least 2-fold enriched in either the pellet or supernatant samples.

Analysis of the hydrogel retention RNA-Seq and the FUS/EWS shRNA knockdown were analyzed in the following manner. The data was analyzed using recent, open-source RNA sequencing tools. The RNA-Seq reads (in FASTQ format) were aligned to the UCSC hg19 assembly of human reference sequence for human samples and to the UCSC mm9 assembly of mouse reference sequence for mouse samples using Tophat (Trapnell et al., 2009). Then the aligned reads were assembled into transcripts and their abundances were estimated using Cufflinks (Trapnell et al., 2010). The relative abundances of transcripts were measured by the normalized read counts known as Reads Per Kilobase of exon per Million mapped reads (RPKM) (Mortazavi et al., 2008). Various quality control checks were in place during analysis to ensure that the data were of

sufficient quality to merit further analysis. Since transcripts abundances were of approximate Poisson distribution, log2 transformations were applied to transcripts abundances (RPKM+1) before comparisons among different groups. Hypergeometric tests were used to determine the statistical significance of the overlaps among gene lists.

MS2 gel shift

MS2 substrate RNA 5'-UAGAAAACAUGAGGAUCACCCAUGUCUGCAG-3' was synthesized with either 6-FAM covalently linked to the 3' end for gel shift or Cy5.5 covalently linked to the 5' end (IDT, Coralville, IA, USA). Hairpin RNA was prepared by denaturing at 70oC for 5 minutes then cooling down for at least 1 minute to allow folding. For gel shift assays, 2uM CFP-tagged constructs were incubated with 1uM 6-FAM-labelled RNA in gel shift buffer containing 20mM Tris-HCl pH 7.5, 50mM KCl, 5mM MgCl2, 0.5% NP-40, 20mM BME, 1:500 RNasin, 0.25mg/mL yeast tRNA for 30 minutes on ice. 6x DNA loading dye was added to the protein-RNA mixture and the reaction was run on 5% native acrylamide gel at 120V for 1.5 hours. The gel was then scanned on a Typhoon scanner using the Cy2 filter to detect the 6-FAM signal.

Lentiviral shRNA-mediated knockdown

GIPz lentivirus shRNAmir constructs to knockdown FUS and EWS were obtained from Open Biosystems (Thermo Scientific, Waltham, MA, USA). For FUS, the shRNA sequence used was 5'-AGGATAATTCAGACAACAA -3' cat# V3LMM_450383 and for EWS, the shRNA sequence used was 5'-AGCAGAGTAGCTATGGTCA-3' cat# V3LHS_376291. To generate lentivirus, near confluent 10-cm plates of HEK 293FT cells were transiently transfected with 5 ug of pGIPz shRNA construct each and 5ug viral packaging vectors (1.6 ug pMD2.G, 2.3 ug psPAX2) with 10uL lipofectamine 2000 (Invitrogen) in Opti-MEM. Media was replaced with DMEM with 10% FBS after 6 hours of transfection and collected 36 hours after transfection. The viral media was filtered through 0.45um low protein binding filter and 10ug/mL polybrene was added. For viral transduction, 5mL of viral media and 5mL of Opti-MEM was added to confluent 15-cm plates of U2OS cells. Transduction efficiency was confirmed by the presence of GFP-positive cells. Knockdown efficiency was confirmed by both qPCR and Western blot using rabbit polyclonal antibodies for FUS cat# A300-294A and EWS cat# A300-417A (Bethyl Labs, Montgomery, TX, USA).

BIBLIOGRAPHY

Ader et al. Amyloid-like interaction within nucleoporin FG hydrogels. *PNAS* 107: 6281 (2010).

Ainger et al. Transport and localization of exogenous myelin basic protein mRNA microinjected into oligodendrocytes. *The Journal of Cell* Biology 123: 431 (1993).

Aizer et al. The dynamics of mammalian P body transport, assembly, and disassembly in vivo. *Mol Biol Cell* 19: 4154 (2008).

Akao et al. The RCK/p54 candidate proto-oncogene product is a 54-kilodalton DEAD box protein differentially expressed in human and mouse tissues. *Cancer Res* 55: 3444 (1995).

Alexandra et al. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* 11: 717 (2005).

Amikura et al. Presence of mitochondria-type ribosomes outside mitochondria in germ plasm of *Drosophila* embryos. *PNAS* 98: 9133 (2001).

