Deciphering the role of liquid-liquid phase separation in oncogenic gene expression

Peter Walsh BSc MSc

This thesis is submitted to the University of Glasgow in fulfilment of the requirements of the Degree of Doctor of Philosophy

School of Cancer Sciences

College of Medical, Veterinary and Life Sciences

University of Glasgow

CRUK Beatson Institute

Garscube Estate

Glasgow

June 2023

Abstract

Liquid-liquid phase separation is a biophysiochemical process that drives the formation of membraneless organelles, termed biomolecular condensates. Stress granules are a biomolecular condensate that form in the cytoplasm in response to stress. Stress granules contain dense networks of RNA and protein. Although much research into stress granules has been conducted, little about their function in regulating translation is known. Stress granules have been implicated in pathophysiology of multiple human diseases including cancer.

Two physiologically relevant stresses in cancer were used for induction of stress granules. Oxidative stress using sodium arsenite and hypoxia. Morphologically distinct granules formed in response to hypoxia or sodium arsenite. Canonical well defined cytoplasmic foci formed in response to sodium arsenite stress with docking P-bodies on their surface. In response to hypoxia, larger nebulous stress granules deficient in docking P-bodies formed. Distinct signalling cascades regulate stress granule formation in each stress. Sodium arsenite induces stress granule formation through phosphorylation of eIF2α, whereas reduced mTOR activity and hypophosphorylation of 4E-BP1 regulates stress granule assembly in hypoxia. Proteomic analysis of stress granules revealed that the DEAD box helicases eIF4A1 and DDX6 were differentially enriched in hypoxia and sodium arsenite respectively. eIF4A1 is a known regulator of stress granule P-body interactions and inhibition of eIF4A1 reduced stress granule size and rescued P-body docking in hypoxia. Proteomic analysis also revealed an enrichment of large ribosomal subunit proteins in hypoxia induced stress granules.

RNA sequencing of stress granules revealed that mRNA localisation is largely stress dependent as distinct subsets of mRNAs with distinct sequence features are recruited to stress granules in a stress dependent manner. Integration of stress granule sequencing with ribosome profiling data revealed that stress granule mRNAs are translationally repressed in response to sodium arsenite in concordance with previously published data sets. Intriguingly stress granule mRNAs were translationally upregulated in response to hypoxia. This taken with the localisation of components of the large ribosomal subunit proteins suggests that hypoxia induced stress granules are a more translationally active compartment than what has previously been described in response to eIF2α phosphorylation.

Table of contents

Table of figures

Table of tables

Abbreviations

Activating transcription factor 4	ATF4
Adenosine triphosphate	ATP
Aminoacyl site	A-site
AMP-activated protein kinase	AMPK
Amyotrophic lateral sclerosis	ALS
ATP-binding cassette sub-family E member 1	ABCE1
Casein kinase 2	CK2
Cleavage and polyadenylation specificity factor	CPSF
DEAD-box helicase	DDX
Decapping MRNA 1A	DCP1A
Decapping mRNA 2	DCP2
Deoxyribonucleic acid	DNA
DEP Domain containing mTOR interacting protein	DEPTOR
E2F transcription factor 3	E2F3
eIF4E binding protein 1	4E-BP1
Enhancer of mRNA decapping 4	EDC4
Eukaryotic elongation factor	eEF
Eukaryotic translation initiation factor	eIF
Eukaryotic translation termination factor	eRF
Exit site	E-site
Fragile X mental retardation protein	FMRP
Gene Set Enrichment Analysis	GSEA
General control nonderepressible 2	GCN2
Glycogen Synthase Kinase 3 Beta	GSK3β
Guanine nucleotide exchange factor	GEF
Guanosine triphosphate	GTP
Heme regulated inhibitor	HRI
Hypoxia Inducible Factor 1 Subunit Alpha	HIF1α
Integrated stress response inhibitor	ISRIB
Internal ribosome entry site	IRES
Intrinsically disordered region	IDR
Kirsten rat sarcoma virus	KRAS
Mammalian target of rapamycin	mTOR
Messenger RNA	mRNA
Methyltransferase 3, N6-Adenosine-Methyltransferase Complex Catalytic Subunit	METTL3
Micro RNA	miRNA
mTOR Associated Protein, LST8 Homolog	MLST8
Non-structural protein 1	NS1
NUAK family SNF1-like kinase 1	NUAK1
Nuclear transport factor 2 like domain	NTF2L
Peptidyl site	P-site
PKR like endoplasmic reticulum kinase	PERK
Processing bodies	P-bodies
Proline-rich Akt/PKB substrate 40 kDa	PRAS40
Protein kinase R	PKR
Quantile regression-based left-censored	QRILC
Ras GTPase-activating protein-binding protein 1	G3BP1

Regulatory associated protein of mTOR	RAPTOR
Ribonucleic acid	RNA
Ribonucleoprotein	RNP
Ribosomal protein S6	RPS6
RNA recognition motif	RRM
TATA-binding protein	TBP
T-Cell-Restricted Intracellular Antigen-1	TIA1
Transcription export	TREX
Traslation efficiency	TE
Tuberous sclerosis complex	TSC
Untranslated region	UTR
Valosin-containing protein	VCP
Vascular endothelial growth factor A	VEGFA
Viral nonstructural protein 3	nsP3
YTH N6-methyladenosine RNA binding proteins	YTHDF

Acknowledgements

Firstly, I would like to thank my supervisor Professor Martin Bushell for all of his support and guidance over the past number of years. Thank you for always pushing me to do the difficult experiments, affording me the opportunities to travel to conferences to present my data and always having an open door. Thank you for fostering an open and collaborative environment in your lab and creating such a stimulating learning environment. I hope there is enough controls to satisfy the both of us!

A special word of thanks to my original supervisor Dr Ania Wilczynska. Thank you for giving me the opportunity to study for a PhD and teaching me so much in such a short space of time (especially planning!). I will never forget the sheer euphoria I felt when you told me I got the position. To my second supervisor Daniel, thank you for your input over the years and doing such a good job representing the west of Ireland on the international stage.

To Drs Tobias Schmidt, Chiara Giacomelli, Joseph Waldron and Ewan Smith, thank you for all of your help and guidance. This thesis truly would not have been possible without all of your input and teaching. More importantly thank you for your friendship and keeping me grounded during the course of my PhD. Additionally, I would like to thank all members of the Bushell lab, past and present for all of your help and input over the years, with the occasional trip to the pub thrown in for good measure.

To my three Beatson amigos, Declan, Bobby and Erin. A sincere thanks for your wonderful friendship and support over the years. Even if most of the coffee breaks were just to procrastinate in the sun, you always listened to my mutterings about RNA. I'm certain our friendships will long outlive our time at the Beatson.

To my family; my Mum, my Dad and my sisters Katie and Sarah. I'm finally finishing school (hopefully)! Thank you for all of the love and support over the course of my life and always encouraging me to go for it.

Lastly to my partner Danielle. Thank you for your unconditional love and support, tolerating me when I've had bad days or working stupid hours. You inspire me to achieve as much as possible every day and hopefully we can now start the chapter of our life together where I'm no longer a student.

Author’s declaration

I declare that, except where explicit reference is made to the contribution of others, this thesis is the rest of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Peter Walsh

1 Chapter 1: Introduction

1.1 What is gene expression?

The title of this thesis contains the term gene expression. It is important to add some context to the term gene expression to understand what it truly means. The central dogma of molecular biology describes the flow of information within a cell (Crick, 1970), (figure 1.1). DNA is transcribed to RNA which in turn is used as the template for the translation of a protein. Therefore, gene expression can be split into two fundamental cellular processes, transcription and translation. Both transcription and translation can be split into three stages. Initiation, elongation and termination. These processes will be outlined in greater detail below.

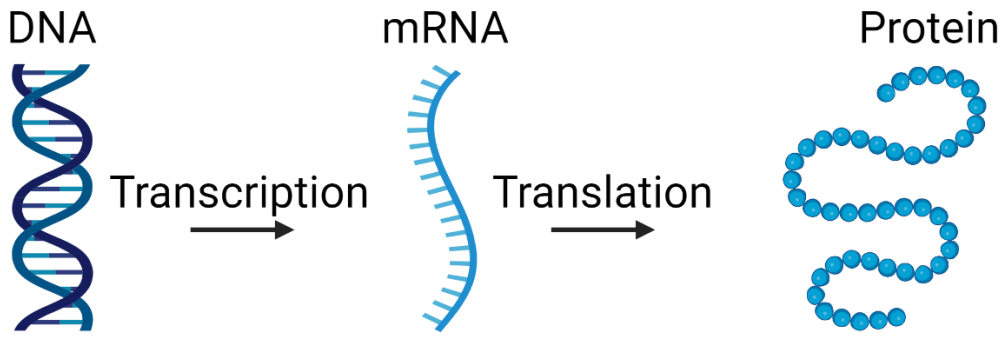


Figure 1.1 Central dogma of molecular biology:
The central dogma of molecular biology was originally proposed by Francis Crick. The central dogma describes the flow of information within a cell during gene expression. DNA is transcribed to mRNA and mRNA is translated into protein.

1.1.1 Transcription

The primary enzymes responsible for transcription are ribonucleic acid (RNA) polymerases. In eukaryotes RNA polymerase I transcribes ribosomal RNAs, including the 28- and 18S RNAs. RNA polymerase III transcribes transfer RNAs and the 5S ribosomal RNA. RNA polymerase II transcribes messenger RNAs (mRNA) in addition to other small RNAs such as micro RNAs (miRNAs) (Cramer et al., 2008). RNA polymerase II is recruited to promoter regions of DNA, which are upstream of the transcription start site by association with transcription factor proteins. General transcription factors regulate the initiation of

transcription in most eukaryotic protein coding genes. General transcription factors bind deoxyribonucleic acid (DNA) regulatory elements, such as TATA-binding protein TBP binding the TATA box which facilitates the formation of the megadalton transcription preinitiation complex (Butler and Kadonaga, 2002). The transcription preinitiation complex then melts the bound promoter region in an adenosine triphosphate (ATP) dependent manner (Holstege et al., 1997). Promoter melting results in a conformational switch of the transcriptional preinitiation complex, from a closed to open conformation and thus initiation of transcription.(He et al., 2013) RNA polymerase II dissociates from the preinitiation complex and synthesises mRNA molecules, using the reverse DNA strand as a template (Pokholok et al., 2002). Transcription termination is regulated by the recognition of a poly adenylation signal by the protein complex cleavage and polyadenylation specificity factor (CPSF) in the nascently transcribed mRNA. This facilitates the recruitment of further proteins that cleave the mRNA. Poly-A polymerase then catalyses the addition of the poly-A tail to the mRNA (Richard and Manley, 2009). Another vital RNA processing step, the addition of the 5' mRNA cap, occurs co-transcriptionally for the majority of mRNAs. The addition of the 5' m7GpppN cap is facilitated by a multi-step enzymatic process. Firstly, RNA triphosphatase cleaves the 5' end of the nascently transcribed mRNA. Guanylyl transferase then adds a guanosine monophosphate to the 5' end of the mRNA. The guanosine cap is then methylated by RNA methyl transferase at the N-7 position (Moteki and Price, 2002, Shatkin and Manley, 2000).

1.1.2 Post transcriptional processing of mRNAs and nuclear export.

After polyadenylation and capping, mRNAs need to be further processed before they are used in translation. Splicing removes non-protein coding segments, introns, from pre-mRNAs. Splicing also adds an additional layer of complexity to the regulation of gene expression. Splicing enables the generation of multiple different mRNAs from a single gene and is largely regulated by recruitment of RNA binding proteins that regulate splice site selection co-transcriptionally (Pandya-Jones, 2011). Exon-intron boundaries are often defined by a consensus 5' AG, 3' GU sequence spanning the boundary (Burset et al., 2000). Splicing is facilitated by the spliceosome complex, which removes the intronic sequence and ligates the two exons together by means of a transesterification reaction (Brody and Abelson, 1985, Grabowski et al., 1985, Moore and Sharp, 1993). After splicing, the spliceosome

recruits the exon junction complex (Le Hir et al., 2000), which in turn regulates mRNA export. This is a mechanism that prevents the transport of immature mRNAs to the cytoplasm.

The export of mRNA to the nucleus is simultaneously regulated by multiple protein complexes. Both the exon junction complex (Le Hir et al., 2001) and the cap binding complex (Cheng et al., 2006) regulate the recruitment of the transcription export (TREX) complex. The TREX complex interacts with a heterodimeric NXF1 NXT1 complex, which results in conformational switching of NXF1, from a closed to an open conformation (Viphakone et al., 2012). Conformational switching facilitates mRNA binding and export to the cytoplasm via a nuclear pore.

1.1.3 Translation initiation

After nuclear export, the mRNA is now ready for the next step of the gene expression cycle, translation initiation. Translation initiation is controlled by eukaryotic translation initiation factors (eIFs). Translation initiation is a multi-step process which results in the formation and maturation of multiple complexes (figure 1.2). The first of these complexes, the ternary complex is composed of guanosine triphosphate (GTP) and the initiator methionine tRNA bound to eIF2 (Kapp and Lorsch, 2004). The ternary complex then matures to the 43S preinitiation complex. The 43S preinitiation complex is composed of the ternary complex, eIF3, eIF1, eIF1A, eIF5 and the 40S small ribosomal subunit (Valášek et al., 2004). Each member of the complex plays a distinct role in regulating translation. eIF1 and eIF1A stabilise the interaction with the 40S ribosomal subunit. eIF1 and eIF1A binding to the 40S also induces a conformational switch in the 40S. The 40S is in a more open conformation after eIF1 and eIF1A bind, enabling the scanning function of the 40S (Zeman et al., 2019). eIF3 interacts with eIF1 and wraps around the 40S ribosomal subunit, which is a critical step in the maturation of the 43S to the 48S preinitiation complex (Aylett et al., 2015).

The 48S preinitiation complex is composed of the 43S preinitiation complex, an mRNA and the eIF4F complex. The eIF4F complex contains eIF4E, the mRNA cap binding protein,

eIF4G, a scaffold protein, and the DEAD box helicase eIF4A1 (Gingras et al., 1999). eIF4E binding the mRNA cap facilitates recruitment of eIF4G (Morley et al., 1997). eIF4G recruitment increases the affinity of eIF4E for the mRNA cap (Haghighat and Sonenberg, 1997). eIF3 interacts with eIF4G, which enables the formation of the 48S preinitiation complex (Korneeva et al., 2000). After successful formation of the 48S preinitiation complex, the next step in translation initiation is scanning for the AUG start codon. eIF4A1 plays multiple roles in the regulation of scanning. Displacement of the eIF3J subunit of eIF3 from the mRNA entry channel of the 40S is stimulated by eIF4A1 when bound by its co-factor eIF4B in a helicase independent manner (Sokabe and Fraser, 2017). Displacement of eIF3J is required for the scanning function of the 40S. The helicase activity is also involved in unwinding secondary structure in the 5'untranslated region (UTR) (Svitkin et al., 2001). eIF4A1 helicase activity is stimulated by eIF4G binding (Lorsch and Herschlag, 1998). Secondary structures are inhibitory to the scanning function of the 40S (Gray and Hentze, 1994). The 40S scans the 5'UTR of the mRNA until it reaches the AUG start codon, which is usually contained within a consensus sequence known as a Kozak sequence (Kozak, 1987).

Upon start codon recognition the 40S ceases its scanning activity. The start codon is recognised by base pairing of the initiator methionine tRNA with the AUG start codon (Cigan et al., 1988). In response to start codon recognition, eIF1A and eIF5 interact resulting in a conformational change from open to closed in the 40S which stabilises the complex at the start codon (Maag et al., 2005). eIF1 is subsequently released and eIF5 stimulates the hydrolysis of the eIF2 bound GTP (Paulin et al., 2001). The recruitment of the 60S large ribosomal subunit is regulated by eIF5B (Pestova et al., 2000). 40S and 60S ribosomal subunits combine to form an 80S ribosome. Formation of the 80S ribosome results in the dissociation of eIF3, GDP bound eIF2 and eIF5 (Unbehaun et al., 2004). eIF1A and eIF5B then dissociate from the 80S ribosome (Fringer et al., 2007), which is now ready to start the next step in the gene expression cascade, translation elongation.

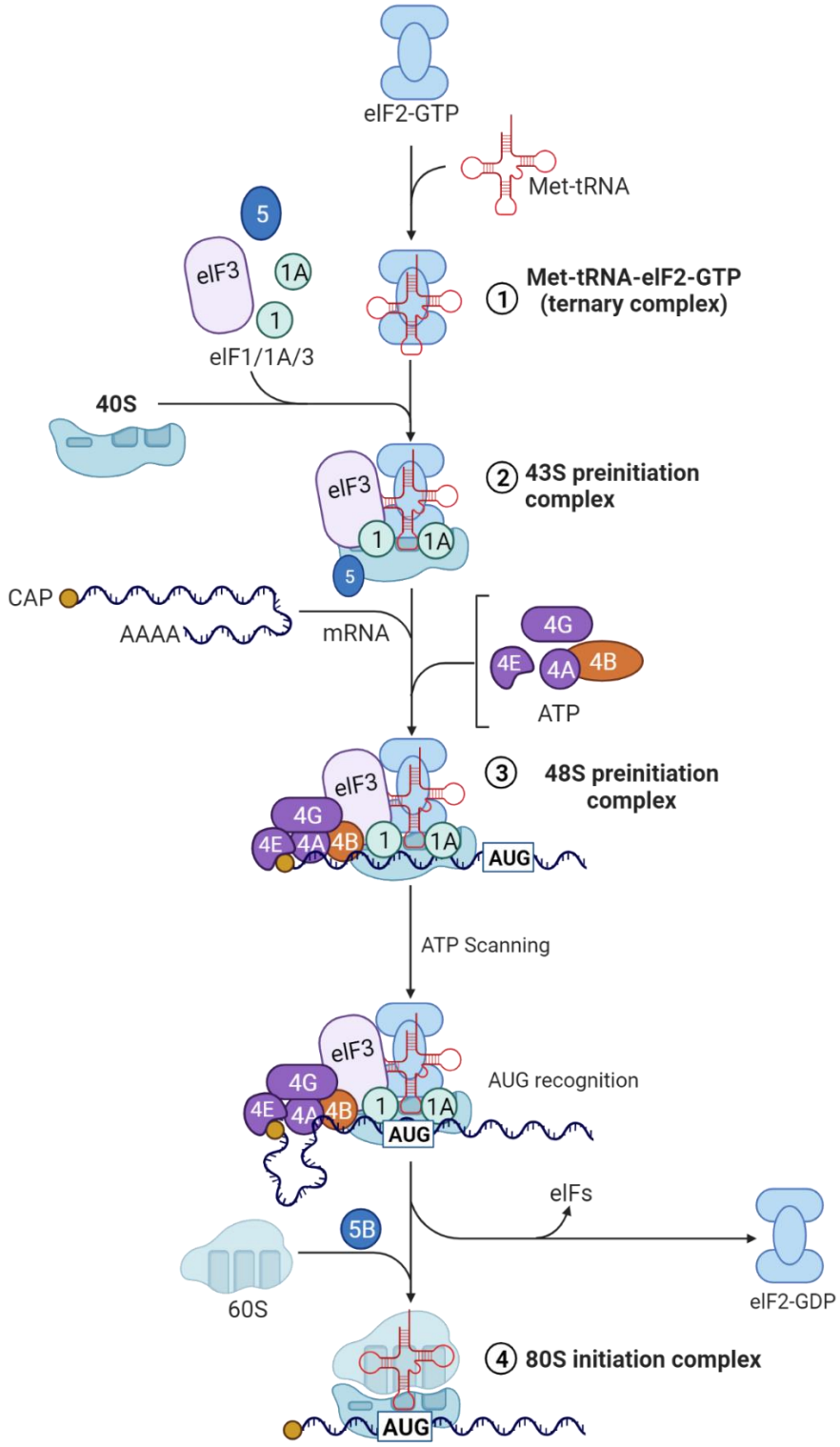


Figure 1.2 Translation initiation:
Translation initiation is regulated by the sequential formation of multiple preinitiation complexes. 1 the ternary complex contains eIF2 bound to GTP and the initiator methionine tRNA. 2 the 43S preinitiation complex contains the ternary complex plus eIF- 3, 5, 1 and 1a. 3 the 48S preinitiation complex contains the 43S preinitiation complex plus the eIF4F cap binding complex and an mRNA. eIF4E binds the cap and recruits eIF4G. eIF4G interacts with eIF3 and recruits the 43S complex to the mRNA. eIF4A1 regulates translation initiation by displacing eIF3 from the mRNA entry channel and unwinding secondary structures in the 5'UTR. Both these activities are stimulated by recruitment of the co-factor eIF4B. The scanning 40S ribosomal subunit stops when it encounters a start codon. eIF2 then hydrolyses GTP 4 eIF5B regulates the recruitment of the 60S subunit. Upon 60S recruitment, the eIF proteins dissociate and the 80S ribosome can begin elongation.

1.1.4 Translation initiation in stress

Translation is the most energetically exhaustive cellular process (Buttgereit and Brand, 1995). In response to various stresses, cells often downregulate translation initiation as a means of energy conservation.

1.1.4.1 The integrated stress response

The eIF2 subunit eIF2α can be phosphorylated at Ser51 by a number of kinases in response specific stresses, termed the integrated stress response (figure 1.3). eIF2α is phosphorylated by heme-regulated inhibitor (HRI) in response to heat shock (Lu et al., 2001) or oxidative stress (McEwen et al., 2005). PKR like endoplasmic reticulum kinase (PERK) phosphorylates eIF2α in response to endoplasmic reticulum stress (Srivastava et al., 1998). Protein kinase R (PKR) phosphorylates eIF2α in response to viral infection (McInerney et al., 2005). General control nonderepressible 2 (GCN2) phosphorylates eIF2α in response to amino acid starvation (Dever et al., 1992). Phosphorylation of the 2α subunit of eIF2 inhibits the guanine nucleotide exchange factor (GEF) function of eIF2B, thus stalling translation initiation (Pavitt et al., 1997). The integrated stress response simultaneously down regulates bulk translation and allows for the selective translation of certain mRNAs, such as activating transcription factor 4 (ATF4) (Lu et al., 2004). Activation of the integrated stress response pathway upregulates the expression of multiple classes of proteins that aid stress circumvention, such as protein folding chaperones, amino acid synthesis enzymes and aminoacyl-tRNA synthetases (Pakos-Zebrucka et al., 2016).

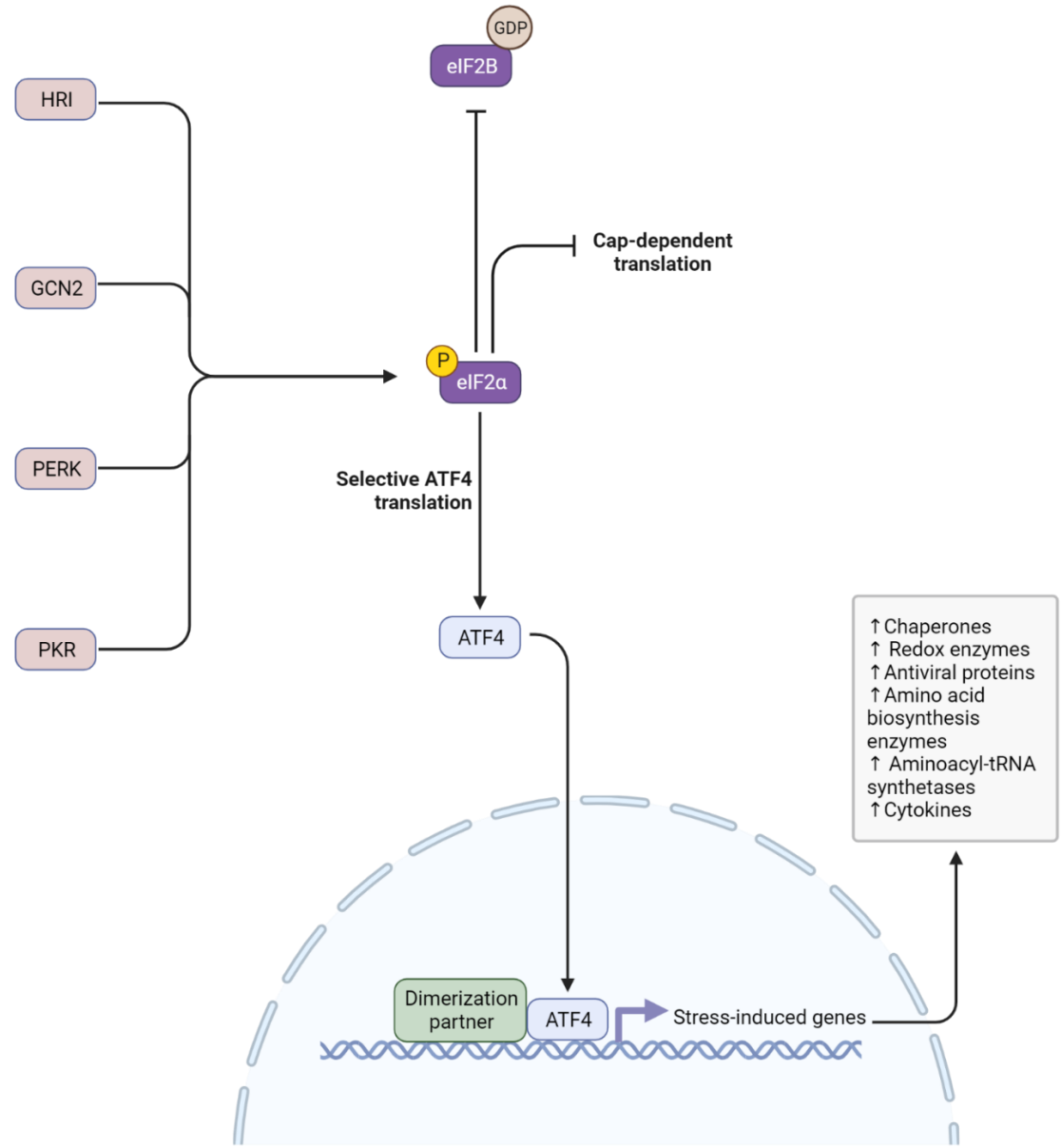


Figure 1.3 The integrated stress response:
In response to various stresses eIF2α is phosphorylated at Ser51 by one of four kinases. HRI, GCN2, PERK or PKR. eIF2α phosphorylation inhibits cap dependent translation by inhibiting the GEF function of eIF2B. The integrated stress response promotes selective translation of stress responsive proteins such as ATF4. ATF4 is a transcription factor that in turn up-regulates the expression of stress responsive proteins.

1.1.4.2 mTOR signalling

Another stress responsive pathway that cells use to limit translation initiation rates is mammalian target of rapamycin (mTOR) signalling (figure 1.4). mTOR is the catalytic subunit of two protein complexes, mTORC1 and mTORC2. The activity of mTORC1 directly regulates translation. mTORC1 is composed of the mTOR kinase, the regulatory-associated protein of mTOR (RAPTOR) regulatory subunit, mTOR Associated Protein, LST8 Homolog (MLST8), proline-rich Akt/PKB substrate 40 kDa (PRAS40) and DEP Domain Containing mTOR Interacting Protein (DEPTOR) (Kim et al., 2002, Kim et al., 2003). mTORC1 activity is regulated by GTPases, although the exact mechanism of this interaction is still unknown (Durán and Hall, 2012). mTORC1 can be inhibited by direct phosphorylation of RAPTOR by AMP-activated protein kinase (AMPK) (Gwinn et al., 2008), or activation of the tuberous sclerosis complex (TSC) regulatory complex (Mihaylova and Shaw, 2011). mTORC1 activity is downregulated in response to a variety of stimuli, including glucose deprivation, hypoxia, and amino acid starvation. The TSC complex stimulates the GTPase activity of RHEB GTPases which render mTOR inactive (Durán and Hall, 2012). mTORC1 phosphorylates eIF4E binding protein 1 (4E-BP1) which results in its dissociation from eIF4E. 4E-BP1 binding eIF4E is inhibitory to eIF4G recruitment, which prevents eIF4F complex assembly and translation initiation (Marcotrigiano et al., 1999). mTORC1 also activates the Ribosomal Protein S6 (RPS6) kinases which in turn phosphorylate RPS6 (Saitoh et al., 2002). The consequences of RPS6 phosphorylation are to date not fully characterised although it has been implicated in promoting the translation of specific subsets of mRNAs (Bohlen et al., 2021).

1.1.4.3 Non canonical translation initiation

It has been known for a number of years that in response to stress cells can use alternative mechanisms of translation initiation. Cap independent mechanisms include internal ribosome entry sites (IRES) (Bushell et al., 2006). IRESs are RNA elements that facilitate the recruitment of ribosomes independently of the eIF4F complex. To date a limited number of endogenous mRNAs have been identified with IRESs although specific mRNAs regulating apoptosis and mitosis have been identified (Bushell et al., 2006, Spriggs et al., 2008). Alternative cap binding proteins and complexes have also been shown to be critical mediators of translation initiation in response to stress. In hypoxia, an alternative eIF4F complex forms. This complex contains the alternative cap binding protein eIF4E2, eIF4G3

and eIF4A1 (Uniacke et al., 2012, Melanson et al., 2017). The Hypoxic eIF4F complex has been shown to promote the translation of certain mRNAs such as epidermal growth factor receptor (EGFR) (Melanson et al., 2017, Uniacke et al., 2012). The eIF3D subunit of eIF3 can also bind the mRNA cap (Lee et al., 2016). In response to glucose deprivation casein kinase 2 (CK2) kinase activity is reduced. CK2 under nutrient replete conditions phosphorylates eIF3D. When this phosphorylation is lost it enhances eIF3D's cap binding capability and drives the translation of pro-survival mRNAs such as JUN (Lamper et al., 2020).

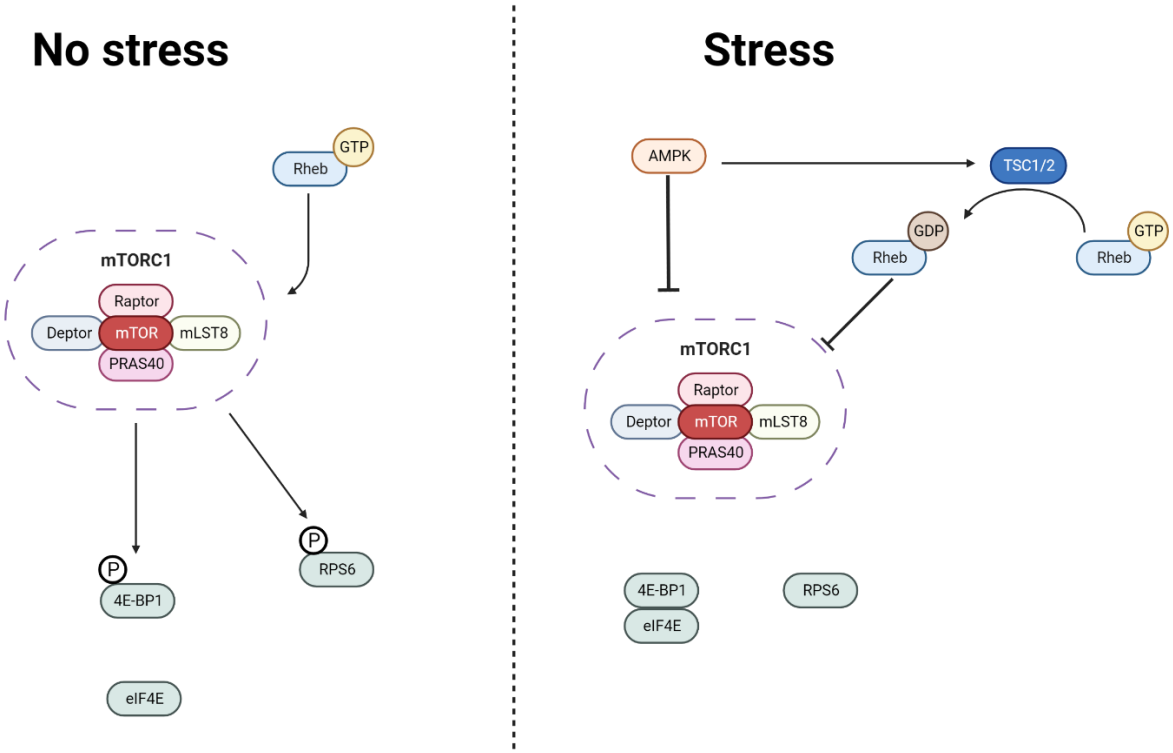


Figure 1.4 mTOR signalling regulates translation:
In normal growth conditions, mTORC1 promotes cap dependent translation by phosphorylating 4E-BP1, which causes it to dissociate from eIF4E, the cap binding member of the eIF4F complex. mTORC1 also activates the RPS6 kinases, which in turn phosphorylates RPS6. GTP bound Rheb promotes mTORC1 activity. In response to various stresses, mTORC1 is inactivated. This can be achieved by direct phosphorylation of RAPTOR by AMPK, or activation of the TSC1/2 complex which promotes GTP hydrolysis of the GTP bound by Rheb. GDP bound Rheb is inhibitory to mTORC1 activity. Hypophosphorylation of 4E-BP1 prevents eIF4F complex assembly and inhibits translation.

1.1.5 Translation elongation and termination

The ribosome is the molecular machine responsible for protein synthesis. Ribosomes are composed of two subunits. The 40S small ribosomal subunit and the 80S large ribosomal subunit. The 40S subunit is composed of 33 ribosomal proteins and the 18S rRNA (Anger et al., 2013, Rabl et al., 2011). The 60S subunit is composed of 47 ribosomal proteins and the 28-, 5.8- and 5S rRNAs (Anger et al., 2013, Ban et al., 2000). 80S ribosomes contain three tRNA binding sites. Namely, the aminoacyl (A site), peptidyl (P-site) and exit (E-site) sites (Wettstein and Noll, 1965). rRNA is critical for the formation of the peptidyl bond between two amino acids as no proteins are proximal to this reaction within the ribosome (Nissen et al., 2000).

After the initiation phase of translation has finished, an 80S ribosome will be bound to an mRNA with the initiator methionine tRNA occupying the P-site. Translation elongation is regulated by a family of proteins known as eukaryotic elongation factors (eEFs). The first step of translation elongation involves GTP bound eEF1A recruiting an aminoacylated tRNA to the A-site of the ribosome (Andersen et al., 2000, Ozturk and Kinzy, 2008, Sasikumar et al., 2012). Base pairing between the aminoacylated tRNA anti-codon and the codon of the mRNA stimulates eEF1A GTP hydrolysis and release. Release of eEF1 results in a conformational switch of the ribosome. This conformational switch brings the amino and carbonyl groups of the amino acids bound to the tRNAs occupying the A and P sites in close proximity facilitating the formation of a peptide bond (Rodnina et al., 2007). This ribosomal conformation is stabilised by the association of GTP bound eEF2. eEF2 GTP hydrolysis stimulates translocation of the ribosome to the next codon of the mRNA (Rodnina et al., 1997). This process is then repeated until the ribosome encounters a stop codon.

Termination of translation is regulated by a family of proteins known as the eukaryotic translation termination factors (eRFs). Stop codons are not decoded by tRNAs. eRF1 recognises the stop codon and occupies the A site of the ribosome (Song et al., 2000). eRF1 is bound by a GTP bound co-factor eRF3 which hydrolyses the GTP (Kobayashi et al., 2004, Salas-Marco and Bedwell, 2004). GTP hydrolysis enables eRF1 to promote release of the polypeptide chain of the ribosome. This is achieved by hydrolysis of the peptidyl tRNA by

the ribosome (Alkalaeva et al., 2006). After the polypeptide has been released by the ribosome the large and small subunits dissociate in an ATP-binding cassette sub-family E member 1 (ABCE1) dependent manner (Pisarev et al., 2010).

Gene expression is a highly regulated process that involves compartmentalisation of the two main processes. Transcription occurs in the nucleus whereas translation occurs in the cytosol. Recent work has revealed a new paradigm in the regulation of gene expression. Gene expression can be further regulated by compartmentalisation through formation of biomolecular condensates. Biomolecular condensates will be the focus of the next section of this introduction.

1.2 Biomolecular condensates

One of the greatest challenges any living cell must overcome to thrive is the simultaneous coordination of thousands of biochemical reactions (Shin and Brangwynne, 2017). The advent of light microscopy in the 1800's revealed that cells contain highly ordered structures termed organelles. The first organelle to be described was the nucleus (Brown and Bennett, 2012). Dozens of subcellular compartments have since been described, often spatially defined by an outer phospholipid membrane. The term organelle is so synonymous with membrane-defined structures that the concept of membraneless structures is often overlooked (Shin and Brangwynne, 2017). Despite this, some of the first subcellular compartments described, including the nucleolus (Pederson, 2011) are indeed lacking a phospholipid membrane. Subsequent research has revealed that cells contain a vast array of intracellular membraneless condensates or organelles, often composed of rich networks of RNA and protein, hence the term ribonucleoprotein (RNP) (Mao et al., 2011b, Anderson and Kedersha, 2009, Handwerger and Gall, 2006, Shin and Brangwynne, 2017). Both nuclear (e.g., nucleolus, paraspeckle, Cajal body) and cytoplasmic (e.g., P-granules, stress granules, processing-bodies (P-Bodies)) RNP condensates have been described with shared biophysical properties (figure 1.5).

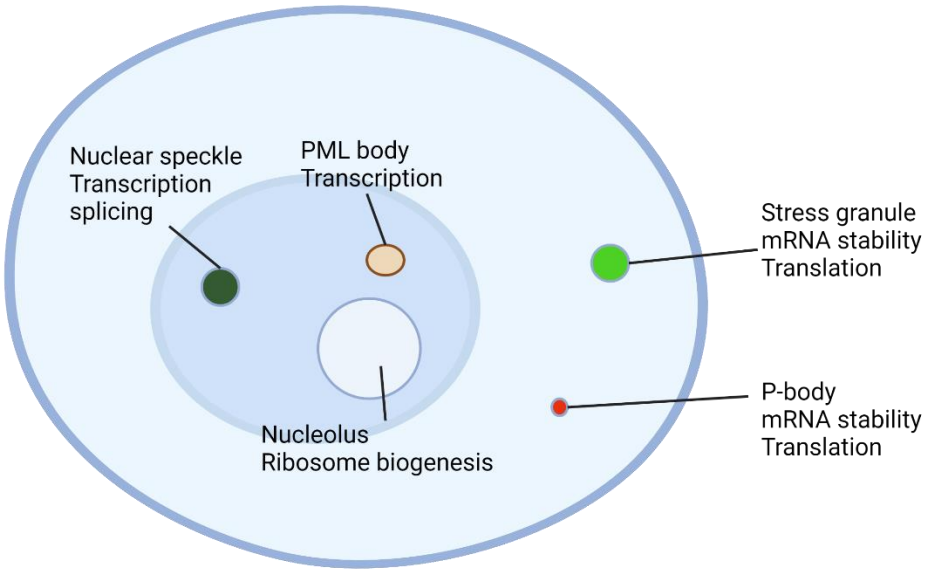


Figure 1.5 Examples of nuclear and cytoplasmic biomolecular condensates:
Liquid-liquid phase separated membraneless organelles, referred to as biomolecular condensates represent an exciting new field of biology and cellular organisation. Biomolecular condensates have been identified in the nucleus (e.g., nuclear speckles, PML bodies and the nucleolus) and the cytoplasm (e.g. stress granules and p-bodies). Biomolecular condensates are of particular research interest as they have been implicated in the regulation of gene expression.

1.2.1 Biomolecular condensates exhibit liquid dynamics.

Seminal work characterising germline P-granules in *Caenorhabditis elegans* revealed the liquid nature of biomolecular condensates (Brangwynne et al., 2009). A key observation in this study showed that when two P-granules came into contact with one another they coalesced into a single spherical assembly. In addition to this, P-granules showed the dynamic properties of a liquid. Namely in response to shear stress, P-granules were shown to share the movement properties of a simple liquid like water as they flowed off nuclei and dripped (Brangwynne et al., 2009). Fluorescent labelling of P-granule proteins followed by fluorescence recovery after photobleaching showed rapid molecular reorganisation as would be expected in a liquid due to high diffusion rates (Brangwynne et al., 2009). Subsequent further investigation showed that liquid like behaviour was conserved for the nucleolus (Brangwynne et al., 2011) and other RNP biomolecular condensates such as stress granules (Molliex et al., 2015).

1.2.2 Liquid-liquid phase separation drives biomolecular condensate formation.

Since the discovery of the liquid like properties of RNP biomolecular condensates, it is now understood that a physiochemical phenomenon known as liquid-liquid phase separation is the driving force behind their assembly (Shin and Brangwynne, 2017, Banani et al., 2017, Hirose et al., 2023). The concept of liquid-liquid phase separation is often explained using the example of oil and water. If you take a mixture of oil and water and shake it, droplets of oil will be seen in the water. However, these two liquids are immiscible and will always demix and form two distinct liquid phases. RNP biomolecular condensate formation is regulated by subsets of cellular proteins that facilitate the condensation of biomolecules, such as RNA and protein (Shin and Brangwynne, 2017, Banani et al., 2017). Characteristic molecular features of these proteins will be outlined below.

1.2.2.1 The contribution of proteins to liquid-liquid phase separation

Liquid-liquid phase separated biomolecular condensates are known to be enriched with proteins that exhibit multivalency. Multivalent proteins harbour multiple protein-protein or protein-nucleic acid interaction domains (Pawson and Nash, 2003, Cohen et al., 1995).

One of the best characterised domain families that contribute to liquid-liquid phase separation are intrinsically disordered regions (IDRs, sometimes referred to as low complexity domains). IDRs are sequences of polypeptides described as lacking hydrophobic amino acids. Hydrophobic amino acids facilitate protein folding and formation of ordered tertiary structures (Chothia, 1975, Chothia, 1974). IDRs instead contain a higher proportion of charged and polar amino acids and as a result lack a defined tertiary structure (Uversky et al., 2000, Babu, 2016). Lacking a defined tertiary structure enables IDRs to be present in multiple conformational states that can be rapidly remodelled in response to physiological stimuli (Babu, 2016), such as the IDRs of the stress granule nucleating protein G3BP1 (Guillén-Boixet et al., 2020, Yang et al., 2020). IDRs can facilitate protein-protein interactions through multiple mechanisms, including hydrogen bonds and charge-charge interactions (Hirose et al., 2023). Purification of reconstituted IDRs of multiple proteins have been shown to be sufficient to promote phase separation and droplet formation *in vitro*. These include IDRs originating from the DEAD box helicase (DDX) DDX4 (Nott et al., 2015) and LAF-1 (Elbaum-Garfinkle et al., 2015). Proteins containing IDRs have been shown to be pivotal regulators of the formation of multiple biomolecular condensates, including stress granules (Gilks et al., 2004), nucleoli (Brangwynne et al., 2011), paraspeckles (Naganuma et al., 2012) and P-bodies (Ayache et al., 2015, Hondele et al., 2019). Proteins that contain IDRs are frequently multivalent. IDRs are often located proximal to other protein domains that facilitate nucleic acid binding (Hirose et al., 2023). The stress granule nucleating protein Ras GTPase-activating protein-binding protein 1 (G3BP1) contains 3 IDRs, two of which flank G3BP1's RNA recognition motif (RRM). In response to G3BP1 binding RNA, the IDRs change conformation and facilitate homodimerisation and promote liquid-liquid phase separation (Yang et al., 2020, Guillén-Boixet et al., 2020).