Amiri et al. An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in *C. elegans. Development* 128: 3899 (2001).

Anderson and Kedersha. RNA granules. *JCB* 172: 803 (2006).

Anderson and Kedersha. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol* 10: 430 (2009).

Andrei et al. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* 11: 717 (2005).

Andressi and Riccio. To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends in Cell Biology* 19: 465 (2009).

Arkov and Ramos. Building RNA-protein granules: insight from the germline. *Trends in Cell Biology* 20: 482 (2010).

Bashirullah et al. RNA localization in development. *Annu Rev Biochem* 67: 335 (1998).

Bannai et al. An RNA-interacting protein, SYNCRIP (heterogeneous nuclear ribonuclear protein Q1/NSAP1) is a component of mRNA granule transported with inositol 1,4,5-trisphosphate receptor type 1 mRNA in neuronal dendrites. *JBC* 279: 53427 (2004).

Barbee et al. Staufen- and FMRP- containing neuronal RNPs are structurally and functionally related to somatic P bodies. *Neuron* 52: 997 (2006).

Berleth et al. The role of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO J.* 7: 1749 (1988).

Böckers et al. Differential expression and dendritic transcript localization of Shank family members: identification of a dendritic targeting element in the 3' untranslated region of Shank1 mRNA. *Mol Cell Neurosci.* 26: 182 (2004).

Bosco et al. Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum Mol Gen* 19: 4160 (2010).

Boswell and Mahowald. Tudor, a gene required for assembly of the germ plasm in Drosophila melanogaster. *Cell* 43: 97 (1985).

Brangwynne et al. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* 324: 1729-32 (2009).

Broadus et al. Staufen-dependent localization of *prospero* mRNA contributes to neuroblast daughter-cell fate. *Nature* 391: 792 (1998).

Buchan et al., Stress-specific composition, assembly and kinetics of stress granules in *Saccharomyces cerevisiae*. *JCS* 124: 228 (2011).

Buchan and Parker. Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* 36: 932 (2009).

Burgin et al. *In situ* hybridization histochemistry of Ca2+/Calmodulin-dependent protein kinase in developing rat brain. *J. Neurosci* 10: 1788 (1990).

Caceres et al. Immunocytochemical localization of actin and microtubule-associated protein MAP2 in dendritic spines. *PNAS* 80: 1738 (1983).

Calvio et al. Identification of hnRNP P2 as TLS/FUS using electrospray mass spectrometry. *RNA* 1: 724 (1995).

Campenot and Eng. Protein synthesis in axons and its possible functions. *Journal of Neurocytology* 29: 793 (2000).

Cassola et al. Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. *Molecular Microbiology* 65: 655 (2007).

Chalupníková et al. Recruitment of the RNA helicase RHAU to stress granules via a unique RNA-binding domain. *JBC* 253: 35186 (2008).

Chartrand and Singer. RNP localization and transport in yeast. *Annu. Rev. Cell Dev. Biol.* 17: 297 (2001).

Chu and Rana. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PloS Biology* 4: 1122 (2006).

Coller et al. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with boththe decapping and deadenylase complexes. *RNA* 7: 1717 (2001).

Colombrita et al. TDP-43 is recruited to stress granules in conditions of oxidative insult. *Journal of Neurochemistry* 111: 1051 (2009).

Comery et al. Abnormal dendritic spines in fragile X knockout mice: maturation and pruning deficits. *PNAS* 94: 5401 (1997).

Cougot et al. Cytoplasmic foci are sites of mRNA decay in human cells. *JCB* 165: 31 (2004).

Cougot et al. Dendrites of mammalian neurons contain specialized P-body-like structures that respond to neuronal activation. *J. Neurosci* 28: 13793 (2008).

Crino et al. Presence and phosphorylation of transcription factors in developing dendrites. *PNAS* 95: 2313 (1998).

Crozat et al. Fusion of CHOP to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* 363: 640 (1993).

Dash et al. Injection of the cAMP-responsive element into the nucleus of *Aplysia* sensory neurons blocks long-term facilitation. *Nature* 345, 718 (1990).

Davis et al. Protein synthesis and memory: A review. *Psychological Bulletin* 96: 518 (1984).

Davis et al. Selective dendritic transport of RNA in hippocampal neurons in culture. *Nature* 330: 477 (1987).