1.2.2.2 The contribution of RNA to liquid-liquid phase separation

Multivalent RNA mediated interactions have also been implicated in regulating the formation of biomolecular condensates. Long non-coding RNAs can serve as an architectural scaffold for recruitment of RNA binding proteins and promotion of phase separation through multivalent interactions (Hirose et al., 2023). A model has been proposed for the formation of nuclear paraspeckles which is dependent on multivalent RNA-RNA and RNA-protein interactions of the NEAT1 long non-coding RNA (Yamazaki et al., 2019). In this model, the nascent NEAT1 transcript is bound by RNA binding proteins (e.g., SPFQ and NONO) (Mao et al., 2011a, Lee et al., 2015). Simultaneously, nascent NEAT1 transcripts also interact with each other through RNA-RNA mediated interactions (Yamazaki et al., 2018). This facilitates oligomerisation of the NEAT1 RNP complexes which then facilitate the recruitment of additional RNA binding proteins (e.g., FUS, RBM14) that promote liquid-liquid phase separation and paraspeckle formation (Yamazaki et al., 2019). Long non-coding RNAs have also been identified in other biomolecular condensates, such as stress granules (Khong et al., 2017), however their role in regulating organelle dynamics is not fully understood.

Another finding that supports a role of RNA in regulating liquid-liquid phase separation is that when at sufficient concentration, RNA will spontaneously phase separate and form droplets *in vitro* (Tauber et al., 2020). This study also proposed that stress granule P-body docking interactions (please note, stress granule P-body docking interactions will be explained in greater detail in the following section) are facilitated by RNA-RNA mediated interactions (Tauber et al., 2020). RNA-RNA mediated interactions are believed to contribute to liquid-liquid phase separation through a number of mechanisms, including Watson-Crick base pairing, non-Watson-Crick base pairing and hydrogen bonding (Hirose et al., 2023).

RNA post-transcriptional modifications have also been implicated in promoting liquid-liquid phase separation. RNA methylation has been shown to promote liquid-liquid phase separation *in vitro*. Two independent studies showed that m^6A RNA methylation promoted the formation of liquid droplets through enhanced recruitment of YTH N6-methyladenosine RNA binding proteins (YTHDF) RNA binding proteins (Anders et al., 2018, Ries et al.,

2019). It is important to note that a follow up study showed that m^6A methylation had a limited, if any effect on RNA partitioning into stress granules and may only account for 6% of the variation observed between wild type and Methyltransferase 3, N6-Adenosine-Methyltransferase Complex Catalytic Subunit (METTL3) knockout cells (Khong et al., 2022). Nonetheless, with technological advancements that utilise high throughput methodologies, a greater understanding of how RNA modifications contribute to liquid-liquid phase separation in a cellular context is now perhaps attainable (Schwartz and Motorin, 2017).

1.2.3 Biomolecular condensates regulate gene expression at multiple steps.

Multiple biomolecular condensates have been shown to regulate gene expression at various steps of the process (figure 1.5). This mediated through the regulation of ribosome biogenesis, and both transcriptional and translational regulation (Hirose et al., 2023). The role that specific biomolecular condensates play in these processes will be outlined in greater detail below.

1.2.3.1 Biomolecular condensates regulate ribosome biogenesis.

The best described function of the nucleolus is ribosome biogenesis through coordination of the initial steps of the process, such as rRNA transcription by RNA polymerase I (Pederson, 2011, Lafontaine et al., 2021). Perturbations in nucleolar morphology through altered phase separation dynamics contribute to multiple diseases including cancer (Derenzini et al., 2009) and ribosomopathies (Farley-Barnes et al., 2019). This change in nucleolar morphology is caused by excess or insufficient functional ribosomes for cancers and ribosomopathies respectively. A greater understanding of nucleolar phase separation dynamics is required to exploit these observations for the development of therapeutics (Lafontaine et al., 2021).

1.2.3.2 Biomolecular condensates regulate transcription.

Biomolecular condensates have also been implicated in the regulation of transcription. Nuclear speckles form in the interchromatin space (Thiry, 1993). Although their exact functions remain to be fully characterised two functions that impact gene expression have been proposed. Firstly, they are proposed to be reservoirs of mRNA processing factors (Saitoh et al., 2004). The second function proposed implicates nuclear speckles in gene expression through promoting transcription and post-transcriptional processing of mRNAs at proximal genomic loci (Chen and Belmont, 2019, Fei et al., 2017, Smith et al., 1999). Promyelocytic leukemia protein (PML) bodies are another species of nuclear biomolecular condensate implicated in transcriptional regulation. PML bodies associate with genomic loci that are transcriptionally active and exclude promoter methyltransferases to maintain promoter activity (Kurihara et al., 2020).

1.2.3.3 Biomolecular condensates regulate translation.

Cytoplasmic biomolecular condensates regulate gene expression through modulation of RNA stability. P-bodies are constitutively present in the cytoplasm but increase in number in response to stresses such as sodium arsenite treatment (Ayache et al., 2015). Proteins involved in mRNA decay pathways, in particular the mRNA decapping complex proteins Decapping mRNA 1A (DCP1A), enhancer of mRNA decapping 4 (EDC4) and Decapping mRNA 2 (DCP2), co-localise to P-bodies and facilitate mRNA decapping followed by 5'-3' exoribonuclease II mediated decay within the P-body liquid-liquid phase separated compartment (Sheth and Parker, 2003). Further to regulating mRNA decay, P-bodies act as a site of storage of translationally repressed mRNAs (Hubstenberger et al., 2017). The majority of P-body mRNAs are in fact not degraded. Therefore, the dominant mechanism in the regulation of translation by P-bodies is believed to be the sequestration of mRNAs away from the translation machinery (Hubstenberger et al., 2017).

Elucidating a better understanding of how stress granules dynamically regulate translation is the primary focus of the experimental work presented in this thesis. A detailed introduction to stress granules is presented below.

1.3 Stress granules

Stress granules are an evolutionarily conserved cytoplasmic biomolecular condensate composed of protein and RNA that form in response to a variety of cellular stresses including viral infection (McInerney et al., 2005), oxidative stress, heat and nutrient deprivation (Anderson and Kedersha, 2009, Anderson and Kedersha, 2002, Anderson and Kedersha, 2006). Stress granules are said to be cytoprotective, as inhibition of stress granule formation results in reduced cell viability in response to stress.

Stress granules have a biphasic substructure, with a more dense core structure, which can be biochemically purified (Jain et al., 2016), and a more fluid like shell (figure 1.6). This has been observed by multiple imaging techniques. Electron micrographs first showed this substructure, by identifying areas that had a higher electron density within stress granules (Souquere et al., 2009). Super-resolution fluorescence microscopy has shown areas within stress granules that had higher densities of protein and mRNA (Jain et al., 2016). The movement of stress granule components between the cytoplasm, liquid shell and the more solid core region, often referred to as re-modelling, is an ATPase dependent highly dynamic process (Jain et al., 2016). In response to ATP ablation stress granules become less dynamic, indicating that remodelling is an energetically regulated active process. A key observation that supports the more liquid like nature of the shell is that the stress granule nucleating protein G3BP1 can still exchange in and out of the shell to the cytoplasm, whereas the core structure appears more stable and less dynamic. Biphasic properties are conserved in a number of other biomolecular condensates, including the nucleolus which contains a more dense fibrillarin core (Brangwynne et al., 2011). To date, it is unknown what the function of the core and shell stress granule sub-compartments are.

Stress granule formation is regulated by upstream signalling pathways that result in the formation of non-canonical translation pre-initiation complexes; thus they are said to be formed from RNP complexes stuck at the initiation step of translation (Panas et al., 2016, Protter and Parker, 2016), (figure 1.6). Their formation has been shown to be linked with the proportion of mRNAs engaged in active translation and associated with polysomes. A

reduction in polysome associated mRNAs promotes stress granule formation (Kedersha et al., 2000).

1.3.1 Translation initiation regulates stress granule assembly.

Phosphorylation of eIF2α at serine 51 is perhaps the best characterised upstream signalling event that leads to stress granule formation (Panas et al., 2016). Four stress responsive kinases can phosphorylate this residue in response to various stresses, which leads to the depletion of available ternary complex (see section 1.1.41 for more detail or refer to figure 1.3). eIF2a phosphorylation drives stress granule formation in multiple contexts. In response to sodium arsenite treatment and heat shock, HRI mediated phosphorylation results in stress granule formation (McEwen et al., 2005, Lu et al., 2001). Endoplasmic reticulum stress results in eIF2α phosphorylation by PERK (Srivastava et al., 1998). Nutrient stress results in eIF2α phosphorylation, which is mediated by GCN2 (Dever et al., 1992) and viral infection by PKR (Meurs et al., 1990). The fact that eIF2α mediates translational repression and thus stress granule assembly in response to a vast array of cellular stresses highlights its importance as a conserved stress responsive mechanism. Integrated stress response inhibitor (ISRIB), a compound which potently inhibits the repression of translation mediated by eIF2α phosphorylation (Anand and Walter, 2020), also prevents stress granule formation in response to sodium arsenite and thapsigargin stress (Sidrauski et al., 2015).

Mechanisms independent of eIF2α phosphorylation that facilitate stress granule formation have also been described (figure 1.6). In this instance, inhibition of the eIF4F translation initiation complex results in the repression of translation and formation of stress granules (Mazroui et al., 2006). Inhibition of the eIF4F complex member eIF4A1 by multiple modalities results in stress granule formation. These included inhibition by the compounds hippuristanol and pateamine. This study also showed that depletion of *eIF4A1* by siRNA was sufficient to promote stress granule formation (Mazroui et al., 2006). Previous to this, stress granule formation was believed to be solely regulated by eIF2α phosphorylation. These findings implicated the eIF4F complex as another stress granule regulatory axis. As a result, it has since been theorised that signalling events that perturb eIF4F complex formation may also promote stress granule formation (Panas et al., 2016).

One other mechanism of eIF4F complex inhibition that has been implicated in stress granule formation is mTOR dependent phosphorylation of 4E-BP1 (Panas et al., 2016). Hyperphosphorylation of 4E-BP1 results in its dissociation from eIF4E which enables eIF4G, the scaffold protein of the eIF4F complex, to bind eIF4E, promote eIF4F complex assembly and translation initiation (Marcotrigiano et al., 1999). Further evidence that supports a role of mTOR regulation of stress granule formation is that multiple agents that result in reduced mTOR activity promote stress granule assembly, including treatment with hydrogen peroxide (Emara et al., 2012) and sodium selenite (Fujimura et al., 2012). Genetic perturbation of eIF4E 4E-BP1 complexes resulted in reduced numbers of stress granules per cell in both studies. More recent work has also shown that mTOR kinase activity is required for stress granule formation and maintenance in response to sodium arsenite treatment (Sfakianos et al., 2018). mTOR inhibition through treatment with rapamycin or depletion of the mTOR regulatory subunit *RAPTOR* results in reduced stress granule formation (Sfakianos et al., 2018). In addition to this, concomitant phosphorylation of eIF2α was observed in response to sodium selenite treatment (Fujimura et al., 2012) but not with hydrogen peroxide (Emara et al., 2012). These findings together show that mTOR signalling regulates stress granule assembly. However, the context and stress applied is important. It is also unknown whether mTOR regulates stress granule assembly by different means depending on the phosphorylation status of eIF2α. Further research is needed to fully understand the contributions of mTOR signalling to stress granule dynamics.

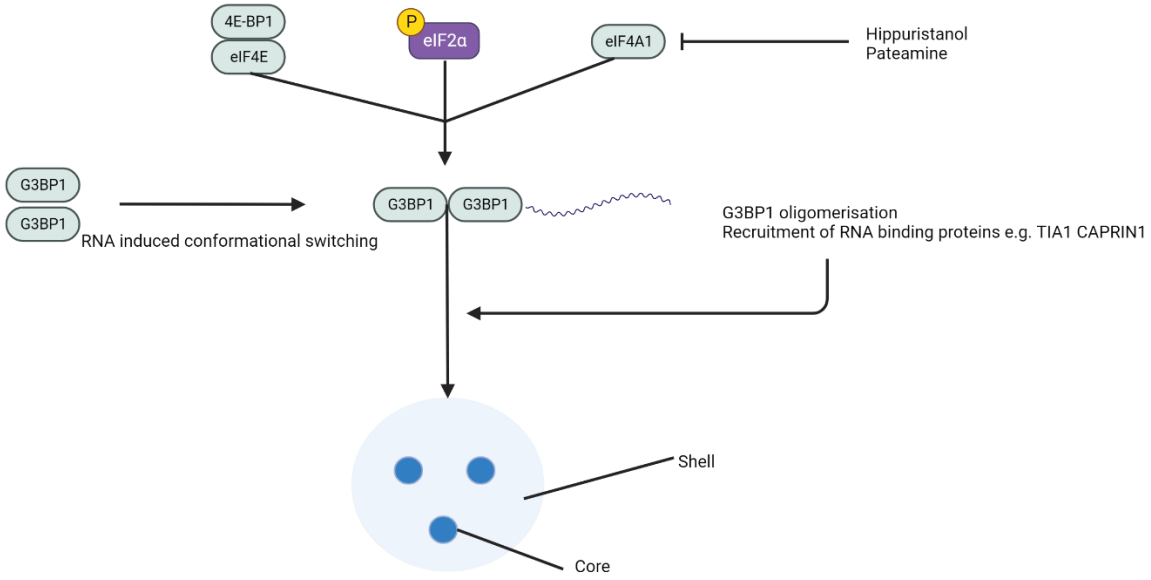


Figure 1.6 Stress granule assembly:
Stress granule assembly is regulated by translation initiation. Phosphorylation of eIF2α, hypophosphorylation of 4E-BP1 or inhibition of eIF4A1 have been shown to promote stress granule assembly. G3BP1, the stress granule nucleating protein binds mRNAs which induces a conformational change. This facilitates G3BP1 oligomerisation and recruitment of other RNA binding proteins that promote liquid-liquid phase separation and stress granule assembly. Like many other biomolecular condensates, stress granules contain a biphasic substructure. With a more fluid shell and more dense core regions.

1.3.2 Stress granule assembly

Once upstream signalling events have resulted in the formation of RNP complexes stalled at translation initiation, the next step in stress granule assembly is condensation of these complexes by liquid-liquid phase separation. As mentioned in section 1.3.2.1, proteins with IDRs can facilitate the formation phase separated liquid droplets. The contributions of stress granule proteins with IDRs to stress granule assembly will be outlined below.

1.3.2.1 The contribution of IDRs to stress granule assembly

G3BP1 is a stress granule protein that has long been implicated in stress granule formation. The first direct evidence for its role in stress granule formation came when it was observed that overexpression of G3BP1 causes spontaneous stress granule formation in unstressed cells (Tourrière et al., 2003). G3BP1 contains three IDRs. An acidic IDR (IDR 1), which is located proximal to its N-terminal nuclear transport factor 2 like domain (NTF2L). G3BP1 contains two additional basic IDRs, which flank its RRM domain (Tourrière et al., 2003, Yang et al., 2020, Guillén-Boixet et al., 2020). Due to the multiple domains that can support protein-protein or protein-RNA interactions it exhibits multivalency, which is a known characteristic of liquid-liquid phase separation promoting proteins (Shin and Brangwynne, 2017) (figure 1.7).

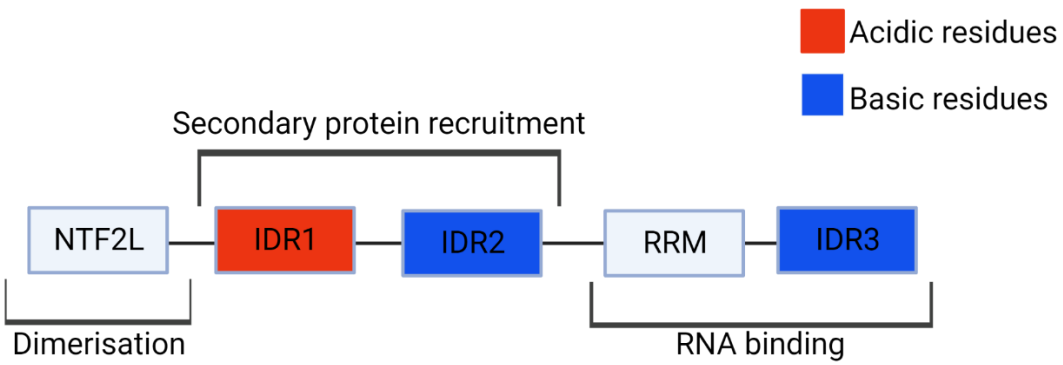


Figure 1.7 G3BP1 protein domains:
G3BP1 contains multiple domains that facilitate protein-protein or protein-RNA interactions, thus it said to be multivalent. Multivalency is a pivotal promoter of liquid-liquid phase separation.

Subsequent years of research have since enabled the development of a model as to the mechanism of G3BP1 mediated stress granule formation. A crucial study in the development of this model was published in 2016 which provided multiple mechanistic insights (Kedersha et al., 2016). Firstly, *G3BP1 G3BP2* double knockout rendered cells unable to form stress granules in response to eIF2α phosphorylation or eIF4A1 inhibition. G3BP2 is a paralogue of G3BP1, both proteins can form homo- and heterodimers (Matsuki et al., 2013). They have been shown to be functionally redundant (Yang et al., 2020). G3BP2 protein expression increases in response to loss of G3BP1 and can facilitate stress granule formation, therefore both proteins were targeted simultaneously to counteract this (Kedersha et al., 2016). G3BP1 was shown to interact with the 40S ribosomal subunit through its third IDR, which provided a link between G3BP1 and the translation machinery. Expression of full length G3BP1 rescued stress granule formation in the *G3BP1 G3BP2* knockout cells whereas expression of a G3BP1 mutant lacking its third IDR did not (Kedersha et al., 2016). The importance of secondary proteins binding G3BP1 and how that regulates stress granule assembly was also shown in this study (Kedersha et al., 2016). Two proteins bound G3BP1 in a mutually exclusive manner, with binding of Caprin1 promoting stress granule assembly and USP10 binding was inhibitory to stress granule formation. Both of these proteins interact with G3BP1 at its N-terminal acidic IDR. This IDR is proximal to G3BP1's NTF2L domain, which is required for G3BP1 oligomerisation (Tourrière et al., 2003, Matsuki et al., 2013). At the time of publication of Kedersha et al, the exact mechanism as to how this regulates stress granule assembly was yet to be elucidated (Kedersha et al., 2016).

Three groundbreaking studies were published in 2020 that to date have provided the most complete model of how G3BP1 promotes stress granule formation (Yang et al., 2020, Guillén-Boixet et al., 2020, Sanders et al., 2020). These studies demonstrated that it is not merely the presence of G3BP1's IDRs that promote stress granule assembly through liquid-liquid phase separation but that G3BP1's IDRs change in conformation in response to binding mRNA. In response to binding RNA via its C-terminal RRM and IDR 3 domains, G3BP1 switches from a closed conformation to an open conformation. When in this open conformation, G3BP1 rapidly oligomerises and promotes liquid-liquid phase separation. This serves as a seed for stress granule formation, which is propagated by recruitment of additional proteins that also promote liquid-liquid phase separation such as CAPRIN1 or T-Cell-Restricted Intracellular Antigen-1 (TIA1) (Yang et al., 2020, Sanders et al., 2020, Guillén-Boixet et al., 2020). This is mediated through interactions with G3BP1's IDR 1 and

IDR 2. A model has now been proposed whereby G3BP1 and G3BP2 act as the central node of a dense network of protein and RNA that coalesce around the G3BP1 seed (Yang et al., 2020, Guillén-Boixet et al., 2020, Sanders et al., 2020). This model also supports the findings that some proteins such as TIA1 promote stress granule assembly when overexpressed (Gilks et al., 2004) but are not essential for stress granule formation when knocked out (Yang et al., 2020).

1.3.2.2 The contribution of RNA to stress granule assembly

Multiple lines of evidence support a role of RNA in regulating stress granule assembly. Firstly, RNA is required for G3BP1 to phase separate and form liquid droplets *in vitro* (Yang et al., 2020). Total RNA isolated from yeast forms stress granule like droplets *in vitro* with the addition of polyamines (Van Treeck et al., 2018). Polyamines were added to mimic cellular stress conditions *in vitro* as they are produced in response to a variety of stresses including oxidative stress (Rhee et al., 2007). Activation or addition of RNAse enzymes inhibited stress granule formation in A549 cells (Decker et al., 2022), and prevented RNA dependent G3BP1 liquid-liquid phase separation *in vitro* respectively (Yang et al., 2020). Multiple independent studies have shown that stress granule RNAs are longer than stress granule depleted RNAs (Khong et al., 2017, Namkoong et al., 2018, Padrón et al., 2019). RNA length is believed to contribute to stress granule assembly due to longer molecules being able to act as a platform to provide more available nucleotides to multivalent protein-RNA or RNA-RNA interactions (Khong et al., 2022). *In vitro* experiments support this hypothesis as longer RNAs enhance G3BP1's ability to phase separate (Yang et al., 2020). There is now a large body of evidence supporting the essentiality of RNA in regulating liquid-liquid phase separation, however exact mechanisms are not yet known. Further work is needed to understand how RNA acts as a regulatory molecule in phase separation. It is likely that protein-protein, protein-RNA, and RNA-RNA interactions all contribute to stress granule assembly. The RNA content of stress granules is largely dominated by mRNAs, with ~80% of stress granule RNA being mRNA (Khong et al., 2017). When analysed by smFISH all mRNAs can localise to some extent to stress granules with 10-15% of global mRNA being recruited to stress granules (Khong et al., 2017, Somasekharan et al., 2020). Recruitment to stress granules can vary greatly in efficiency, with some lowly abundant mRNAs showing < 1% recruitment (e.g. GAPDH), whilst highly abundant mRNAs can show 95% recruitment (e.g. AHNAK) (Khong et al., 2017). Non-coding RNAs have also

been shown to localise to stress granules, however, as only 0.5% of non-coding RNAs localise to stress granules they are believed to play a lesser role (Khong et al., 2017). Similarly, to what was observed for mRNAs, non-coding RNAs that localise to stress granules also tend to be longer. Limited research to date has been conducted to date elucidating whether non-coding RNAs play a regulatory role in stress granule assembly dynamics.

1.3.3 Stress granule disassembly

The exact mechanisms of stress granule disassembly are yet to be fully characterised. Stress granule disassembly has been linked to a number of fundamental cellular processes. As mentioned in section 1.4.1, stress granule formation is regulated by translation initiation signalling cascades. Stress granules are said to be in a dynamic state of equilibrium with polysomes (Riggs et al., 2020). This is evidenced by treatment with translation elongation inhibitors, such as cycloheximide or emetine, which stabilises polysomes by preventing elongation. Both of these compounds prevent stress granule formation in response to sodium arsenite (Kedersha et al., 2000). Conversely, treatment with puromycin, which causes premature termination of translation promotes stress granule assembly (Kedersha et al., 2000). Treatment with cycloheximide or emetine will reduce the number of stalled preinitiation complexes whereas treatment with puromycin will increase the number of mRNA molecules not engaged in the polysomes by initiating premature termination. The stoichiometric ratio of stalled versus active preinitiation complexes is clearly important in regulating both stress granule assembly and disassembly. However, little is actually understood about how this mechanistically regulates disassembly.

Stress granule assembly and disassembly has also been shown to correlate with cellular ATP levels (Wang et al., 2022). Reduced cellular ATP levels promote stress granule assembly (Somasekharan et al., 2020, Wang et al., 2022). Once cellular ATP levels have returned to normal stress granule disassembly occurs (Wang et al., 2022). As mentioned in section 1.4, stress granules are remodelled in an ATPase dependent manner (Jain et al., 2016). Further to this, inhibition of ATP production results in stress granules becoming less dynamic (Jain et al., 2016). This led to the hypothesis that ATP regulated processes may also contribute to

stress granule disassembly (Protter and Parker, 2016). It was theorised that the interactions that result in stress granule formation, protein-RNA, or RNA-RNA, could therefore be disrupted by ATP dependent helicases. Other ATP dependent processes that regulate stress granule dynamics such as the microtubule-dependent motor proteins (Nadezhdina et al., 2010) were also theorised to play a potential role (Protter and Parker, 2016). There is to date no experimental evidence to support these models and further research is required.

In addition to stress granule disassembly, stress granules can be degraded by autophagy dependent mechanisms. This process was originally described in (Buchan et al., 2013) and termed granulophogy. Clearance of stress granules by autophagic mechanisms was conserved in yeast and mammalian cells (Buchan et al., 2013). Knockout or inhibition of the autophagy regulatory protein valosin-containing protein (VCP) resulted in increased number of spontaneous stress granules in non-stressed conditions and stress granule persistence after the alleviation of stress in mouse embryonic fibroblasts. As with stress granule assembly, it is likely that a combination of factors simultaneously contribute to stress granule disassembly or clearance. Further investigation is needed to understand the rate limiting steps in these processes.

1.3.4 What is the function of stress granules?

As discussed in Mateju and Chao, a limitation in the stress granule field is that many of the publications (which number over 5000 based on PubMed data) have been very observational. (Mateju and Chao, 2022). Modulation of stress granule dynamics can revert many observable phenotypes yet the mechanistic underpinnings of how they do this often remains elusive. There is evidence supporting stress granules having two distinct roles, regulating signalling pathways by recruitment of effector proteins and regulation of mRNA stability and translation.

1.3.4.1 Stress granules as modulators of signalling pathways

Recent advancements in methodologies have allowed for the study of the stress granule proteome on an unprecedented scale. Two approaches have thus far been used. Biochemical

purification of stress granule cores in tandem with affinity purification of G3BP1 (Jain et al., 2016) and proximity ligation assays, using tagged stress granule proteins (Markmiller et al., 2018, Padrón et al., 2019, Youn et al., 2018, Somasekharan et al., 2020). Stress granules have a diverse proteome, containing hundreds of proteins that varies dependent on the stress applied (Markmiller et al., 2018, Padrón et al., 2019). Approximately 50% of proteins identified in stress granules are RNA binding proteins (Jain et al., 2016), suggesting that stress granules may have a role outside of RNA condensation. Previous work has shown that stress granules can supress apoptosis through inhibition of the JNK signalling cascade (Arimoto et al., 2008). This is mediated by recruitment of RACK1 to stress granules, which is a scaffold protein required for the activation of the effector kinase MTK1. Stress granule recruitment limits RACK1-MTK1 association and thus supresses apoptosis. Stress granules have also been implicated in the regulation of mTOR signalling. Multiple independent studies have shown the recruitment of the mTOR pathway proteins RAPTOR (Thedieck et al., 2013) and TORC1 (Takahara and Maeda, 2012) which blunts mTOR signalling and prevents apoptosis through mTOR hyperactivation. The scope of pathways stress granules regulate by selective recruitment of regulatory proteins is still yet to be understood.

1.3.4.2 Regulation of mRNA stability and translation.

Stress granules have long been implicated in the regulation of mRNA stability (Anderson and Kedersha, 2009). Stress granules were proposed to be sites of mRNA storage, that protect mRNAs from degradation in response to various cellular stresses. Once the stress was relieved, stress granules disassemble and provide the cell with a repertoire of protected mRNAs that can now be used for protein translation. Although this model has persisted within the field for a number of decades, there is still limited mechanistic evidence. Stress granules have also been implicated in regulating mRNA stability through their interactions with another biomolecular condensate, P-bodies. P-bodies have previously been shown to be able to facilitate RNA degradation (section 1.3.3.3). Stress granules and P-bodies are said to be dynamically linked biomolecular condensates as P-bodies can physically interact with or dock onto stress granules (Kedersha et al., 2005). Very little is known to date about this interaction, but it has been proposed within the field that the physical interaction between these biomolecular condensates facilitates the exchange of mRNA between them. Reporter mRNAs have been shown to move between stress granules and P-bodies (Mateju et al., 2020). This condensate specific localisation of mRNAs may then differentially regulate

mRNA stability (Kedersha et al., 2005). How P-bodies dock onto stress granules is also poorly understood. It has been proposed to be facilitated by RNA-RNA mediated interactions, as overexpression of eIF4A1, which limits RNA condensation *in vitro,* also limits stress granule P-body docking interactions *in cellulo* (Tauber et al., 2020). A number of proteins, including DDX6 are known to localise to both stress granules and P-bodies (Youn et al., 2018). It has also been proposed that certain proteins may coordinate the stress granule P-body interaction by providing multivalent interaction domains (Sanders et al., 2020). Both stress granules and P-bodies are tethered to the ER microtubule network, which facilitates fission and fusion events of both biomolecular condensates (Lee et al., 2020). Whether this also regulates stress granule P-body interactions is yet to be characterised.

Stress granules have also long been proposed to regulate translation by sequestering mRNAs away from the translation machinery which in turn inhibits their ability to engage in translation. It was therefore believed that stress granules inhibited protein synthesis as a means of energy conservation. This model originally developed due to the supposed absence of large ribosomal subunit proteins within stress granules when visualised by immunofluorescent microscopy (Kimball et al., 2003, Kedersha et al., 2002) . It is however important to note that these findings were perhaps overstated, as discussed in (Mateju et al., 2020). It is fair to say that large ribosomal proteins were depleted compared to the cytosol but there is insufficient evidence to support the claim that they are completely absent. More recent studies have indeed shown that large ribosomal proteins can indeed localise to stress granules (Seguin et al., 2014). The model of stress granules repressing translation was also supported by more recent studies, which showed that stress granule mRNAs were translationally repressed compared to stress granule excluded transcripts in response to eIF2α phosphorylation (Khong et al., 2017, Namkoong et al., 2018, Padrón et al., 2019).

The model of stress granules serving as a translationally repressive entity is now being challenged for a number of reasons. Firstly, stress granule formation is not required for bulk translational repression in *G3BP1 G3BP2* double knockout cells (Kedersha et al., 2016). Despite these cell lines being incapable of stress granule formation, translation is still repressed to a similar extent. Furthermore, stress granules have been shown to form in conditions where translation is not repressed, for example, in response to UVC (Moutaoufik

et al., 2014). Only 10-15% of cellular mRNAs localise to stress granules in response to sodium arsenite treatment whereas translation is repressed by ~95% (Khong et al., 2017, Tauber et al., 2020). It is therefore unlikely that enough mRNA is localised to stress granules to facilitate this robust downregulation of translation. Moreover, stress granule mRNAs are translationally repressed in response to eIF2α phosphorylation, but translational repression was not observed for stress granule mRNAs in response to eIF4A1 inhibition by hippuristanol (Padrón et al., 2019). Stress granule mRNAs had no statistically significant difference in translational efficiency compared to stress granule excluded mRNAs. Using single molecule imaging methods, a reporter construct containing the 5'UTR of the ATF4 mRNA was shown to be able to undergo all steps of translation (initiation, elongation, termination) within a stress granule (Mateju et al., 2020). In contrast, a reporter mRNA with a 5' TOP motif was unable to be translated in response to stress. These findings led to the formation of the hypothesis that stress granules may be able to support selective translation of certain mRNAs with concomitant suppression of others. This paper presented the greatest challenge as of yet to the model of stress granules regulating translation by bulk downregulation. Complementary work has also shown that not all G3BP1 bound mRNAs are translationally repressed in response to sodium arsenite treatment (Somasekharan et al., 2020). In response to sodium arsenite stress, some G3BP1 bound mRNAs were shown to be translationally supressed. For example, the pro-apoptotic protein BAX. Another sub-population of G3BP1 bound mRNAs were shown to be actively associated with the polysomes. This sub-population included pro-survival proteins, such as the transcription factors FOS and JUN. This leads to the question as to whether the role of stress granules in regulating translation is supporting a pro-survival gene expression programme? Understanding the contribution of stress granule localisation to translational regulation of mRNAs is the primary focus of the work presented in this thesis.

1.3.5 Stress granules in disease

Stress granules have been of particular research interest due to their pathophysiological implications in human diseases. Perhaps the best studied link to disease to date is the role of aberrant stress granule composition, formation and disassembly or clearance in neurodegenerative disorders. Mutations in RNA binding proteins that promote stress granule formation proteins are observed in multiple disorders. Mutations in Fragile X mental retardation protein (FMRP) are the most common cause of fragile X mental retardation

(Bassell and Warren, 2008). The stress granule proteins G3BP1 and TIA1 co-localise to the protein aggregates present in Alzheimer's disease (Ash et al., 2014). TDP43, another stress granule associated RNA binding protein is present in the pathological aggregates of amyotrophic lateral sclerosis (Liu-Yesucevitz et al., 2010). Mutations in VCP, which regulates granulophy mediated clearance of stress granules, have been shown to be causative in amyotrophic lateral sclerosis (ALS) (Johnson et al., 2010) and frontotemporal dementia (Kimonis et al., 2008). Formation of pathological protein aggregate assemblies is a conserved feature of the neurological disorders mentioned above. A proposed mechanism for the assembly of these aggregates is increased stress granule assembly or reduction of clearance in unstressed cells. Over time, stress granules transition from a liquid to a more solid state and contribute to aggregate formation (Protter and Parker, 2016). How related pathogenic aggregation is to the formation of stress granules is yet to be determined.

Viral infection induces stress granule formation (McInerney et al., 2005). This is believed to be a conserved mechanism of viral resistance as multiple viruses target stress granule formation through a variety of mechanisms. One such mechanism is the targeting of the stress granule nucleating protein, G3BP1, for degradation by viral proteases (White et al., 2007). In response to polio virus infection, stress granules form. However, as the infection persists stress granules dissipate. This is mediated through cleavage of G3BP1 by poliovirus 3C proteinase. Stress granule formation can also be impeded by sequestration of stress granule promoting proteins, including G3BP1. This mechanism was originally shown in response to Semliki Forest virus infection (Panas et al., 2012) but has also been shown in response to other viruses, including SARS-CoV-2 (Zheng et al., 2021). This is facilitated by G3BP1 interacting with viral proteins such as viral nonstructural protein 3 (nsP3) in response to Semliki Forest virus infection (Panas et al., 2012). Viruses also target the eIF2α kinase PKR (Khaperskyy et al., 2012). In response to Influenza A infection, non-structural protein 1 (NS1) inhibits PKR and prevents stress granule formation. Due to the targeting of stress granule formation by multiple viruses, it is clear that stress granules play a regulatory role in response to viral infection. More investigation is required to fully understand the role of stress granules in viral immunity and whether this can be exploited therapeutically.

The tumour microenvironment constantly bombards cells with stress (Ackerman and Simon, 2014). Many of these stresses, including hypoxia and ER stress result in stress granule formation in cultured mammalian cells. Little is known about the potential roles of stress

granules in cancer disease progression. High expression of the stress granule protein G3BP1 is associated with a poor prognostic outcome in lung cancer (Zheng et al., 2019), liver cancer (Dolicka et al., 2021) and sarcomas (Somasekharan et al., 2015). Increased G3BP1 expression correlated with increased metastatic potential in murine sarcoma models (Somasekharan et al., 2015). Sections of both human pancreatic ductal adenocarcinoma and KRAS mutant driven mouse models showed elevated levels of stress granules compared to normal pancreatic tissue (Grabocka and Bar-Sagi, 2016). In addition to this, stress granule formation is upregulated by mutant Kirsten rat sarcoma virus (KRAS) signalling (Grabocka and Bar-Sagi, 2016). KRAS G12 mutant cell lines have an enhanced ability to induce stress granule formation in response to sodium arsenite treatment (Grabocka and Bar-Sagi, 2016). This is mediated through KRAS dependent upregulation of the prostaglandin 15-d-PGJ2 which inhibits eIF4A1. This was shown to be a cell non-autonomous process. 15-d-PGJ2 is secreted and conditioned media from KRAS mutant cell lines is sufficient to induce stress granule formation in KRAS WT cell lines to similar levels observed in KRAS mutant cell lines. Conditioned media also conferred resistance to the chemotherapeutic agent oxaliplatin further implicating stress granules in cancer by promoting drug resistance (Grabocka and Bar-Sagi, 2016). It is important to note that other commonly used chemotherapeutic agents have been shown to promote stress granule formation, including sorafenib (Adjibade et al., 2015) and 5-fluorouracil (Kaehler et al., 2014). For the reasons outlined above, it is evident that stress granules are involved in the formation, progression and therapeutic response of human cancers. A better understanding of their contributions to disease progression and chemotherapeutic resistance may improve patient treatment outcomes and requires more research.

1.4 Environmental stresses in cancer

Cancer cells are subjected to a constant barrage of environmental stresses within the tumour microenvironment. Not only do cancer cells need to overcome environmental stresses to survive, but environmental stress can also promote gene expression programmes that contribute to disease progression.

1.4.1 Hypoxia in cancer

Hypoxia is a stress synonymous with cancer. Cancer can overcome hypoxia by inducing angiogenesis which is considered a hallmark of cancer (Hanahan, 2022), (figure 1.8). In response to low oxygen levels, Hypoxia Inducible Factor 1 Subunit Alpha (HIF1α) is stabilised. In normoxic conditions HIF1α is rapidly degraded. Ubiquitination by the VHL complex in an oxygen dependent manner promotes HIF1α degradation (Hon et al., 2002, Masson et al., 2001). HIF1α is a transcription factor that promotes transcription of genes such as Vascular endothelial growth factor A (VEGFA) that in turn promote angiogenesis (Holmes et al., 2007). Proangiogenic signalling is implicated in the progression and treatment response of a number of human cancers including cancers of the lung, colon and breast (Semenza, 2003). VEGF inhibitors are used in the treatment of human cancers, although they show a modest efficacy. Over time, cancer cells can become resistant to VEGF inhibition (Itatani et al., 2018). HIF1α associates with HIF1β to enable nuclear translocation which thus enables the transcription of response genes such as VEGFA (Lee et al., 2019).

Hypoxia, and indeed hypoxia inducible factors have also been implicated in the regulation of translation in response to oxygen depletion. HIF2α can promote the translation of cancer associated genes such as EGFR by direct binding of its 3’UTR (Uniacke et al., 2012). Additional oxygen dependent enzymes have been implicated in the regulation of translation by hydroxylation of ribosomal protein subunits. MINA and NO66 are known to hydroxylate RPL27A and RPL8 in an oxygen dependent manner. The modification of these residues is suspected to play a role in the translation of specific stress responsive mRNAs and in ribosome biogenesis (Bundred et al., 2018). As discussed in section 1.1.4.3, alternative mechanisms of translation initiation such as formation of the hypoxic eIF4F complex and

IRES mediated imitation are also known to contribute to the regulation of protein synthesis in hypoxia.

Exposure to hypoxia is known to promote stress granule assembly in mammalian cells (Timalsina et al., 2018, Gottschald et al., 2010). Inhibition of stress granule assembly resulted in increased sensitivity to cisplatin and paclitaxel under hypoxic conditions in HeLa cells (Timalsina et al., 2018). Whether stress granules contribute to modifying gene expression programmes in hypoxic regions of tumours is yet to be elucidated.

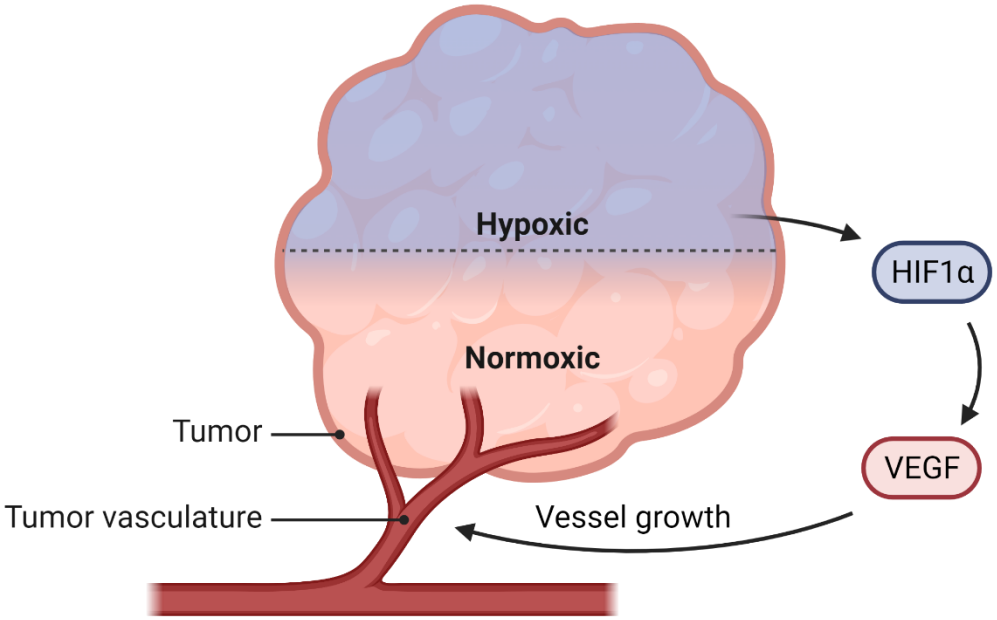


Figure 1.8 Hypoxia promotes tumour angiogenesis:
In response to low oxygen levels in tumours, HIF1α is stabilised. HIF1α promotes angiogenesis by upregulating the expression of VEGF. Angiogenesis is considered a hallmark of cancer and contributes to disease progression.

1.4.2 Oxidative stress in cancer

In addition to hypoxia, oxidative stress has been implicated in cancer progression. Reactive oxygen species can accumulate in a cell from a variety of sources including altered intracellular metabolism or environmental exposure (Hayes et al., 2020). Oxidative stress has also been shown to elicit alterations in gene expression that can promote tumorigenesis. In response to elevated oxidative stress transcription factors such as NRF2 or HIF1α are activated (Marinho et al., 2014). VEGFA has also been implicated in combating oxidative stress by promoting angiogenesis (Schäfer et al., 2003). Agents that cause oxidative stress within a cell such as sodium arsenite and hydrogen peroxide are known stress granule inducers. Sodium arsenite is the most commonly used molecule to induce stress granule formation is a potent inducer of oxidative stress which results in phosphorylation of eIF2α, activation of the integrated stress response and stress granule formation (sections 1.1.4.1, 1.3.2). As with hypoxia, the role that stress granules play in modulating antioxidant responses is still not fully understood.

1.4.3 Lung cancer

Some of the research presented in this thesis implicates stress granules in the pathophysiology of lung cancer, therefore a brief introduction section to lung cancer has been added. Lung cancer represents one of the greatest unmet needs in modern medicine. According to Cancer Research UK data, lung cancer is the most common cause of cancer associated deaths in the UK accounting for 21% of all cancer associated deaths. This represents almost 35,000 deaths annually with a ten-year survival rate of 10%. Treatment outcomes have remained almost stagnant in their progression over the last 50 years (cancerresearchuk.org). As mentioned in section 1.3.5, stress granules have been implicated in poor outcomes in lung cancer. Lung tumours have decreased oxygen levels (median % of oxygen = 2.2) when compared to normal lung tissue (median % of oxygen = 5.6) and are therefore considered a hypoxic tumour (Le et al., 2006). The leading cause of lung cancer, smoking, exposes cells to reactive species (Valavanidis et al., 2013). eIF2α phosphorylation and activation of the integrated stress response is tumorigenic in KRAS driven lung cancers and inhibition with the compound ISRIB extended survival in murine models (Ghaddar et al., 2021). All of this evidence suggests that stress granules may represent a novel therapeutic target to improve treatment outcomes.