Decker et al. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *JCB* 179: 437 (2007).

De Leeuw et al. The cold-inducible RNA-binding protein migrates from the nucleus to cytoplasmic stress granules by a methylation-dependent mechanism and acts as a translational repressor. *Exp Cell Res* 313: 4130 (2007).

Denning et al. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *PNAS* 100: 2450 (2003).

Dewey et al. TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Molecular and Cellular Biology* 31: 1098 (2011).

Di Nardo et al. Dendritic localization and activity-dependent translation of *Engrailed1* transcription factor. *Molecular and Cellular Neuroscience* 35: 230 (2007).

Dostie et al. A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO J.* 19: 3142 (2000).

Doyle and Kiebler. Mechanisms of dendritic mRNA transport and its role in synaptic tagging. *EMBO J.* 30: 3540 (2011).

Dreyfuss et al. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol.* 3: 195 (2002).

Driever et al. Autonomous determination of anterior structures in the early *Drosophila* embryo by the bicoid morphogen. *Development* 109: 811 (1990).

Dunker et al. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* 19: 26 (2001).

Dynes and Steward. Dynamics of bidirectional transport of Arc mRNA in neuronal dendrites. *Journal of Comparative Neurology* 500: 433 (2007).

Eberwine et al. Transcription factors in dendrites: dendritic imprinting of the cellular nucleus. *Results Probl Cell Differ*. 34: 57 (2001).

Eddy. Germ plasm and the differentiation of the germ cell line. *Int. Rev. Cytol.* 43: 229–280 (1975).

Elvira et al. Characterization of an RNA granule from developing brain. *Molecular & Cellular Proteomics* 5.4: 635 (2006).

Ephrussi and Lehmann. Induction of germ cell formation by oskar. *Nature* 358: 387 (1992).

Ephrussi et al., Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell* 66: 37 (1991).

Eyastathioy et al. A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Mol Biol Cell* 13: 1338 (2002).

Eyastathioy et al. The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA* 9: 1171 (2003).

Falley et al. Shank1 mRNA: Dendritic transport by kinesin and translational control by the 5' untranslated region. *Traffic* 10: 844 (2009).

Feng et al. Fragile X mental retardation protein: nucleocytoplasmic shuttling and association with somatodendritic ribocomes. *J. Neurosci.* 17: 1539 (1997).

Ferraiuolo et al. A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *JCB* 170: 913 (2005).

Ferrandon et al. Staufen protein associates with the 3'UTR of *bicoid* mRNA to form particles that move in a microtubule-dependent manner. *Cell* 79: 1221 (1994).

Fraser and Doudna. Structural and mechanistic insights into hepatitis C viral translation. *Nature Reviews Microbiology* 5: 29 (2007).

Frey et al. FG –rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science* 314: 815 (2006).

Frey and Görlich. FG/FxFG as well as GLFG repeats form a selective permeability barrier with self-healing properties. *EMBO J.* 28: 2554 (2009).

Frey and Görlich. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complex. *Cell* 130: 512 (2007)

Fujii et al. The RNA binding protein TLS is translocated to dendritic spines by mGluR5 activation and regulates spine morphology. *Current Biology* 15: 587 (2005).

Fujii and Takumi. TLS facilitates transport of mRNA encoding an actin-stabilizing protein to dendritic spines. *Journal of Cell Science* 118: 5755 (2005).

Fukuda et al. hnRNP K interacts with RNA binding motif protein 42 and functions in the maintenance of cellular ATP level during stress conditions. *Genes to Cells* 14: 113 (2009).

Fuller-Pace. DexD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research* 34: 4206 (2006).

Gal et al. Nuclear localization sequence of FUS and induction of stress granules by ALS mutants. *Neurobiology of Aging* 32: 2323e27 (2011).

Gallo et al. Cytoplasmic partitioning of P granule components is not required to specify the germline in C. elegans. *Science* 330: 1685-9 (2010).

Garner et al. Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites. *Nature* 336: 674 (1988).

Gavis et al. Localization of *nanos* RNA controls embryonic polarity. *Cell* 71: 301 (1992).

Gavis et al. Identification of *cis*-acting sequences that control *nanos* RNA localization. *Developmental Biology* 176: 36 (1996).

Gilks et al. Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* 15: 5383 (2004).