1.5 Project aims

The title of this thesis is deciphering the role of liquid-liquid phase separation in oncogenic gene expression. More specifically, the aim is to understand how cells alter their translational output in response to stress via the formation of stress granules. In order to understand the contribution of stress granules, physiologically relevant stresses in cancer will be used to induce stress granule formation. Namely oxidative stress and hypoxia. Methods will also be developed to biochemically purify stress granule cores. This will enable characterisation of both the stress granule proteome, by means of mass spectrometry, and RNA content by RNA sequencing. Ribosome profiling will be used in parallel to understand the gene expression programmes in response to stress. These datasets will then be integrated to understand how stress granule localisation alters translational output of stress granule mRNAs in a stress specific manner. To date, a limited number of studies have integrated stress granule proteomics, RNA sequencing and ribosome profiling to develop a model of stress granule mediated translational control. No high throughput studies have been performed on hypoxia induced stress granules and this represents a great unmet need within the field as understanding how stress granules protect tumours from environmental stresses has the potential to improve patient outcomes in cancer.

2 Chapter 2: Materials and methods

Unless otherwise stated, all chemicals and reagents were purchased from MERCK/Sigma Aldrich. The vast majority of work presented in this thesis was performed by myself. Some experiments were performed by or in collaboration with other people but have been included in this thesis for clarity and because they were used in hypothesis development. It will be clearly stated in this methods section if work was carried out by another person. All schematics were made using Biorender.

2.1 Cell culture methods

2.1.1 General Tissue culture

U2OS cells were cultured in DMEM (Thermo Fisher), supplemented with 10% FBS, 100 µg/ml penicillin-streptomycin and 2mM l-glutamine. A549 cells were cultured in identically supplemented RPMI (Thermo Fisher) media. Cells were routinely checked for Mycoplasma contamination in house and kept in culture for fewer than 20 passages. Cells were passaged at 80% confluency Cells were frozen in complete media plus 10% DMSO in a Mr Frosty (Thermo Fisher) at -80°C for 24 hours, then transferred to liquid nitrogen storage. For experiments cells were plated at a density of $5.0x10^6$ per 15cm dish. Cell number was adjusted by surface area when using different diameter culture vessels. All cell counts were performed using a CellDrop automated cell counter (DeNovix). Unless otherwise stated, samples were harvested by trypsinising cells. Media was removed and added to a 50 ml falcon, which was placed on ice. Trypsin was then added and cells were incubated at 37°C until they detached. Plates were placed on ice and washed with trypsin ten times, ensuring all cells had detached. Cells were then added to back to their media to deactivate the trypsin and spun at 200 G for five minutes to pellet cells. Media was then aspirated and cells were flash frozen in liquid nitrogen. All cell pellets were stored at -80°C until further processing.

2.1.2 Induction of stress and drug treatments

Culture media was exchanged for fresh media 1 hour prior to the induction of stress to improve reproducibility in stress granule formation (Khong et al., 2018). Oxidative stress

was induced by treatment with 200 or 500 µM sodium arsenite for 3 hours or 1 hour respectively. A hypoxic environment was created by culturing cells in a H35 hypoxystation (Don Whitley Scientific) at an oxygen concentration of 0.1% O_2 for 6 hours. Cells were treated with 300 nM Hippuristanol (gift from John Le Quesne) for 30 minutes to inhibit eIF4A1. An equal volume of DMSO was added as a vehicular control for hippuristanol experiments. For puromycin incorporation assays, puromycin was added 10 minutes prior to harvesting at a concentration of 10 µg/ml.

2.1.3 Expression of GFP plasmids

U2OS cells were cultured as described above. Following plating cells were allowed to recover overnight prior to transfection. pEGFP-C1 (GFP, source: Aldo Bader, Beatson Institute) and pEGFP-C1-G3BP1 (GFP-G3BP1, source: Gerald McInerney, Karolinska Institute) were transfected using Lipofectamine 2000 (Thermo Fisher) at a ratio of 3:1 of µl of transfection reagent to µg of DNA. 5-20 µg of plasmid was transfected per 15cm dish which was optimised for each plasmid preparation. Cells were visualised using a Primovert (Zeiss) fluorescent microscope 16 hours post transfection to ensure G3BP1 overexpression was not causing spontaneous stress granule formation. Plasmid was added to 500 µl of Opti-MEM (Thermo Fisher) which had been pre-warmed to 37°C. Lipofectamine was then added to pre warmed Opti-MEM and then mixed with the Opti-MEM plasmid solution and allowed to incubate for 5 minutes. 1 ml of transfection mixture was then added dropwise per 15cm plate. Samples were harvested 24 hours post transfection.

2.1.4 Generation of APEX2-GFP-G3BP1 cell line

U2OS cells were cultured as described above. In order to generate an endogenously tagged APEX2-GFP-G3BP1 cell line a CRISPR approach was used as previously described in (Markmiller et al., 2018). Short guide RNA sequences (G3BP1_gRNA5_F: caccTCCATGAAGATTCACTGCCG, G3BP1_gRNA5_R: aaacCGGCAGTGAATCTTCATGGA) were cloned into pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene) by Bbs1 (New England Biolabs) restriction digest of 1 µg of plasmid for 15 minutes at 37 °C, according to manufacturer's protocol. 50 ng of linearised plasmid was then used for the ligation reaction using T4 DNA Ligase (New England Biolabs) adding

equimolar amounts of the short guide RNA oligos according to manufacturer's protocol. Ligation reaction was carried out at 16°C overnight. G3BP1 homology template HR_G3BP1-V5-APEX2-GFP (Addgene) was linearised by dual restriction digest by BamH1 (New England Biolabs) and SbF1 (New England biolabs) according to manufacturer's protocol. All enzymes were heat inactivated by incubation at 65°C for 10 minutes.

Cells were plated in 6 well plate format at a density of $4x10^5$ per well. After 14 hours in culture cells were transfected with 1.5 µg of Cas9 and linearised homology plasmids respectively. After 24 hours cells were treated with 3 µg/ml puromycin (Invivogen) and left in selection for 72 hours. Cells were washed with PBS (137mM NaCl, 2.7mM KCl, 10mM Na2HPO4, KH2PO4) and fresh media was added. Cells were allowed to recover for 2-3 days before they were re-plated for monoclonal selection. A serial dilution method as described in (John A. Ryan) was used. Half of the media was removed and replaced with fresh media every three days. Cells were left in monoclonal selection for 14 days and then GFP positive clones expanded to 24 well plates, 6well plates and then T75 flasks respectively. Stress granule formation was visualised by treating with arsenite as described above and cells were imaged. Samples were harvested for western blot analysis and promising clones were frozen down.

2.1.5 siRNA transfections

APEX2-GFP-G3BP1 U2OS cells were cultured as described above and plated in 6 well plates prior to knockdown. Cells were allowed to recover overnight before transfection of siRNAs. G3BP1 siRNA (Thermo Scientific, AM16708A) was transfected at a concentration of 20 nM using Lipofectamine RNAiMAX (Thermo Fisher). 3 µl of 20 µM siRNA stock was added to 150 µl Opti-MEM. 9 µl of RNAiMAX reagent was added to 150 µl Opti-MEM. Diluted siRNA and transfection reagent solutions were then mixed and allowed to incubate for five minutes. 250 µl of transfection mixture was then added dropwise to culture media. Cells were harvested 48 hours post transfection.

2.2 DNA methods

2.2.1 Plasmid preparations

100 µl of competent DH5alpha *E. coli* cells were thawed on ice and then transformed with 10 ng of DNA by incubation at 42°C for 45 seconds and were then placed back on ice for 5 minutes. Transformed cells were then added to 5 ml of LB broth containing 50 µg/ml kanamycin or ampicillin and incubated at 37°C for 4 hours with shaking at 200 rpm. 5 ml cultures were then used to inoculate 200 ml of LB broth containing correct concentration of antibiotic. Cultures were then incubated at 37°C with shaking at 200 rpm overnight. 500 µl of the culture was added to 500 µl of 50% glycerol and frozen at -80°C. Cultures were spun at 2000 rpm for 15 minutes to pellet cells. LB broth was decanted and pellets were frozen. All plasmid preparations and Sanger sequencing of plasmids was carried out in house by the Beatson Institute Molecular Technologies Services.

2.3 Protein methods

Please refer to appendix for tables of primary and secondary antibodies used.

2.3.1 Whole cell lysates

Whole cell lysates were generated using RIPA buffer (150mM NaCl, 50mM Tris (pH7.5), 1% NP-40 (v/v), 0.5% Sodium deoxycholic acid (w/v), 1% SDS (v/v), 1 complete mini EDTA free protease inhibitor tablet (Roche) per 10 mL of lysis buffer, 10mM NaF). Pellets were allowed to thaw on ice for five minutes. 500 µl of lysis buffer was added per 15 cm plate. Cell pellets were resuspended in lysis buffer and left on ice for 10 minutes. Samples were then spun at 13,000 rpm for 3 minutes to remove cell debris. If samples were viscous, they were sonicated on high 30 seconds on, 30 seconds off in a BIORUPTOR water bath sonicator (diagenode) for 5 minutes at 4°C. Supernatant was added to a fresh tube. 5X SDS loading buffer (62.5mM Tris-HCl pH 6.8, 7% SDS (w/v), 20% sucrose (w/v), 0.01% bromophenol blue, 5% beta-mecarptoethanol (v/v) was added directly before lysis) was then diluted to 1X in the lysate and boiled at 95°C for ten minutes. All denatured protein samples were stored at -20°C.

2.3.2 Immunofluorescent microscopy

U2OS cells were cultured as described above. Sterile coverslips (VWR, 18x18 mm) were added to the culture vessel. Media was removed and cells were washed with PBS. Glass cover slips were fixed in 4% PFA for 10 minutes and washed three times in PBS. Cells were permeabilised with PBS-0.125% TritonX for 5 minutes, followed by blocking with 1% BSA-PBS for one hour. Primary antibody incubation was done overnight at 4°C. All primary antibodies were used at a 1:100 dilution. Glass cover slips were washed three times in PBS and incubated with secondary antibody for 1 hour at room temperature. Cover slips were washed three times in PBS and dipped in milliQ H2O to dissolve any remaining salt. Cover slips were allowed to air dry and were mounted on microscope slides using VectaShield hard set antifade mounting medium with DAPI (Vector Laboratories). Slides were stored at 4°C until they were imaged. Prior to imaging slides were allowed to warm to room temperature. Images were acquired using a LSM 710 confocal microscope (Zeiss). Laser intensities were set ensuring stress granules and P-bodies were not overexposed. Images were acquired at a bit depth of 12, using averages from four images at maximum speed. CZI files were processed and converted to TIFFs using Fiji (Schindelin et al., 2012).

2.3.3 Stress granule core isolation

Stress granule core isolation method was based on the methods developed in (Khong et al., 2017, Jain et al., 2016). U2OS cells were cultured and transfected as described above. Two 15 cm plates were used per condition. Media was removed and collected in 50 ml falcon tubes. 5 ml of complete media which had been warmed to 37°C was added and cells were scraped and added to the 50 ml falcon tubes. Cells were pelleted by spinning at 1500 G at room temperature. Media was aspirated and samples flash frozen in liquid nitrogen and stored at -80°C. Pellets were thawed on ice for five minutes and resuspended in 1 ml of lysis buffer (50 mM Tris HCl pH 7.4, 100mM Potassium acetate, 2mM Magnesium acetate, 0.5mM DTT, 50 µg/mL Heparin, 0.5% NP40 (v/v), 1:5000 Antifoam B, 1 complete mini EDTA free protease inhibitor tablet (Roche) per 50 mL of lysis buffer. Add RnaseIN (Promega) 1 U/µL right before lysis.). Lysate was passed through a 25G 5/8 needle ten times to aid lysis. All spins post lysis are at 4°C. Lysate was spun at 1000 G for five minutes to remove cell debris and nuclei. Supernatant was transferred to a fresh 1.5 ml tube. 50 µl was taken for protein input. Samples were spun at 18,000 G for 20 minutes and supernatant was discarded. Pellet was resuspended in 300 µl of lysis buffer with 1% NP4O. Samples were

left on ice for ten minutes and intermittently pipetted up and down to aid resuspension. 65 µl of GFP-Trap magnetic agarose beads were washed three times in lysis buffer and added to samples. Samples were incubated at 4°C tumbling end over end for 2 hours. Tubes were then placed on a magnet and buffer was removed. Beads were washed in lysis buffer plus 2M urea for 2 minutes then buffer was aspirated. Beads were then washed in lysis buffer plus 300 mM potassium acetate for five minutes. Buffer was removed and beads were washed three times in lysis buffer. Beads were then boiled in 100 µl 2X SDS loading buffer at 95°C for 10 minutes. Samples that were submitted for mass spectrometry analysis were stored at 4°C, resuspending beads in 50 µl of lysis buffer with no NP40 added. Samples were then handled and processed by the Beatson Institute proteomics facility.

2.3.4 Antibody conjugation to Dynabeads

Protein G magnetic Dynabeads (Thermo Fisher) were used for immunoprecipitations. Beads were washed three times in the buffer being used for the experiment. 4 µg of antibody was added per 8 µl of beads. Antibody dynabeads mixture was incubated for 30 minutes at room temperature tumbling end over end. Supernatant was then removed and beads were washed three times with buffer. Beads were resuspended in buffer and stored on ice until use. Beads were prepared directly before immunoprecipitation.

2.3.5 eIF4A1 immunoprecipitation

A549 cells were cultured as described above. One 15 cm plate was used per immunoprecipitation. Beads were conjugated with eIF4A1, or rabbit IGG isotype control as described above. Cell pellets were allowed to thaw on ice for five minutes. Cells were lysed in immunoprecipitation buffer (20mM Tris pH7.5, 200mM NaCl, 5mM MgCl2, 0.5% Triton-X100, (v/v), 1% BSA (w/v)) on ice for ten minutes. Lysates were then spun at 5000 rpm for 10 minutes to remove cell debris. Supernatant was then added to a fresh tube, 50 µl of lysate was taken for input and 50 µl of beads were added to remainder of lysate. Samples were incubated at 4°C for 1 hour tumbling end over end. Tubes were briefly spun down and then placed on a magnet. Supernatant was then aspirated and beads were washed three times in immunoprecipitation buffer. Beads were then resuspended in 100 µl SDS loading buffer and boiled for 10 minutes at 95°C.

2.3.6 m7G cap pulldowns

U2OS cells were cultured as described above using 1 15cm plate per condition. m7G cap pulldowns were performed by Tobias Schmidt. Pellets were thawed on ice for 5 minutes and then lysed in 1 ml cytoplasmic extraction buffer (20 mM tris-HCl pH 7.4, 40 mM KCl, 3 mM MgCl2, 1% glycerol (v/v) 1 mM NaPPi, 1 mM NaF, 1 complete mini EDTA protease inhibitor tablet (Roche), Made to 50 ml with water) and left on ice for 15 minutes. Lysate was passed through a 5/8 G needle 10 times to aid lysis. Lysates were spun at 16,000 G to clear cell debris. Supernatant was added to clean tube and protein was quantified by Bradford (Bio-Rad) assay. m7G Sepharose beads (GE healthcare) were washed twice in cytoplasmic extraction buffer. 50 µl of beads were used per sample. 300 µg total protein was added to the beads, 10 µg of total protein was saved as input. Samples were incubated for 2 hours with the beads, tumbling end over end at 4°C. Beads were then washed twice with 500 µl cytoplasmic extraction buffer. Beads were resuspended in 30 µl 2X SDS buffer and boiled for 10 minutes at 95°C.

2.3.7 *In vitro* immunoprecipitation

50 µg of 6X-His-G3BP1 (Origene) was used per immunoprecipitation. Equimolar amounts of eIF4A1 (source: Tobias Schmidt Beatson Institute) was added. Reaction was made up to 100 µl using protein storage buffer (20 mM Tris-HCl pH 7.4, 100 mM KCl, 0.1 mM EDTA 1% (v/v) glycerol, 1 mM TCEP). Samples were incubated for 15 minutes at room temperature tumbling end over end. 15 µl of Dynabeads pre conjugated with 6X-His antibody or IGG isotype control were then added and samples were incubated for a further 30 minutes tumbling end over end at room temperature. Samples were then briefly spun down and placed on a magnet. Samples were then washed three times in protein storage buffer. Beads were then resuspended in 30 µl 2X SDS loading buffer and boiled for 10 minutes at 95°C.

2.3.8 G3BP1 protein purification

Please note G3BP1 recombinant protein purification was performed in collaboration with Tobias Schmidt. G3BP1 CDS was PCR amplified (G3BP1_BsaI_fw:

TTGGTCTCATGGTGTGATGGAGAAGCCTAGTCC, G3BP1_NotI_rv: AATAGCGGCCGCTTACTGCCGTGGC) from GFP-G3BP1 plasmid and cloned into pET-SUMO vector using BsaI and NotI (New England Biolabs) restriction digest. G3BP1 protein was expressed as a His-SUMO fusion protein in *E. coli* BL21 (DE3) CodonPlus-RP (Agilent) was induced by adding 1mM IPTG to culture when cultures were at an OD600 of 0.8. Cultures were harvested 4 hours post induction. Cells were lysed in buffer A (20 mM Tris-HCl pH 7.5, 1 M NaCl, 30 mM imidazole, 10% (v/v) glycerol.) 1 mM PMSF and complete EDTA-free protease inhibitor were added directly before lysis. Lysate was then cleared by centrifugation at 75,000 G and 0.45 µm filtered. Lysate was then added to a HisTrap (GE Heatlthcare) column. Bound protein was then eluted by a linear imidazole gradient and fractions were pooled. Pooled fractions were then diluted in buffer B (20 mM Tris-HCl pH 7.4, 10% (v/v) glycerol, 0.1 mM EDTA) and incubated with SUMO-protease (generated in house by Tobias Schmidt) for 1 hour at 8°C to remove His-SUMO tag. Samples were then further diluted in buffer B and applied to a ResourceQ (GE Healthcare) anion exchange column. Bound protein was eluted with a linear KCl gradient. Pooled fractions were purified by means of size exclusion chromatography using a Superdex 200 column which had been pre equilibrated in storage buffer. Samples were then concentrated using Vivaspin protein concentrator spin columns (Cytiva) and snap frozen in liquid nitrogen.

2.3.9 SDS PAGE

Denatured protein samples were loaded onto 4-12% NuPAGE Bis-Tris gels (Thermo Fisher). Samples were run using MOPS or MES buffer (Thermo Fisher) in a novex mini cell gel tank system (Thermo Fisher). Buffer was selected based on what molecular weight proteins needed to be visualised. Precision Plus Protein Dual Color Standards (Bio-Rad) were used as a molecular weight marker. Prior to loading, the wells of the gel were washed using a P1000 pipette and MOPS buffer. Gels were run at 60 V for 15 minutes and then 120 V for 75 minutes.

2.3.10 Western blotting

Proteins were transferred to Amersham Protran 0.2 µm nitrocellulose membranes (GE healthcare) using the Bio-Rad protean II wet transfer system at 100 V for 2 hours (Transfer

buffer recipe: 0.1% SDS, 25mM Tris, 0.192M Glycine, 20% Methanol). Membranes were blocked in 5% (w/v) milk TBS-tween (TBS: 10mM Tris HCl pH7.4, 150mM NaCl, TBS-tween: TBS, 0.1% Tween-20 (v/v)) for 1 hour. Membranes for use with phospho-antibodies were blocked using 1% BSA TBS-tween. Primary antibodies were incubated overnight on a roller in blocking solution at a dilution of 1:1000. Membranes were then washed for 5 minutes in TBS-tween with gentle rocking. Membranes were then incubated with secondary antibodies diluted 1:10,000 in blocking solution for one hour with gentle rocking. Membranes were then washed as described above and imaged using a LI-COR Odyssey imager with Image Studio software. Western blot images were exported as 600 DPI TIFFs.

2.3.11 Coomassie staining

For Coomassie staining, SDS PAGE gels were incubated in 10 ml InstantBlue Coomassie Protein Stain (Abcam) for 15 minutes with gentle rocking. Gels were then washed three times in Milli-Q water and imaged using a LI-COR Odyssey imager.

2.4 RNA methods

2.4.1 RNA precipitations

For total RNA precipitations 50 µl of lysate was added to 500 µl of TRIzol (Thermo Fisher). 100 µl of chloroform was added and samples were vigorously shaken for 30 seconds. Samples were then spun at 13,000 rpm for 15 minutes at 4°C. The aqueous layer was retained and precipitated in 250 µl isopropanol with 1 µl of glycogen (Roche, 20mg/ml) on ice for 30 minutes. Samples were then spun 13,000 rpm for 20 minutes. Supernatant was then removed and pellet was washed in 500 µl 75% EtOH. Pellet was the resuspended in 100 µl of water. 10 µl of 3M NaOAc pH 5.2, 300 µl 100% EtOH and 1 µl of glycogen were added, samples were vortexed and incubated overnight at -20°C. Samples were then spun at 13,000 rpm for 30 minutes, pellet washed with 500 µl 75% EtOH, allowed to air dry for 3 minutes and resuspended in water.