Gingras et al. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annual Review of Biochemistry* 68: 913 (1999).

Goetze et al. The brain-specific double-stranded RNA-binding protein Staufen2 is required for dendritic spine morphogenesis. *JCB* 172: 221 (2006).

Goulet et al. TDRD3, a novel Tudor domain-containing protein, localizes to cytoplasmic stress granules. *Hum Mol Gen* 17: 3055 (2008).

Grooms et al. Activity bidirectionally regulates AMPA receptor mRNA abundance in dendrites of hippocampal neurons. *J. Neurosci* 26: 8339 (2006).

Guil et al. hnRNP A1 relocation to the stress granules reflects a role in the stress response. *Molecular and Cellular Biology* 26: 5744 (2006).

Hafner et al. PAR-CliP – a method to identify transcriptome-wide the binding sites of RNA binding proteins. *J Vis Exp* 41 (2010).

Han et al. Functional diversity of the hnRNPs: Past, Present, and Perspectives. *Biochem J.* 430: 379 (2010).

Harris et al. Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C. *Development* 128: 2823 (2001).

Hay et al. Identification of a component of Drosophila polar granules. *Development* 103: 625 (1988a).

Hay et al. A protein component of Drosophila polar granules is encoded by vasa and has extensive sequence similarity to ATP-dependent helicases. *Cell* 55: 577 (1988b).

Hay et al. Localization of vasa, a component of Drosophila polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity. *Development* 109: 425 (1990).

Hegner. Effects of removing the germ-cell determinants from the eggs of some chrysomelid beetles. Preliminary report. *Biol. Bull.* 16: 19–26 (1908).

Hegner. Experiments with chrysomelid beetles. III. The effects of killing parts of the eggs of Leptinotarsa decemlineata. *Biol. Bull.* 20: 237–251 (1911).

Herb et al. Prominent dendritic localization in forebrain neurons of a novel mRNA and its product dendrin. *Molecular and Cellular Neuroscience* 8: 367 (1997).

Hilliker et al. The DEAD-box protein Ded1 modulates translation by the formation and resolution of an eIF4F-mRNA complex. *Molecular Cell* 43: 962 (2011).

Hoell et al. RNA targets of wild-type and mutant FET family proteins. *Nat. Struct. Mol. Biol.* (2011)

Huttenlocher et al. Adhesion in cell migration. *Current Opinion in Cell Biology* 7: 697 (1995).

Hüttelmaier et al. Spatial regulation of β-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* 438: 512 (2005).

Huynh et al. The *Drosophila* hnRNPA/B homolog, Hrp48, is specifically required for a distinct step in *osk* mRNA localization. *Developmental Cell* 6: 625 (2004).

Illmensee and Mahowald. Transplantation of posterior polar plasm in Drosophila. Induction of germ cells at the anterior pole of the egg. *PNAS* 71: 1016 (1974).

Ingelfinger et al. The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci. *RNA* 8: 1489 (2002).

Ivanov et al. Disruption of microtubules inhibits cytoplasmic ribonucleoprotein stress granule formation. *Exp Cell Res* 290: 227 (2007).

Janknecht. EWS-ETS oncoproteins: the linchpins of Ewing tumors. *Gene* 363: 1 (2005).

Job and Eberwine. Localization and translation of mRNA in dendrites and axons. *Nature Reviews Neuroscience* 2: 889 (2001).

Kanai et al. Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron* 43: 513 (2004).

Kashikawa et al. Mitochondrial small ribosome RNA is a component of germinal granules in *Xenopus* embryos. *Mech Dev* 101: 71 (2001).

Kausik et al. Aplysia CPEB can form prion-like multimers in sensory neurons that contribute to long-term facilitation. *Cell* 140: 421 (2010).

Kayali et al. Prolonged translation arrest in reperfused hippocampal comu Ammonis 1 is mediated by stress granules. *Neuroscience* 134: 1223 (2005).

Kedersha et al. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *JCB* 169: 871 (2005).

Kedersha et al. RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF2a to the assembly of mammalian stress granules. *JCB* 147: 1431 (1999).

Kedersha et al. Evidence that ternary complex (eIF2-GTP-tRNAi^{Met})-deficient preinitiation complex are constituents of mammalian stress granules. *Mol Biol Cell* 13: 195 (2002).

Kedersha et al. Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *JCB* 151: 1257 (2000).

Kelley et al. Initial organization of the *Drosophila* dorsoventral axis depends on an RNA-binding protein encoded by the *squid* gene. *Genes & Dev.* 7: 948 (1993).

Keryer-Bibens et al. Tethering of proteins to RNAs by bacteriophage proteins. *Biol. Cell* 100: 125 (2008).

Kiebler and DesGroseillers. Molecular insights into mRNA transport and local translation in the mammalian nervous system. *Neuron* 25: 19 (2000).

Kiebler et al. The mammalian Staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: implications for its involvement in mRNA transport. *J Neurosci* 19: 288 (1999).

Kim-Ha et al. *oskar* mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* 66: 23 (1991).

Kim-Ha et al. Multiple RNA regulatory elements mediate distinct steps in localization of *oskar* mRNA. *Development* 119: 169 (1993).

King et al. Polarizing genetic information in the egg: RNA localization in the frog oocyte. *BioEssays* 21: 546 (1999).

Kislauskis et al. Isoform-specific 3'-untranslated sequences sort α-cardiac and β-cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J. Cell. Biol.* 127: 441 (1993).

Kislauskis et al. Sequences responsible for intracellular localization of b-actin messenger RNA also affect cell phenotype. *J. Cell Biol.* 127: 441 (1994).

Kislauskis et al. b-actin messenger RNA localization and protein synthesis augment cell motility. *J. Cell Biol.* 136: 1263 (1997).

Kloc et al. RNA localization and germ cell determination in Xenopus. *Int Rev Cytol.* 203: 63 (2001).

Knowles et al. Translocation of RNA granules in living neurons. *The Journal of* Neuroscience 16: 7812 (1996).

Kobayashi and Okada. Restoration of pole-cell-forming ability to u.v.-irradiated Drosophila embryons by injection of mitochondrial lrRNA. *Development* 107: 733 (1989).

Köhrmann et al. Microtubule-dependent recruitment of Staufen-green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons. *Molecular Biology of the* Cell 10: 2945 (1999).
Kole et al. Intracellular mechanics of migrating fibroblasts. *MboC* 16: 328 (2005).

Kramer et al. Prion-like propagation of cytosolic protein aggregates. *Prion* 3: 206 (2009).

Krichevsky and Kosik. Neuronal RNA granules: a link between RNA localization and stimulation-dependent translation. *Neuron* 32: 683 (2001).

Kwiatkowski et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323: 1205 (2009).

Kye et al. Somatodendritic microRNAs identified by laser capture and multiplex RT-PCR. *RNA* 13: 1224 (2007).

Lagier-Tourenne and Cleveland. Rethinking ALS: The FUS about TDP-43. *Cell* 136: 1001 (2009).

Lai et al. The DEAD-box RNA helicase DDX3 associates with export mRNPs as well as TAP and participates in translational control. *Mol Biol Cell* 19: 3847 (2008).

Lécuyer et al. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131: 174 (2007).

Lehmann and Nusslein-Volhard. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in Drosophila. *Cell* 47: 141 (1986).

Linder. Dead-box proteins: a family affair – active and passive players in RNP-remodeling. *Nucleic Acids Research* 34: 4168 (2006).

Linder and Jankowsky. From unwinding to clamping – the DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology* 12: 505 (2011).

Link et al. Somatodendritic expression of an immediate early gene is regulated by synaptic activity. *PNAS* 92: 5734 (1995).

Liu et al. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nature Cell Biology* 7: 719 (2005).

Liu-Yesucevitz et al. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *Plos One* 5: e13250 (2010).

Long et al. Mating type switching in yeast controlled by asymmetric localization of *ASH1* mRNA. *Science* 277: 383 (1997).

Lyford et al. Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* 14: 433 (1995).

Mahowald. Fine structure of pole cells and polar granules in Drosophila melanogaster. *J. Exp. Zool.* 151: 201–215 (1962).

Mahowald. Polar granules of Drosophila: II. Ultrastructural changes during early embryogenesis. *J. Exp. Zool.* 167: 237–262 (1968).

Mahowald. Polar granules of Drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis. *J. Exp. Zool* 176: 345 (1971).

Mallardo et al. Isolation and characterization of Staufen-containing ribonucleoprotein particles from rat brain. *PNAS* 100: 2100 (2003).