For stress granule core isolation experiments, protocol outlined in section 2.3.3 was used. In place of SDS loading buffer, beads were resuspended in 200 µl proteinase K buffer (10mM Tris-HCl pH7.4, 100mM NaCl, 1mM EDTA 0.2% SDS plus one unit of CHIP grade proteinase

K (Invitrogen)). Samples were incubated at 1100 rpm at 50°C for 20 minutes. 600 µl TrizolLs was added. Trizol RNA extractions were carried out as described above with an additional lithium chloride precipitation step added at the end to remove any residual heparin from the lysis buffer. The RNA pellet from the sodium acetate precipitation step was resuspended in 150 µl of H_2O. 2 µl of glycogen and 50 µl of 10M lithium chloride was added and incubated overnight at 4°C. Samples were spun at 13,000 rpm for one hour. Supernatant was removed and pellets were washed in 500 µl 75% EtOH, vortexed and spun at 13,000 rpm for 10 minutes at 4°C. Pellets were aspirated dry and resuspended in 20 µl of H_2O. All RNA samples were quantified using a NanoDrop One (Thermo Fisher)

2.4.2 Ribosome profiling

2.4.2.1 Preparation of sucrose gradients

10% and 50% sucrose were made using 5ml of 10x gradient buffer (3M NaCl, 150mM MgCl2, 150mM Tris-HCl pH 7.4, 1 mg/ml cycloheximide, up to 50 ml with milliQ H2O). 6 ml of 10% sucrose solution was added to a 14 ml Thinwall Polypropylene tube (Beckman Coulter). 6.4 ml of the 50% solution was layered beneath. Gradients were made using a Biocomp Gradient Master using the 10-50% fast setting. Gradients were allowed to chill to 4°C before samples were loaded.

2.4.2.2 Preparation of samples

U2OS cells were cultured as described above. 2 15cm plates were seeded per condition. One plate was used for undigested gradient and one plate was used for digestion. Media was exchanged for fresh media one hour prior to harvest or induction of stress. Pellets were allowed to thaw on ice for 5 minutes. Cells were lysed in 400 µl lysis buffer (1X gradient buffer plus 1% TritonX 100, 3 µl Turbo DNAse (Invitrogen) per ml and 2 µl of Ribolock per ml) and left on ice for 10 minutes. Samples were then spun at 13,000 rpm for five minutes to remove cell debris and supernatant was transferred to a fresh tube. 7.5 µl of RNAse I (Ambion) was then added to sample for digestion and was incubated at 22°C for 40 minutes, with shaking at 650 rpm. Undigested lysates were left on ice with 3 µl of SUPERaseIn (Invitrogen). Digestion reaction was quenched with the addition of 15 µl SUPERaseIn. 300 µl of lysate was added carefully to the top of sucrose gradients. Tubes were spun in a pre-

chilled SW40Ti rotor (Beckman Coulter) at 38,000 rpm for 2 hours. Samples were run on a Biocomp Gradient Station and fractions were collected using a Gilson fractionator. Undigested fractions were collected in 3 ml 7.7 M Guanidine HCl 9 µl of glycogen and 4 ml of EtOH was added. Undigested fractions were stored at -20°C. Digested fractions were collected and acid phenol chloroform was added at 1:1 volume ratio to fractions 4, 5 and 6. Samples were vortexed and spun at 13,000 rpm for 20 minutes. Aqueous layer was retained and added to a fresh tube. Chloroform was then added at 1:1 volume ratio, samples were vortexed and then spun at 13,000 rpm for 10 minutes. Aqueous layer was retained, added to a fresh tube and RNA was precipitated using the NaOAc method described above overnight. Fractions were merged and resuspended in a total volume of 10 µl.

2.4.2.3 Isolation of ribosome protected fragments

10 µl of 2X urea loading buffer (12% ficoll, 7M urea, 0.03% bromophenol blue and xylene cyanol) was added to samples and the 28-(AGCGUGUACUCCGAAGAGGAUCCAACGU) and 34 (GCAUUAACGCGAACUCGGCCUACAAUAGUGACGU) nucleotide markers. Markers were made to a final concentration of 250 nM Samples were run on pre-cast TBE-Urea gels (Thermo Fisher). Gels were pre-run at 200 V for 15 minutes. Lanes were rinsed prior to loading of samples. Samples were denatured for 90 seconds at 80°C and then placed on ice. Gels were run for 1 hour. Gels were then stained with Sybr Gold (Thermo Fisher), diluted to 1X in TBE for two minutes. Gels were then rinsed twice in TBE and imaged on a Typhoon-Phospho imager. Gel images were printed at actual size and placed under a glass plate to visualise region that needed to be cut. A clean scalpel was used to excise the region of the gel exclusive of the 34 nt marker and inclusive of the 28 nt marker. Gel piece was then placed in a tube and broken up with the scalpel to aid RNA extraction. 500 µl RNA extraction buffer (300 mM NaOAc, 1 mM EDTA, 0.25% SDS) was then added to the gel pieces, ensuring all gel pieces were submerged in buffer. Samples were incubated at 16°C overnight with shaking at 600 rpm. Supernatant was added to a SpinX-column (Corning) and spun at maximum speed in benchtop centrifuge for 1 minute to remove residual gel pieces. RNA was then precipitated overnight with the addition of 2 µl of glycogen and 500 µl isopropanol. Samples were spun at maximum speed for 45 minutes to pellet the RNA. Pellet was then washed twice in 75% EtOH and resuspended in 43 µl of water.

2.4.2.4 PNK treatment

In order to make the RPFs compatible for the generation of next generation sequencing libraries, the 3' end needs to be dephosphorylated and the 5' end needs to have a phosphate group added. This is achieved by using the T4 PNK enzyme, which facilitates both reactions. Samples were incubated at 80°C for 1 minute and then placed on ice. 5 µl 10X T4PNK buffer, 1 µl SUPERaseIN and 1 µl T4 PNK (New England Biolabs) were added. Samples were incubated at 37°C for 15 minutes with shaking at 600 rpm. 5 µl 10 mM ATP was then added and samples were incubated at 37°C for 20 minutes with shaking at 600 rpm. Samples were then diluted to 300 µl with water. Acid phenol chlroform was then added at a 1:1 volume ratio. RNA precipitations were then carried out as described above, precipitating overnight in NaOAc.

2.4.3 Generation of RNAseq libraries

2.4.3.1 Ribosomal RNA depletion

All total RNA, and stress granule RNA samples were ribosomal RNA depleted prior to library preparation using the RiboCop V2 kit (Lexogen). 1 µg of total RNA was used as input for total RNA samples. The entirety of the stress granule isolation precipitated RNA was used as input for stress granule samples. Procedure was carried out according to manufacturer's protocol. In brief, RNA samples and affinity probes are first denatured. The affinity probes are then hybridised with the ribosomal RNA present in the sample. Depletion beads are then added to the mixture. Hybridised probes then bind to the beads removing the ribosomal RNA from the samples.

2.4.3.2 Generation of total and stress granule RNAseq libraries

Total RNA and stress granule RNA libraries were generated using a CORALL Total RNA-Seq library prep kit (Lexogen). Ribsomal RNA depleted samples were used as the starting point. Procedure was carried out according to manufacturer's protocol. In brief, RNA is reverse transcribed by random hybridisation of displacement stop primers. Displacement

stop primers contain a partial Illumina compatible adaptor sequences which are added to the 5’ end of the reverse transcription product. Linker oligos are then ligated, which adds a unique molecular identifier (UMI), which is used for removal of PCR artefacts from RNAseq data, and a partial Illumina compatible adapter sequence to the 3’ end. Samples are then PCR amplified. Primers used in the reaction complete the Illumina adaptor sequence and also incorporate i7 indexes which are used to identify samples in multiplexed sequencing runs.

For stress granule sequencing experiments, an additional qPCR step was added to calculate the optimum number of cycles for PCR amplification. qPCR was performed on a QuantStudio 3 (Applied Biosystems) qPCR machine using SYBR green (Thermo Fisher) dye. Using linear amplification curves 50% of maximum fluorescence was calculated and cycle number was noted. As only 10% of sample is used for the qPCR, three cycles were subtracted from the 50% of maximum fluorescence CT value. 13 PCR cycles were used to amplify total libraries and 15 cycles were used to amplify stress granule libraries.

2.4.3.3 Generation of ribosome profiling libraries

PNK treated ribosome protected fragments were used as the starting point for library generation. NEXTflex-Small-RNA-Seq-v3_19.01 (PerkinElmer) kit was used according to manufacturer’s protocol. In brief 3’adaptors are first ligated. 5’ adaptors are then ligated. Samples are then reverse transcribed using a universal primer that binds the 3’ adaptor sequence. Samples are then PCR amplified using a universal primer for the 5’end and a primer that includes an index sequence for the 3’end to allow for deduplication of multiplexed samples. Libraries were amplified using 12 PCR cycles. PCR products were then size selected by running them on 6% TBE acrylamide gel according to the library kit protocol provided.

2.4.3.4 Determination of library fragment size and concentration

Library fragment size was determined by running 2 µl on a D1000 high sensitivity screen tape (Agilent Technologies). Library concentration was determined using 2 µl of PCR

product in a Qubit 2.0 fluorometer using high sensitivity DNA reagents assay (Thermo Fisher). Both assays were performed according to manufacturer's protocol. Fragment size and concentrations were used to pool libraries to a final concentration of 2 nM. Four times the molar ratio of ribosome profiling libraries was pooled with total RNA sequencing libraries.

2.4.3.5 Next generation sequencing

All next generation sequencing was performed by the Beatson institute molecular technologies services facility. Sequencing was performed on an Illumina NextSeq 500 instrument, using a high output single-end 75 cycle kit (Illumina).

2.4.4 Selective TCP-seq experiments

All selective TCP-seq experiments and data analysis was performed by Joseph Waldron in MCF7 cells. Data was included for clarity, as it was a fundamental part of hypothesis development. Extensive methods are available at (Wagner et al., 2022). Selective TCP-seq combines ribosome profiling methods with immunoprecipitation to identify proteins that are assembling into complexes co-translationally.

2.5 Computational and bioinformatic approaches

All scripts were run on an Ubuntu (version 22.4.2) virtual machine. Shell scripts were run from the Ubuntu terminal. R (version 4.3.0) was installed on the Ubuntu virtual machine and R scripts were run in R studio. All packages required for shell scripts were installed using Conda. All packages required for R scripts were installed using Bioconductor. Packages were updated as required. All data processing and plotting in R was conducted using the Tidyverse environment. Volcano plot for stress granule proteomics was generated using the EnhancedVolcano package. Venn diagrams were generated using Venny 2.0 website. Statistical tests were conducted using the ggpubr R package. Due to the space constraints of this thesis, all scripts have not been included. Scripts can be made available upon request.

2.5.1 Analysis of stress granule proteomics

Raw data from mass spectrometry was handled and analysed by the Beatson institute proteomics facility. Raw data from mass spectrometry were analysed using MaxQuant/Andromeda search engine using an FDR of 0.01. This was used as the input for differential enrichment analysis of proteomics data (DEP) package (Zhang et al., 2018). This package provides a suite of tools for analysing proteomics data. Data was filtered to remove potential contaminants and decoy peptides identified in the MaxQuant output. Proteins were then filtered to only include proteins that were identified in all replicates of at least one condition. Data was then normalised protein wise using a variance stabilising normalisation. Remaining missing values were then imputed using a QRILC imputation method. As mass spectrometry data generates intensities, similar to microarray data, the limma algorithm can be used for differential enrichment analysis. Differential enrichment analysis was performed on proteins identified in sodium arsenite and hypoxia stress granule isolations. Statistical significance thresholds were set at an adjusted p-value of 0.1 and a log2 fold change of 0.5.

2.5.2 Public database mining

TCGA and CPTAC data was accessed through the UALCAN website (Chandrashekar et al., 2017). mRNA and protein expression values in lung adenocarcinoma samples for eIF4A1 and G3BP1 were plotted and statistically tested in UALCAN. Correlation of mRNA expression values was plotted and calculated in UALCAN.

CPTAC quantitative proteomics for lung adenocarcinoma data was accessed through cBioPortal for scatter plot and correlation calculation (Cerami et al., 2012, Gao et al., 2013). Survival analysis of TCGA data was analysed using the XENA platform. G3BP1 and eIF4A1 were split into high and low expression groups by median normalised counts from TCGA RNAseq data.

2.5.3 Protein-Protein interaction prediction

Alphafold predicted structures of eIF4A1 and G3BP1 were downloaded from https://alphafold.ebi.ac.uk/ in PDB file format. Protein-Protein interaction prediction was performed using the HADDOCK v2.4 web service using standard parameters (van Zundert et al., 2016). The first 20 amino acids of eIF4A1 were used as predicted active residues, as this is a predicted IDR. Both basic IDRs of G3BP1 were tested as predicted active residues as these were thought to be the most likely residues to interact with eIF4A1 due to their opposing charge. Predicted structures were exported as images for visualisation.

2.5.4 Analysis of RNAseq data

Scripts used for the analysis of RNAseq data are available at https://github.com/Bushell-lab/Ribo-seq/tree/Ribo-seq2.0. Scripts were originally written by Joseph Waldron. All scripts were downloaded, edited and run by me. For consistency and reproducibility of datasets produced in the Bushell lab all alignments are against the same reference transcriptome (GENCODE v38 human) that has been pre-filtered to only include protein coding transcripts. In order to be classified as a protein coding transcript it must satisfy three criteria. It must be annotated as a protein coding transcript by Havana. It must have a 5' and 3'UTR. The CDS must be divisible by 3, start with a start codon and end with a stop codon. A feature properties file containing gene nucleic acid composition information was curated by Joseph Waldron. This file was used for the feature properties plots presented in this thesis.

2.5.4.1 RNAseq initial data processing

All total and stress granule RNAseq samples were handled as follows. During the library preparation process, indexes (often referred to as barcodes) are incorporated into the fragments. The addition of indexes enables multiplexing of samples allowing them to be included in the same sequencing run. The first step in data analysis is conversion of the raw sequencing data from BCL to fastq format and demultiplexing. The Illumina software package bcl2fastq is used for this process and simultaneously performs file conversion and assigns the reads to their respective samples. FASTQC is used to assess quality of sequencing data, ensuring fragments are of expected size and have a representative

nucleotide distribution. The next step in the process is the removal of the Illumina adaptor sequences and trimming of low-quality bases from reads. This is done using cutadapt (Martin, 2011). During the library preparation process, UMIs are incorporated into the 5' end of fragments prior to PCR amplification. UMIs are short stretches of random bases that allow for identification of fragments that are overrepresented due to bias in the PCR amplification This process is referred to as deduplication as it removes PCR duplicates. For stress granule sequencing experiments UMI sequences were removed using UMI tools (Smith et al., 2017). UMI-Tools extracts the UMI sequence from the read. Reads are then mapped to the transcriptome using bowtie2. After mapping UMI-Tools then performs deduplication and removes overrepresented fragments. For total RNA samples in ribosome profiling experiments cd-hit-dup was used to generate FASTA files containing only unique reads (Fu et al., 2012). UMI sequences are then removed using cutadapt. Reads are then aligned to the transcriptome using bowtie2. Calculation of isoform and gene level expression is the next step in data analysis. RSEM used for this calculation. RSEM output is then used to calculate the most abundant transcript per gene. The protein coding transcriptome FASTA file is then filtered so that it only contains the most abundant transcript per gene. This is required for the alignment of the ribosome protected fragment data.

As positional information is critical in interpreting ribosome profiling data, reads must be aligned to a single transcript. Ribosome profiling reads were demultiplexed and deduplicated as described above for total RNA samples. Reads are then aligned to ribosomal RNAs, tRNAs, and mitochondrial mRNAs using bbmap. Any reads that align are removed from the data. Remaining data is then aligned to the protein coding transcriptome containing only the most abundant transcripts per gene from the total RNAseq data. A python script which has been adapted from the RiboPlot package is then used to generate counts files which are used for all downstream analyses. Ribosome protected fragment quality is then assessed by visualising read length, trinucleotide periodicity and CDS enrichment. As counts of length 27-33 had the features of ribosome protected fragments they were used in downstream analyses. In addition to this, an offset needs to be applied in order to accurately identify the ribosome P-site. This is visualised by identifying the first peak of reads upstream from the AUG start codon, as the P-site of the ribosome will be situated at the start codon. An offset of 12 was used for read lengths 27-30. An offset of 13 was used for read lengths 30-33. CDS counts were then summed for use in differential expression analysis. Ribosome occupancy plots were generated by normalising counts data to transcript abundance from total RNAseq

and then calculating counts per million (cpm). Transcript 5'UTR, CDS, and 3'UTR were then binned into 50 windows (CDS) and 25 windows (UTRs.)

2.5.4.2 Differential expression analysis

RSEM output (total and stress granule RNAseq) or counts files (Ribosome profiling) were loaded into R. DESeq2 was used for differential expression analysis (Love et al., 2014). DESeq2 uses linear regression modelling to estimate changes in expression. Data was filtered to remove any genes that had less than 10 reads across all samples. For ribosome profiling data, counts from the first 20 and last 10 codons were removed to ensure data only represents actively elongating ribosomes. An experimental design accounting for experimental condition and replicate was used. The apeglm log fold change shrinkage method was used to reduce noise within the data. An adjusted p-value < 0.1 was used as a significance cut-off.

Translational efficiency was also calculated using DESeq2 according to the methods described in (Chothani et al., 2019). This method allows for incorporation of ribosome profiling data and total RNAseq data into a single analysis matrix that accounts for sequencing type, condition and replicate.

2.5.5 De novo motif enrichment analysis

FASTA files containing nucleic acid sequence composition of stress granule enriched and depleted mRNAs were generated using the SeqinR R package, using the most abundant transcript per gene from RNAseq data as the reference sequence. These FASTA files were used as the input for STREME analysis (Bailey, 2021). Stress granule depleted transcripts were used as the background. STREME analysis was conducted using a minimum motif width of 6 nucleotides as had previously been used in (Namkoong et al., 2018). All other parameters were left as standard.

2.5.6 Statistical significance calculation of gene overlaps

The GeneOverlap R package was used for calculation of statistical significance of overlap between gene sets. This package uses a Fisher's exact test to calculate significance. Genome size was set to the maximum number of genes identified in the studies being compared.

2.5.7 Enrichment analysis

Stress granule proteomics enrichment analysis was conducted using the MetaCore (Clarivate Analytics) one click analysis pipeline. Enrichment data from the process networks (a pathway database) and disease biomarker databases is presented in this thesis. RNAseq enrichment analysis was conducted using the FGSEA or the enrichR R packages against the GSEA molecular signatures hallmark database. P values <0.05 were considered significant in all instances.

3 Chapter 3: Isolation and proteomic analysis of stress granule cores

3.1 Chapter introduction

Stress granules are cytosolic biomolecular condensates composed of mRNA and protein that form in response to stress. Stress granules are of particular research interest due to their implications in human disease, including neurodegenerative diseases and cancer. High expression levels of stress granule proteins is associated with poor prognostic outcomes in multiple cancers, including sarcomas and lung cancer. Stress granules have also been implicated in contributing to chemotherapeutic resistance. Although many studies on stress granules have been published very little about their functions is known. Stress granules have been proposed to have two functions, regulating cell signalling by sequestering effector proteins and regulating translation through mRNA recruitment (see section 1.3 for in depth introduction to stress granules). The aim of this thesis is to understand how stress granules contribute to stress responsive oncogenic gene expression.

In order to study stress granules, robust methodologies for inducing their formation need to be developed. The two stresses chosen are hypoxia and oxidative stress. Hypoxia is an environmental synonymous with the tumour microenvironment. Previous work has shown that stress granules form in response to hypoxia, however there are no high throughput studies looking at granule composition in response to this stress (section 1.4). The oxygen concentration chosen, 0.1% is based on previously published literature showing that this level of oxygen depletion was required for stress granule formation in human cells (Gottschald et al., 2010). Tumours frequently have lower oxygen concentrations than the surrounding tissue due to poor vascularisation. Oxygen concentrations are known to drop below 0.4%, and this has been shown to contribute to radiotherapy resistance (Le et al., 2006). Oxidative stress is another environmental stress synonymous with cancer (section 1.4). Sodium arsenite, a potent oxidising agent will be used to induce oxidative stress. Sodium arsenite treatment serves two functions in this thesis. Firstly, as it is the compound most frequently used to study stress granules it serves as a positive control. Sodium arsenite treatment induces eIF2α phosphorylation and ATF4 expression which has been shown to be a driver of poor prognostic outcomes in human KRAS driven lung cancers (section 1.4).

Recent studies have enabled the characterisation of the stress granule proteome on a previously unprecedented scale. Two methods have been published to date, differential centrifugation methods (Namkoong et al., 2018, Jain et al., 2016) and proximity ligation based methods (Youn et al., 2018, Somasekharan et al., 2020, Markmiller et al., 2018, Padrón et al., 2019). The method that will be developed in this thesis will involve differential centrifugation, to enrich for stress granule cores, followed by an immunoprecipitation. Therefore, U2OS cells will be used in this thesis to allow comparison with the published datasets based on this methodology. Stress granules contain a diverse proteome rich in RNA binding proteins, and helicases. Helicases have been shown to regulate stress granule dynamics in an ATP dependent manner (Jain et al., 2016). Intriguingly some of the ATPase/helicase complexes identified had opposing effects on stress granule dynamics, with the CCT complex inhibiting stress granule formation whereas the MCM complex promoted stress granule persistence. In order to understand how stress granules are regulating translation, it is important to understand their protein content as stress granules have been shown to have stress specific alterations in their proteome (Markmiller et al., 2018).