Mangiardi et al. Progression of hair cell ejection and molecular markers of apoptosis in the avian cochlea following gentamicin treatment. *J. Comp. Neurol.* 475: 1-18 (2004).

Marión et al., A human sequence homologue of Staufen is an RNA-binding protein that is associated with polysomes and localizes to the rough endoplasmic reticulum. *Mol Cell Biol* 19: 2212 (1999).

McEvoy et al. Cytoplasmic polyadenylation element binding protein 1-mediated mRNA translation in Purkinje neurons is required for cerebellar long-term depression and motor coordination. *J. Neurosci.* 27: 6400 (2007).

Megosh et al. The role of PIWI and the miRNA machinery in Drosophila germline determination. *Curr Biol.* 16: 1884 (2006).

Michelitsch and Weissman. A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *PNAS* 97: 11910 (2000).

Micklem et al. Distint roles of two conserved Staufen domains in *oskar* mRNA localization and translation. *EMBO J.* 19: 1366-1377 (2000).

Miki et al. The role of mammalian Staufen on mRNA traffic: a view from its nucleocytoplasmic shuttling function. *Cell Structure and Function* 30: 51 (2005).

Mili et al. Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* 453: 115 (2008).

Miller et al. Disruption of dendritic translation of CaMKIIalpha impairs stabilization of synaptic plasticity and memory consolidation. *Neuron* 24: 36 (2002).

Mingle et al. Localization of all seven messenger RNAs for the actin-polymerization nucleator Arp2/3 complex in the protrusions of fibroblasts. *J. Cell Sci.* 118: 2425 (2005).

Minshall and Standart. The active form of Xp54 RNA helicase in translational repression is an RNA-mediated oligomer. *Nucleic Acids Res* 32: 1325 (2004).

Miyashiro et al. On the nature and differential distribution of mRNAs in hippocampal neurites: Implications for neuronal functioning. *PNAS* 91: 10800 (1994).

Mortazavi et al. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5, 621-628 (2008).

Moore. From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309: 1514 (2005).

Muddashetty et al. Dysregulated metabotropic glutamate receptor-dependent translation of AMPA receptor and postsynaptic density-95 mRNAs at synapse in a mouse model of Fragile X syndrome. *J. Neurosci* 27: 5338 (2007).

Murata and Wharton. Binding of pumilio to maternal hunchback mRNA is required for posterior patterning in Drosophila embryos. *Cell* 80: 747 (1995).

Nakamura et al. Requirement for a noncoding RNA in Drosophila polar granules for germ cell establishment. *Science* 274: 2075 (1996).

Neuman-Silberberg and Schüpbach. The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGFα-like protein. *Cell* 75: 165 (1993).

Nobes and Hall. Rho GTPases control polarity, protrusion, and adhesion during cell movement. *JCB* 144: 1235 (1999).

Nonhoff et al. Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with p-bodies and stress granules. *MboC* 18: 1385 (2007)

Nover et al. Cytoplasmic heat shock granules are formed from precurson particles and are associated with a specific set of mRNAs. *Molecular and Cellular Biology* 9: 1298 (1989).

Nover et al. Formation of cytoplasmic heat shock granules in tomato cell cultures and leaves. *Molecular and Cellular Biology* 3: 1648 (1983).

Okabe et al. Real time monitoring of endogenous cytoplasmic mRNA using linear antisense 2'-O-methyl RNA probes in living cells. *Nucl. Acids Res.* 39: e20 (2011).

Okada et al. Restoration of fertility in sterilized Drosophila eggs by transplantation of polar cytoplasm. *Dev Bio* 37: 43 (1974).

Paige et al. RNA mimics of green fluorescent protein. *Science* 333: 642 (2011).

Palacios and St. Johnston. Getting the message across: the intracellular localization of mRNAs in higher eukaryotes. *Annu. Rev. Cell Dev. Biol.* 17: 569 (2001).

Parsyan et al. mRNA helicases: the tacticians of translational control. *Nat Rev Mol Cell Biol* 12: 235 (2011).

Pontius, 1993. Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends Biochem. Sci.* 18: 181 (1993).

Pitt et al. P granules in the germ cells of Caenorhabditis elegans adults are associated
with clusters of nuclear pores and contain RNA. *Dev Biol*. 219: 315 (2000).

Plath et al. Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. *Neuron* 52: 437 (2006).