The aims of this chapter are:

1. Develop methodologies for stress granule induction.

2. Develop a methodology for isolation of stress granule cores to enable proteomic characterisation of stress granules.

3.2 Morphologically distinct stress granules form in response to arsenite and hypoxia.

In order to perform high-throughput analyses to determine the constituent RNAs and proteins present in stress granules, robust methodologies for inducing stress granules first needed to be established. U2OS cells were transfected with GFP-G3BP1 24h prior to stress induction. The most widely used method for studying stress granules in the literature is treatment with 0.5mM sodium arsenite for 1 hour. This serves somewhat as a positive control but also is of interest in itself as its mechanism of action is through phosphorylation of eIF2α at Ser51. A hypoxic environment was created by culturing cells in a Donn Whitley Scientific H35 Hypoxystation at 0.1% O_2 for six hours. Cells were paraformaldehyde fixed, then imaged by immunofluorescence, visualising stress granules through GFP-G3BP1 and also staining for EDC3 for P-bodies (figure 3.1).

A pan-cytoplasmic localisation of GFP-G3BP1 was observed in the non-stressed control (Ctrl) condition. It is critical to observe this localisation as overexpression of G3BP1 can in itself cause spontaneous stress granule formation (Tourrière et al., 2003). In response to arsenite stress, canonical stress granules, i.e., punctate cytoplasmic foci with docking P-bodies (Kedersha et al., 2005) were observed (figure 3.1).

Conversely, in response to hypoxia, nebulous stress granules which appeared to be larger and individually occupy a larger proportion of the cytoplasm were observed (figure 3.1). Hypoxia induced stress granules also appeared to be deficient in P-body docking. This lack of interaction between stress granules and P-bodies in this condition may also indicate that the RNA present in these morphologically distinct granules could potentially be exposed to different post-transcriptional regulatory mechanisms, as P-bodies contain many proteins from mRNA decay pathways including LSM1-7, which regulate mRNA de-capping and DCP1/2 which are the catalytic subunit of the mRNA de-capping complex (Ingelfinger et al., 2002).

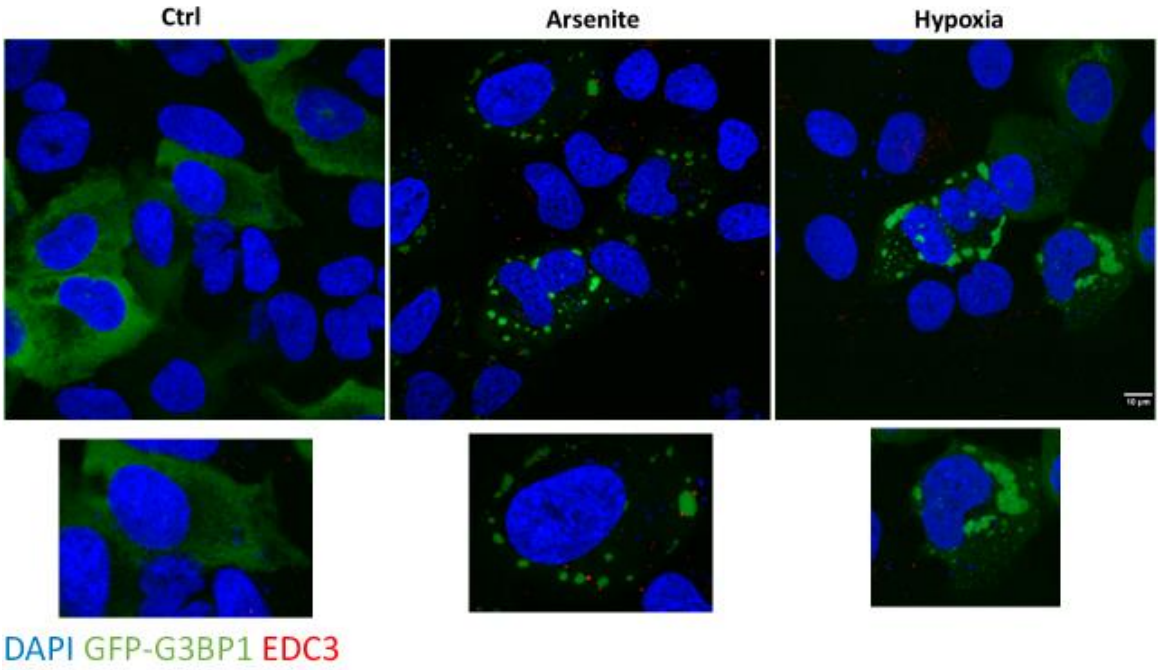


Figure 3.1: Morphologically distinct stress granules form in response to Arsenite and Hypoxia induced stress
Representative immunofluorescence images of U2OS cells expressing GFP-G3BP1, co-stained with DAPI to visualise nuclei and EDC3 to visualise P-bodies. Scale bar represents 10 µm. Representative image of 3 biological replicates.

As numerous upstream signalling pathways feed into stress granule formation, (section 1.3), reviewed in (Panas et al., 2016), U2OS protein lysates were analysed by Western blotting to interrogate the stress signalling landscape in response to arsenite and hypoxia (Figure 3.2).

Treatment with arsenite resulted in increased eIF2α phosphorylation as previously reported (McEwen et al., 2005). In the hypoxia treated samples, a reduction in 4E-BP1 phosphorylation is observed. This is particularly evident in the total 4E-BP1 blot by the downward shift in the banding. This indicates hypophosphorylation of the protein. As 4E-BP1 can also be phosphorylated by the Akt signalling pathway (Gingras et al., 1998), phospho-RPS6 (downstream of mTOR) was also analysed by immunoblotting. A reduction in RPS6 phosphorylation is also observed at the same time points (3, 4 and 6 hours) in hypoxia indicating that there is a reduction in mTOR activity (figure 3.2). The fact that there is a reduction, but not complete ablation of mTOR activity is a key observation, as a recent study implicated mTOR activity and subsequently the two downstream kinases S6K1 & 2 as a requirement for stress granule assembly and maintenance (Sfakianos et al., 2018).

There is a duality of mTOR signalling that needs to be considered in order to mechanistically understand what is leading to the formation of stress granules in hypoxia. Firstly, hypoxia clearly leads to a reduction in mTOR kinase activity as seen through RPS6 and 4E-BP1 phosphorylation states. This reduction in 4E-BP1 phosphorylation results in greater affinity for its binding partner eIF4E which is inhibitory to eIF4G binding and thus eIF4F complex assembly (Marcotrigiano et al., 1999). This is somewhat at odds with the requirement of mTOR activity for stress granule assembly and maintenance. Furthermore, rapamycin treatment or depletion of *RPTOR* did not induce stress granule formation (Sfakianos et al., 2018). Hypoxia may represent a signalling threshold whereby eIF4F complex assembly is inhibited but mTOR may still be active enough to promote and maintain stress granules.

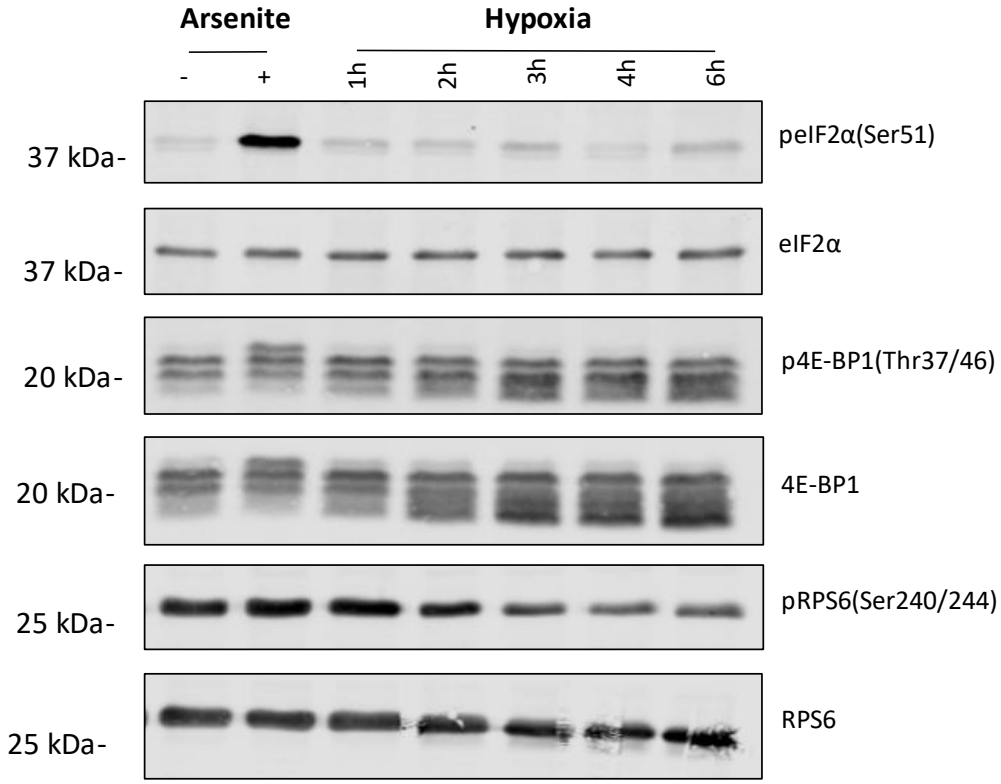


Figure 3.2: Hypoxic stress granules may be a consequence of hypo-phosphorylation of 4E-BP1 due to reduced mTOR kinase activity.
Western Blot of U20S protein lysates with indicated antibodies after being exposed to arsenite or hypoxia. Arsenite results in phosphorylation of eIF2α at Ser51. Hypoxia results in reduction of 4E-BP1 phosphorylation. This is believed to be through a reduction in mTOR kinase activity as a reduction in RPS6 phosphorylation is also observed. Representative image of 3 biological replicates. Equal volume of cell lysate loaded.

In order to determine if the eIF4F complex was disrupted in hypoxia, an affinity purification approach was used. A549 cells were cultured and subsequently exposed to arsenite (0.5mM for 1 hour) and hypoxia (0.1% O_2 for 6 hours). A549 cells were used to expand the scope of the work in this thesis beyond a single cell line. Lysates were then incubated for one hour at 4°C tumbling end over end with ProteinG dynabeads, pre-conjugated with eIF4A1 (eIF4F complex member) antibody or rabbit IGG as a negative control. Dynabeads were then washed three times in lysis buffer and analysed by western blotting (Figure 3.3).

A clear enrichment of eIF4A1 was observed in all three conditions versus the IGG control, thus indicating that the immunoprecipitation had worked. To assess the state of eIF4F complex assembly, two other constituent proteins of the complex were blotted for, namely, eIF4G and eIF4E. In both control and arsenite stressed cells, both eIF4G and eIF4E co-purified with eIF4A1, indicating an intact eIF4F complex in both these conditions (figure 3.3). The same cannot be said for eIF4F complex assembly in hypoxia as neither eIF4G nor eIF4E co-purified with eIF4A1, indicating that complex formation is inhibited.

These data suggest that eIF4F complex formation is being inhibited specifically in hypoxia. Due to the reduced levels of 4E-BP1 phosphorylation observed it is a reasonable hypothesis that the increased affinity of 4E-BP1 for eIF4E is driving this. If this hypothesis is true, an increase of 4E-BP1 would be observed at the mRNA cap.

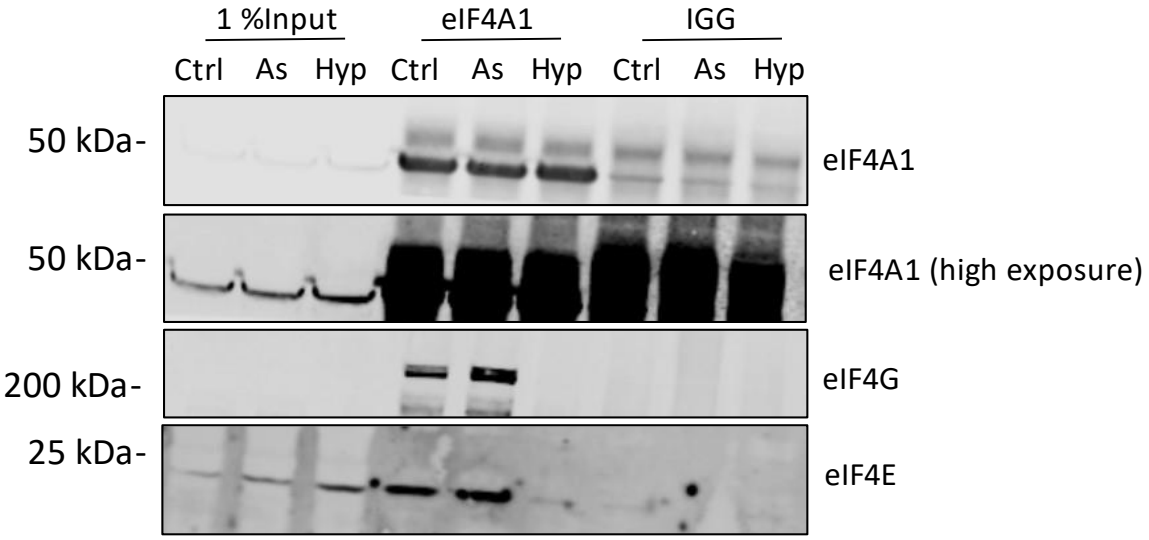


Figure 3.3 eIF4F complex is perturbed in hypoxia:
Western blot of eIF4A1 immunoprecipitations with rabbit IGG as a negative control in A549 cells. eIF4F complex members, eIF4G and eIF4E co purify with eIF4A1 in control (Ctrl) and arsenite (As) treated cells. This interaction is lost in hypoxia (Hyp) suggesting inhibition of complex formation. Representative image of 3 biological replicates. Equal volume of cell lysate and purified fractions loaded.

A widely used methodology for identification of mRNA cap interactors is by m7G-Sepharose affinity purification. In this technique mRNA cap analogues are coupled to agarose beads and used to identify mRNA cap interactors (Sonenberg et al., 1979, Timpano et al., 2016). This approach was used to ask if more 4E-BP1 is present at the mRNA cap in response to hypoxia.

U2OS cells were again stressed with 0.5mM arsenite for 1 hour or hypoxia (0.1% O_2 for 6 hours). Lysate containing 300 µg of total protein was loaded onto the m7G-sepahrose columns and incubated at 4°C for 2 hours. 10 µg of protein was taken as input. Beads were then washed three times with lysis buffer. 4E-BP1 cap interactions were visualised by Western blot using eIF4A1 (a member of the eIF4F cap binding complex) as a positive control (Figure 3.4).

eIF4A1 was successfully eluted in all conditions indicating that cap binding proteins had been purified (figure 3.4). As expected, an enrichment of 4E-BP1 is seen in hypoxia comparatively to the other two conditions. This supports the hypothesis that a reduction in mTOR signalling is leading to hypophosphorylation of 4E-BP1, which is inhibiting eIF4F complex formation through limiting eIF4G binding to eIF4E. Recent work characterising stress granules in response to nitric oxide identified two classes of stress granules. Type I stress granules which form in response to arsenite treatment and initially in response to nitric oxide. Type I stress granules are cytoprotective. Type I stress granules transition to type II stress granules to promote apoptosis. These type II stress granules are characterised by a temporal loss of mTOR dependent 4E-BP1 phosphorylation and an accumulation of 4E-BP1 at the cap. This is coupled with a temporal loss of eIF3 from stress granules (Aulas et al., 2018). It is important to note that although the signalling cascades observed in (Aulas et al., 2018) are similar to those observed in figure 3.2, they are not identical. The key difference is that nitric oxide initially trIgGers eIF2α phosphorylation, which then reduces over time, and loss of 4E-BP1 phosphorylation is a subsequent secondary event. In hypoxia, reduction of 4E-BP1 phosphorylation appears to be the primary event driving stress granule formation and no change in eIF2α phosphorylation is observed.

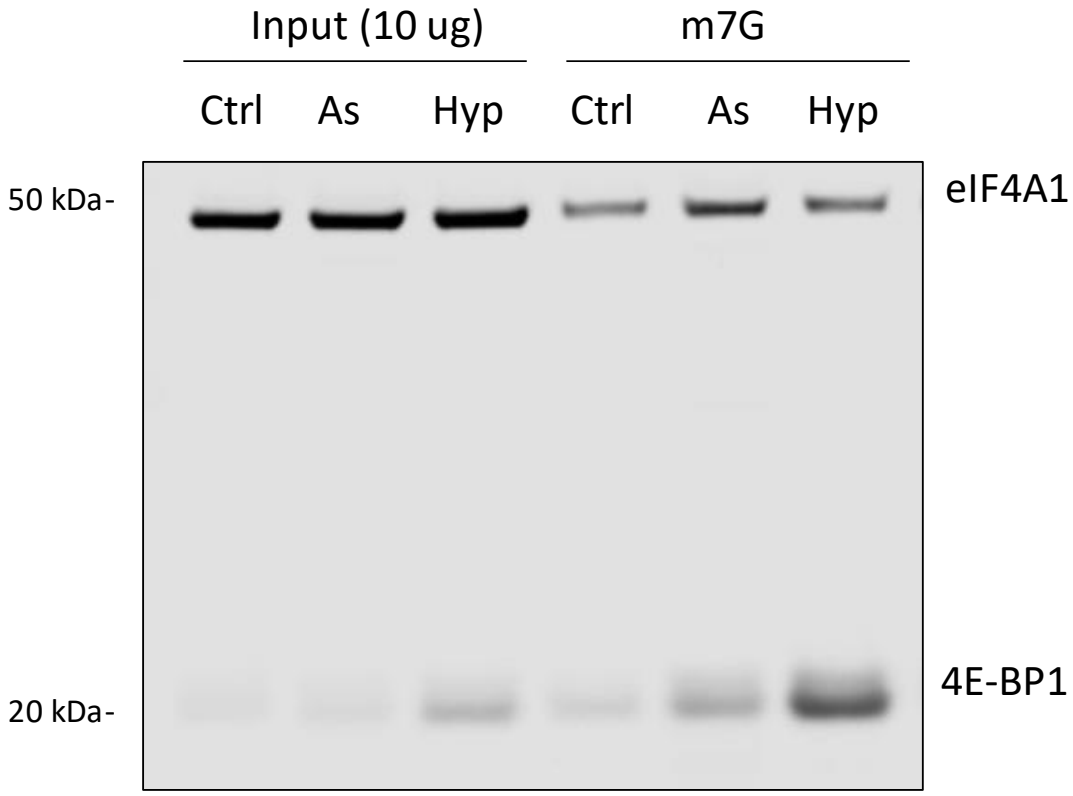


Figure 3.4: 4E-BP1 is enriched at the mRNA cap in hypoxia:
Western blot of m7G cap analogue affinity purification from U2OS cells. An enrichment of 4E-BP1 is observed in hypoxia (Hyp) vs control (Ctrl) and arsenite treated (As) conditions. Representative image of 3 biological replicates. Equal volume of cell lysate and purified fractions loaded.

Type II, pro-apoptotic, stress granules are best characterised by temporal loss of eIF3. As the upstream signalling events observed in hypoxia were similar to those observed in response to nitric oxide stress in (Aulas et al., 2018), an immunofluorescence approach was used to determine eIF3 localisation (figure 3.5). U2OS cells were transfected with GFP-G3BP1 24 hours prior to being exposed to 0.5mM arsenite for 1 hour or 0.1% O_2 for 6 hours respectively. Cells were then fixed and incubated with primary antibody overnight.

GFP-G3BP1 exhibited a pan-cytoplasmic localisation in the non-stressed control. In response to arsenite and hypoxia cytoplasmic stress granules formed similar to those observed in figure 3.1. eIF3B staining showed colocalization with GFP-G3BP1 at stress granules (figure 3.5). These data indicate that hypoxia induced stress granules do not fulfil the criteria to be categorised as type II (pro-apoptotic) stress granules for two reasons. Firstly, they contain eIF3B and secondly, the upstream signalling event kinetics are different in hypoxia compared to nitric oxide stress.

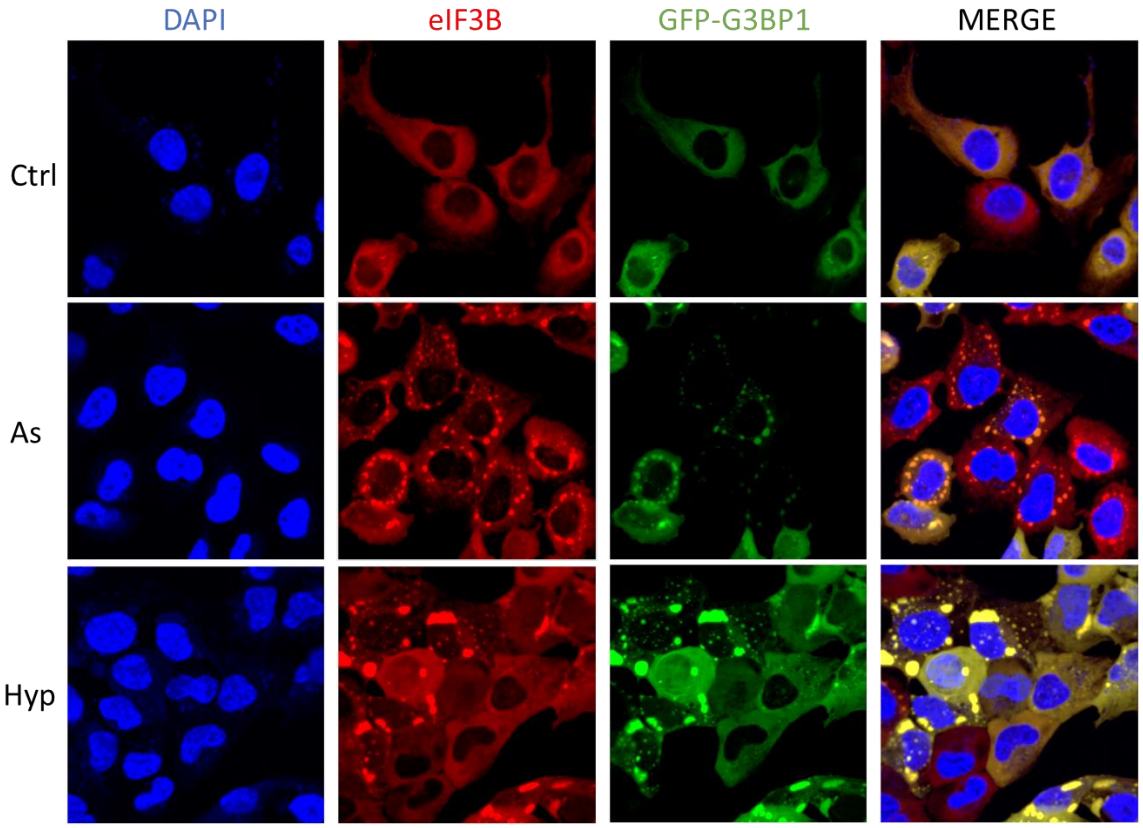


Figure 3.5 eIF3B localises to hypoxic stress granules:
Representative immunofluorescence images of U2OS cells expressing GFP-G3BP1, in non-stressed control (Ctrl), arsenite stressed (As) and hypoxia stressed (Hyp) cells. co-stained with DAPI to visualise nuclei and eIF3B. Representative images of 3 biological replicates.