Quaresma et al. Human hnRNP Q re-localizes to cytoplasmic granules upon PMA, thapsigargin, arsenite and heat-shock treatments. *Exp Cell Res* 315: 968 (2009).

R Development Core Team (2010). R: A language and environment for statistical computing (Vienna, Austria: R Foundation for Statistical Computing).

Ramos et al. RNA recognition by a Staufen double-stranded RNA-binding domain. *EMBO J.* 19: 997 (2000).

Rebagliati et al. Identification and cloning of localized maternal RNAs from Xenopus eggs. *Cell* 42: 769 (1985).

Reijns et al. A role for Q/N-rich aggregation-prone regions in P-body localization. *JCS* 121: 2463 (2008).

Riggi et al. Sarcomas: genetics, signalling, and cellular origins. Part I: The fellowship of TET. *J Pathol* 213: 4 (2007).

Roegiers and Jan. Staufen: a common component of mRNA transport in oocytes and neurons? *Trends in Cell Biology* 10: 220 (2000).

Rook et al. CaMKIIa 3' untranslated region-directed mRNA translocation in living neurons: visualization by GFP linkage. *J. Neurosci.* 20: 6385 (2000).

Sadek et al. Cardiogenic small molecules that enhance myocardial repair by stem cells. *PNAS* 105: 6063 (2008).

Sen et al. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nature Cell Biology* 7: 633 (2005).

Schneider et al. Small-molecule activation of neuronal cell fate. *Nature Chemical Biology* 4: 408 (2008).

Schratt et al. A brain-specific microRNA regulates dendritic spine development. *Nature* 439: 283 (2006).

Schröder et al. Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? *Biochem Soc Trans* 39: 679 (2011).

Schupbach and Wieschauss. Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo. *Roux's Archives of Developmental Biology* 195: 302 (1986).

Shih et al. Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene* 27: 700 (2008).

Si et al. A neuronal isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in *Aplysia*. *Cell* 115: 893 (2003).

Si et al. A neuronal isoform of the *Aplysia* CPEB has prion-like properties. *Cell* 115: 879 (2003b).

Si et al. *Aplysia* CPEB can form prion-like multimers in sensory neurons that contribute to long-term facilitation. *Cell* 140: 421 (2010).

Siegel et al. A functional screen implicated microRNA-138-dependent regulation of the depalmitoylation enzyme APT1 in dendritic spine morphogenesis. *Nature Cell Biology* 11: 705 (2009).

Souquere et al. Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. *J Cell Science* 122: 3619 (2009).

Sun et al. Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS. *PloS Biology* 9: e1000614 (2011).

St Johnston et al. Staufen, a gene required to localize maternal RNAs in the Drosophila egg. *Cell* 66: 51 (1991).

St Johnston et al. A conserved double-stranded RNA-binding domain. *PNAS* 89: 10979 (1992).

Steward and Levy. Preferential localization of polyribosomes under the base of dendritic spines in granule cells of the dentate gyrus. *J. Neurosci.* 2: 284 (1982).

Takizawa et al. Actin-dependent localization of an RNA encoding a cell-fate determinant in yeast. *Nature* 389: 90 (1997).

Teixeira et al. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* 11: 371 (2005).

Tian et al. A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. *Cell* 67: 629 (1991).

Timmusk et al. Dendritic localization of mammalian *neuralized* mRNA encoding a protein with transcription repression activities. *Molecular and Cellular Neuroscience* 20: 649 (2002).

Thomas et al. Mammalian Staufen1 is recruited to stress granules and impairs their assembly. *JCS* 122: 563 (2009).

Thomson and Lasko. Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning. *Genesis* 40: 164 (2004).

Thomson and Lasko. Isolation of new polar granule components in Drosophila reveals P body and ER associated proteins. *Mechanisms of Development* 125: 865 (2008).

Thomson et al. Isolation of new polar granule components in *Drosophila* reveals P body and ER associated proteins. *Mechanisms of Development* 125: 865 (2008).

Torre and Steward. Demonstration of local protein synthesis within dendrites using a new cell culture system that permits the isolation of living axons and dendrites from their cell bodies. *J. Neurosci* 12: 762 (1992).

Tourrière et al. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *JCB* 160: 823 (2003).

Trapnell et al. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105-1111 (2009).

Trapnell et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28, 511-515 (2010).

Tsai and Nussinov. Gene-specific transcription activation via long-range allosteric shape-shifting. *Biochem J.* 439: 15 (2011).

Tyagi. Imaging intracellular RNA distribution and dynamics in living cells. *Nature Methods* 6: 331 (2009).

Tyagi and Kramer. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* 14: 303 (1996).

Udan and Baloh. Implications of the prion-related Q/N domains in TDP-43 and FUS. *Prion* 5: 1 (2011).

Updike and Strome. A genomewide RNAi screen for genes that affect the stability, distribution and function of P granules in *Caenorhabditis elegans*. *Genetics* 183: 1397 (2009).

Updike and Strome. P granule assembly and function in *Caenorhabditis elegans* germ cells. *J. Andrology* 31: 53 (2010).

Updike et al. P granules extend the nuclear pore complex environment in the *C. elegans* germ line. *JCB* 192: 939 (2011).

Vagin et al. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313: 320 (2006).

Vance et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323: 1208 (2009).

Van Dijk et al. Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.* 21: 6950 (2002).

Vessey et al. A loss of function allele for murine Staufen1 leads to impairment of dendritic Staufen1-RNP delivery and dendritic spine morphology. *PNAS* 105: 16374 (2008).

Vessey et al. Mammalian Pumilio 2 regulates dendrite morphogenesis and synaptic function. *PNAS* 107: 3222 (2010).

Villacé et al. The composition of Staufen-containing RNA granules from human cells indicates their role in the regulated transport and translation of messenger RNAs. *Nucleic Acids Research* 32: 2411 (2004).

Voronina et al. RNA granules in germ cells. Cold Spring Harbor Perspectives in Biology 2011.

Voronina and Seydoux. The *C. elegans* homolog of nucleoporin Nup98 is required for the integrity and function of germline P granules. *Development* 137: 1441 (2010).

Wälde and Kehlenbach. The part and the whole: functions of nucleoporins in nucleocytoplasmic transport. *Trends in Cell Biology* 20: 461 (2010).

Wallace et al. Differential intracellular sorting of immediate early gene mRNAs depends on signals in the mRNA sequence. *J. Neurosci* 18: 26 (1998).

Wang et al. Localization of Kv1.1 and Kv1.2, two K channel proteins, to synaptic terminals, somata, and dendrites in the mouse brain. *J. Neurosci.* 14: 4588 (1994).

Wickham et al. Mammalian Staufen is a double-stranded-RNA-and tubulin-binding protein which localizes to the rough endoplasmic reticulum. *Mol Cel Biol* 19: 2220 (1999).

Willett et al. Localization of ribosomes and translation initiation factors to talin/b3-integrin-enriched adhesion complexes in spreading and migrating mammalian cells. *Biology of the Cell* 102: 265 (2010).

Wooten and Federhen. Statistics of local complexity in amino acid sequence databases. *Comput Chem* 17: 149 (1993).

Xu and Powers. Nuclear pore proteins and cancer. *Semin Cell Dev Biol.* 20: 620 (2009).

Yano et al. Hrp48, a *Drosophila* hnRNPA/B homolog, binds and regulates translation of *oskar* mRNA. *Developmental Cell* 6: 637 (2004).

Yin et al. Induction of a dominant negative CREB transgene specifically blocks long-term memory in Drosophila. *Cell* 79: 49 (1994).

Yokota. Historical survey on chromatoid body research. *Acta Histochem Cytochem.* 41: 65 (2008).

Yoon and Mowry. Xenopus Staufen is a component of a ribonucleoprotein complex containing Vg1 RNA and kinesin. *Development* 131: 3035 (2004).

Zalfa et al. The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell* 112: 317 (2003).

Zamore et al. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* 3: 1421 (1997).

Zenklusen et al. Imaging real-time gene expression in living yeast. *CSH Protoc* (2007).

Zhang et al. A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* 390: 477 (1997)

Zhang et al. Small-molecule blocks malignant astrocyte proliferation and induces neuronal gene expression. *Differentiation* 81: 233 (2011).

Zheng et al., 2001. Neurotrophin-induced transport of a β-actin mRNP complex increases β-actin levels and stimulates growth cone motility. *Neuron* 31: 261 (2001).

Zinszner et al. TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. *Journal of Cell Science* 110: 1741 (1997).