3.3 Development of stress granule isolation methodology

Now that robust methodologies for inducing stress granule formation in response to arsenite and hypoxia have been developed, the next step was to develop a technique to purify and characterise these morphologically distinct granules using high throughput techniques. A pioneering method was developed in (Jain et al., 2016) for stress granule core enrichment. This method employs multiple steps of differential centrifugation followed by an immunoprecipitation. This method has been used to characterise both the stress granule proteome (Jain et al., 2016) and transcriptome (Khong et al., 2017). Using this technique will enable two fundamental questions to be asked:

1. Do hypoxia and arsenite granules contain differentially enriched proteins that could be driving the observed morphological differences?

2. Do hypoxia and arsenite granules contain differentially enriched RNAs?

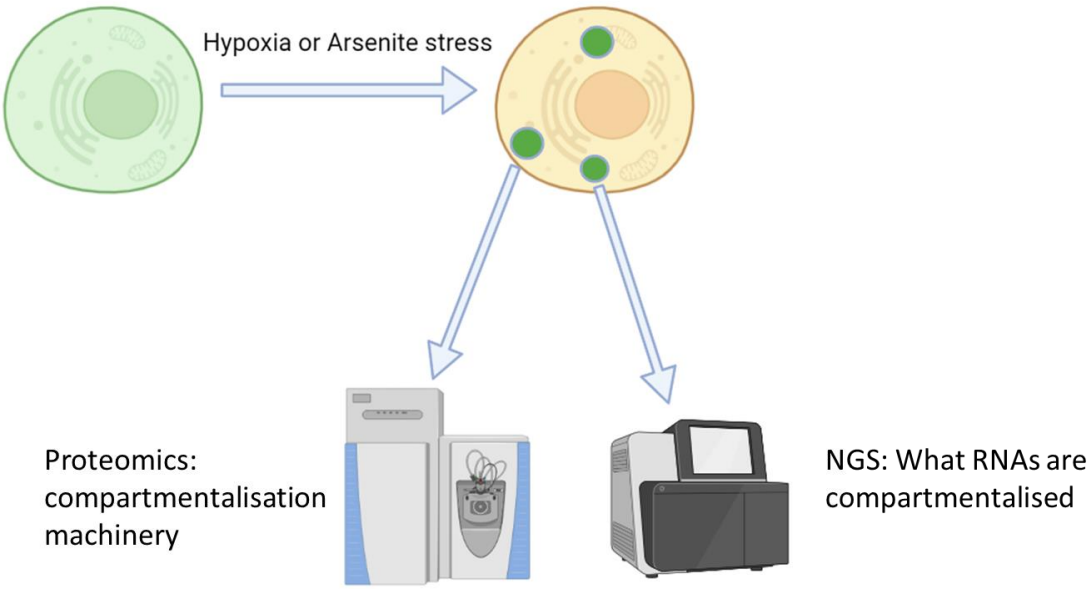


Figure 3.6 Schematic of high throughput techniques following stress granule core enrichment:
Mass spectrometry/proteomics can be used to identify differentially enriched proteins, NGS can be used to identify differentially enriched mRNAs.

U2OS cells were transfected with GFP-G3BP1 prior to induction of stress using 0.5mM arsenite or 0.1% O_2. Media was collected and cells were washed with warm PBS and then scraped and collected in warm media. Media and PBS were equilibrated in the hypoxystation prior to harvesting the hypoxia samples. Cells were then pelleted, media decanted and snap frozen in liquid nitrogen. Stress granule cores were enriched by differential centrifugation, followed by a 2-hour anti GFP immunoprecipitation using GFP-Trap beads at 4°C. GFP-Trap beads were then washed and resuspended in 2X SDS loading buffer. Samples were denatured at 95°C for 10 minutes and then loaded onto SDS PAGE gels. Total protein recovery was visualised by Coomassie staining (figure 3.7).

A clear enrichment of recovered proteins is observed in both arsenite and hypoxia versus the non-stressed control. This indicates that the enrichment process has been successful and indicates low levels of background in the control condition (figure 3.7). In order to account for potential GFP mediated interactors, a GFP only empty vector control was added to the experiment. Samples were prepared as before and analysed by western blotting (figure 3.8).

A clear enrichment of GFP-G3BP1 is observed in arsenite and hypoxia samples versus the non-stressed and GFP empty vector control (figure 3.8). In addition, a clear enrichment of known stress granule proteins eIF4A1 and PABP is evident in the stress conditions. Interestingly, there is an enrichment of eIF4A1 in hypoxia versus arsenite. The fact that no detectable GFP from the GFP empty vector or GAPDH (GAPDH is known to be excluded from stress granules (Jain et al., 2016)) is present further indicates that the stress granule core enrichment protocol is specifically enriching for stress granule proteins (figure 3.8).

A critical optimisation step in this process was increasing the detergent concentration from 0.5% to 1% NP40 to aid in the resuspension of the pellet after differential centrifugation. Samples were also left on ice for 10 minutes at this step to aid resuspension.

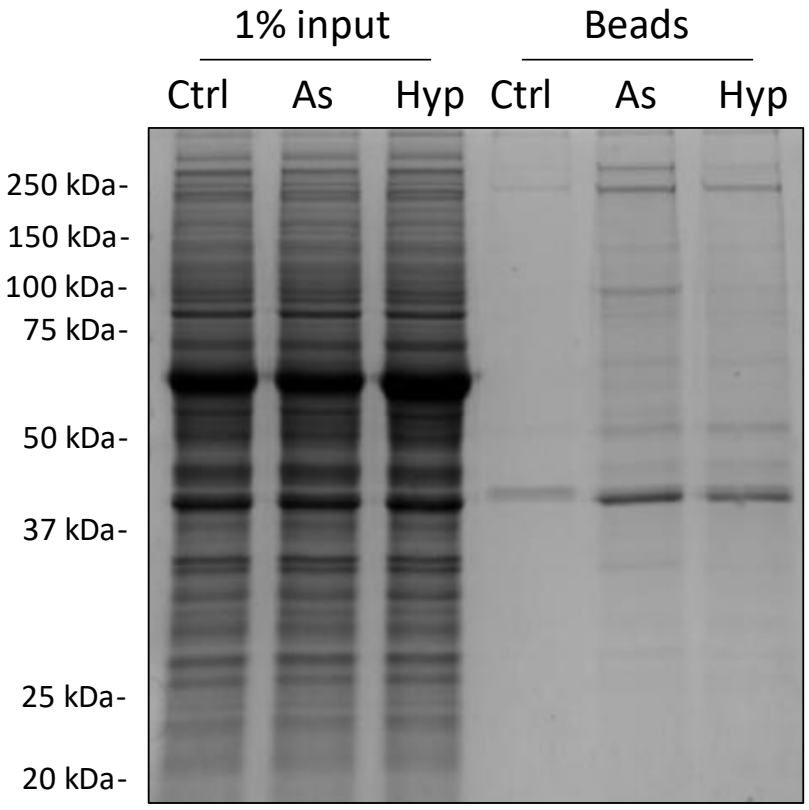


Figure 3.7: Stress granule isolation enriches recovered proteins in response to arsenite and hypoxia:
Coomassie stained SDS PAGE gel of proteins recovered after stress granule core enrichment in non-stressed control (Ctrl), arsenite stressed (As) and hypoxia (Hyp) stressed U2OS cells. Equal volumes of cell lysate and purified fractions loaded. Representative image of 3 biological replicates.

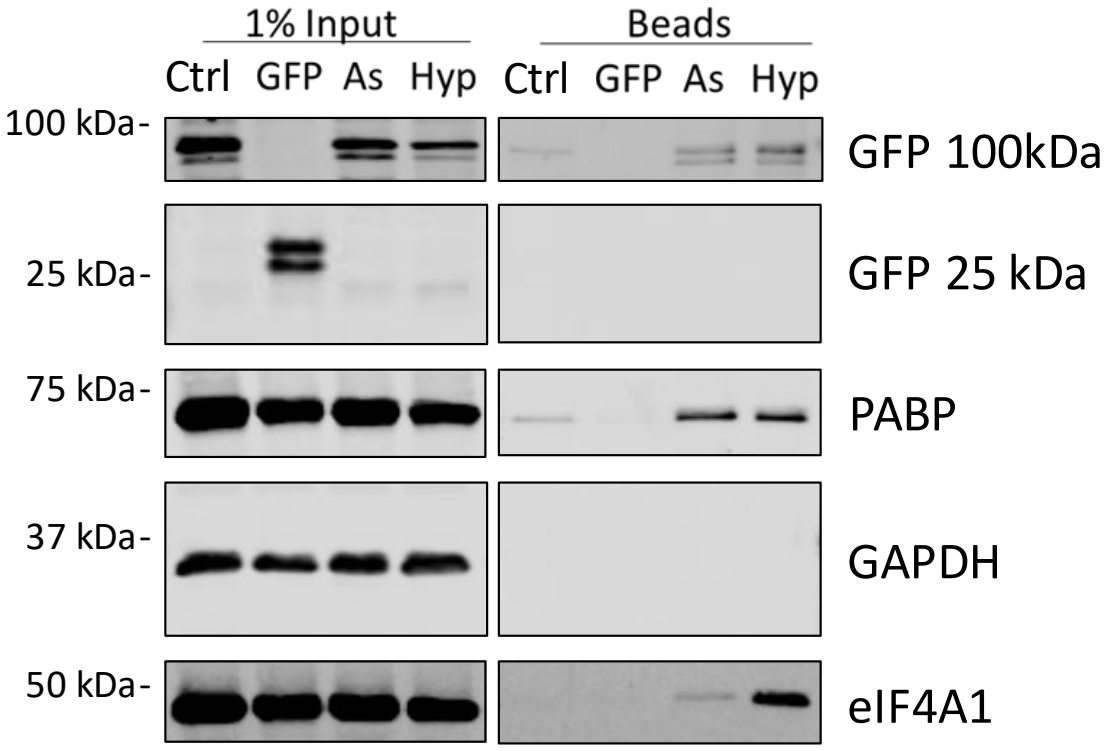


Figure 3.8: Stress granule isolation enriches for stress granule specific proteins in response to arsenite and hypoxia.
Stress granule isolation protocol shows enrichment of GFP-G3BBP1 (GFP 100 kDa) and stress granule proteins eIF4A1 and PABP in arsenite (As) and hypoxia (Hyp) stressed U2OS cells versus non stressed control (Ctrl) and GFP empty vector control (GFP). eIF4A1 is enriched in hypoxia versus arsenite. GAPDH which is excluded from stress granules was not detected in any condition. GFP (GFP 25 kDA) was also not detected after enrichment in the empty vector control. Equal volumes of cell lysate and purified fractions loaded. Representative image of 3 biological replicates.

3.4 Stress granule proteomics: Initial data processing and normalisation

Samples were prepared as described in section 3.2. As an initial trial mass spectrometry run had numerous detergent peaks, after stress granule isolation beads were stored in stress granule isolation buffer with no detergent. Mass spectrometry was conducted by the Beatson Institute proteomics facility. Raw data was analysed using MaxQuant/Andromeda (Cox et al., 2011) against the Uniprot human database accounting for common contaminants. An FDR cut-off of .01 was used for peptide identification.

Analysis of the MaxQuant output was conducted using the Differential Enrichment analysis of Proteomics data (DEP) R package (Zhang et al., 2018). This package provides an all-in-one analysis pipeline that enables data preparation/filtering, normalisation, missing value imputation and differential expression analysis of enriched proteins. Data was first filtered to remove potential contaminants and decoy peptides identified in the MaxQuant output. Proteins with missing intensity values (i.e., were not identified in one of the conditions) were filtered to only include proteins that were identified in all replicates of at least one condition. As MaxQuant generates label free quantification intensities, proteomics data can be analysed in a similar manner to microarray intensities, and many of the packages developed for analysing microarray data are now being applied to proteomic datasets (van Ooijen et al., 2018). Data was normalised using a variance stabilising normalisation (Huber et al., 2002). Remaining missing values were visualised of proteins with at least one missing value (figure 3.9).

This heatmap shows that the samples with the most missing values are three of the control conditions. Namely, two of the GFP empty vector controls and one of the GFP-G3BP1 non stressed controls. As more missing values are present in these control conditions, this could indicate that specific proteins are being enriched for in a condition dependent manner which is in agreement with the results observed in figures 3.7 and 3.8. It is likely that many proteins will be below the detection limit in these control conditions.

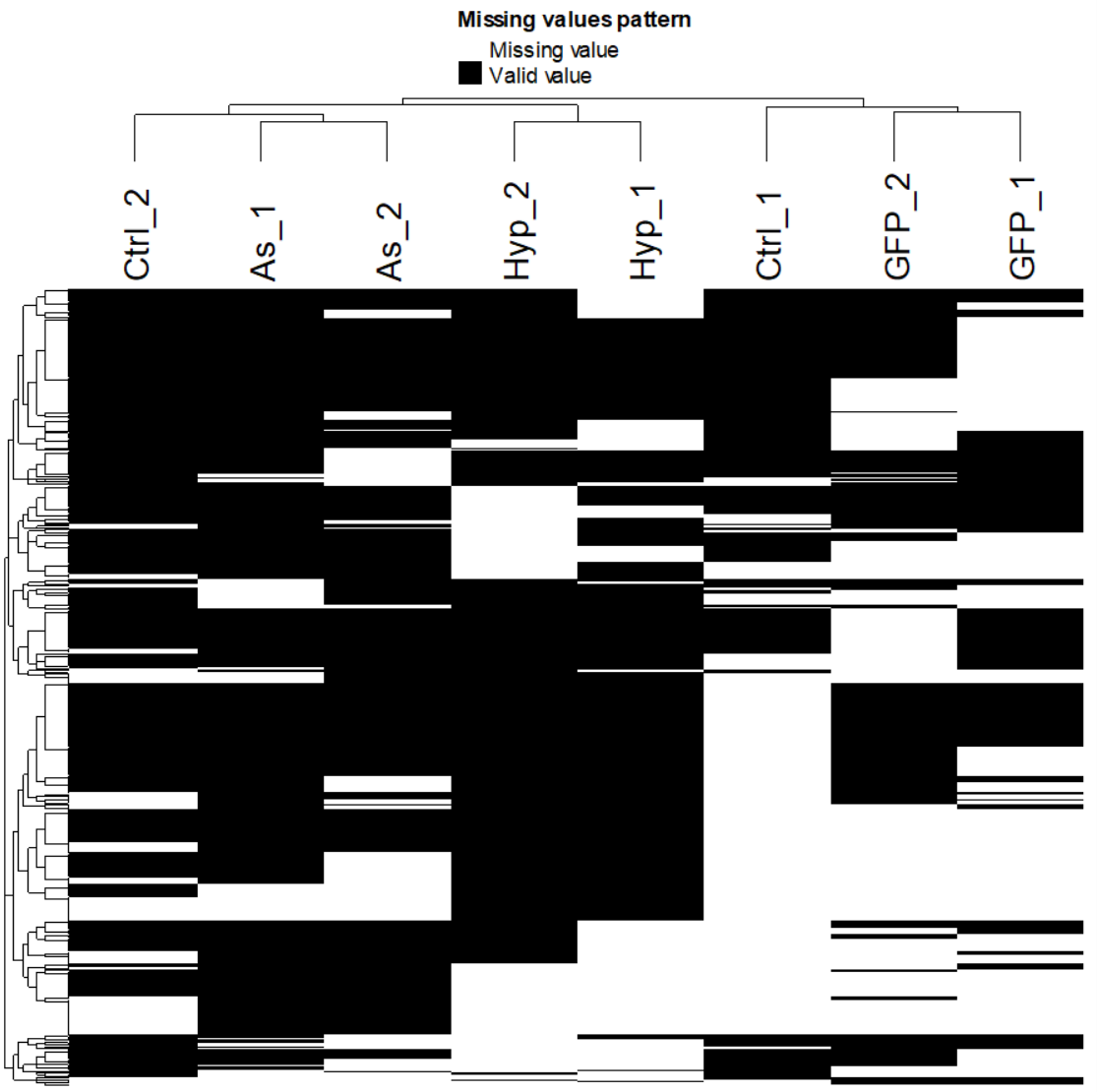


Figure 3.9: Heatmap of missing values in stress granule proteomics data set.
Heatmap shows most missing values are present in three of the control conditions. GFP empty vector control (GFP_1 &2) and the the non-stressed control (Ctrl_1). This indicates that the data is missing not at random. As and Hyp denote the the arsenite and hypoxia controls respectively.

The selection of the method of imputation is a critical step in analysing proteomics data in this manner. Missing values can be defined as “missing at random” if the proteins are identified in a specific replicate but not in another. Alternatively they can be “missing not at random” which indicates that proteins are not identified in specific conditions (Gatto and Lilley, 2012). Although the proteins with missing values appear to be biased towards particular conditions it is important to look at the distribution of these missing values to select the correct method of imputation (figure 3.10).

This density plot shows that the proteins with missing values are biased towards proteins that have lower intensities. This is evident as the distribution of proteins is leftward shifted comparatively to the distribution of proteins without missing values. Figures 3.9 and 3.10 indicate that these missing values are missing not at random and in general the proteins with missing values are close to the detection limit, therefore, a quantile regression-based left-censored (QRILC) imputation was used (Zhang et al., 2018, Gatto and Lilley, 2012). The imputed data can now be used to determine differentially enriched proteins. After data processing and filtering 1285 proteins were included for differential enrichment analysis. The approach adopted was to use limma. limma uses protein/gene wise linear models coupled with empirical Bayesian statistics to estimate centered intensity values for the differential enrichment analysis (Ritchie et al., 2015). A PCA plot for the top 500 variable proteins was generated to visualise the degree of clustering between samples (figure 3.11).

A high degree of clustering was observed for the GFP empty vector control, arsenite and hypoxia treated samples. A lower degree of clustering was observed in the non-stressed control condition. This may be a consequence of the imputation, as replicate 1 had a higher number of missing values comparatively. This may also indicate a greater degree of background in the second replicate of the control condition as it is closer to the arsenite and hypoxia samples on the PCA in comparison to the first replicate.

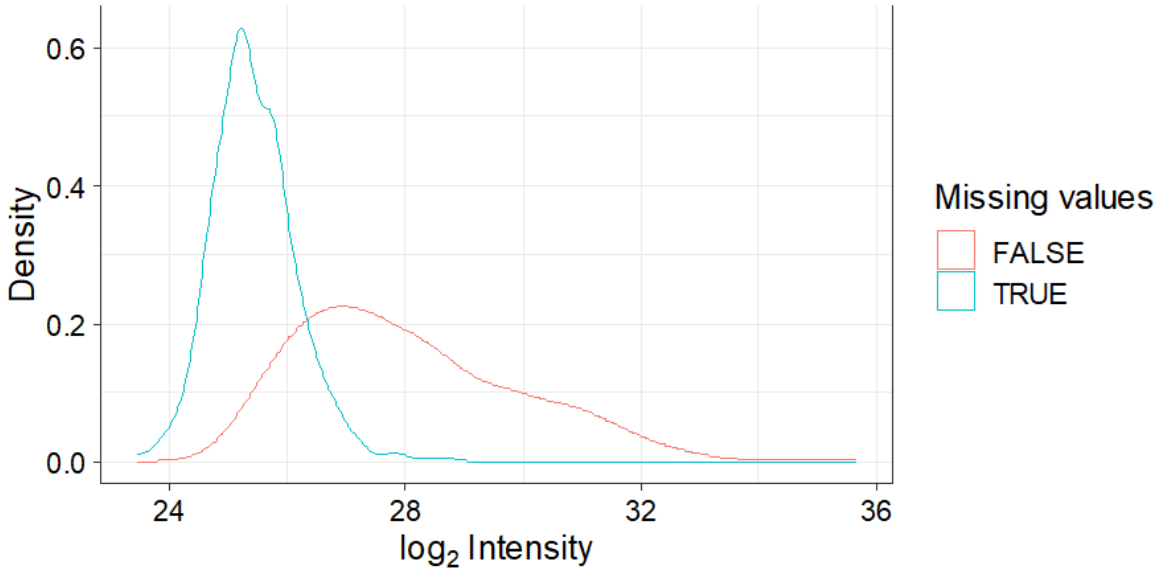


Figure 3.10: Missing values from the stress granule proteomics are biased towards proteins that have lower intensities:
Density plot showing that proteins with missing values have a left shifted distribution compared to proteins without missing values. This indicates that these proteins may be close to the detection limit.

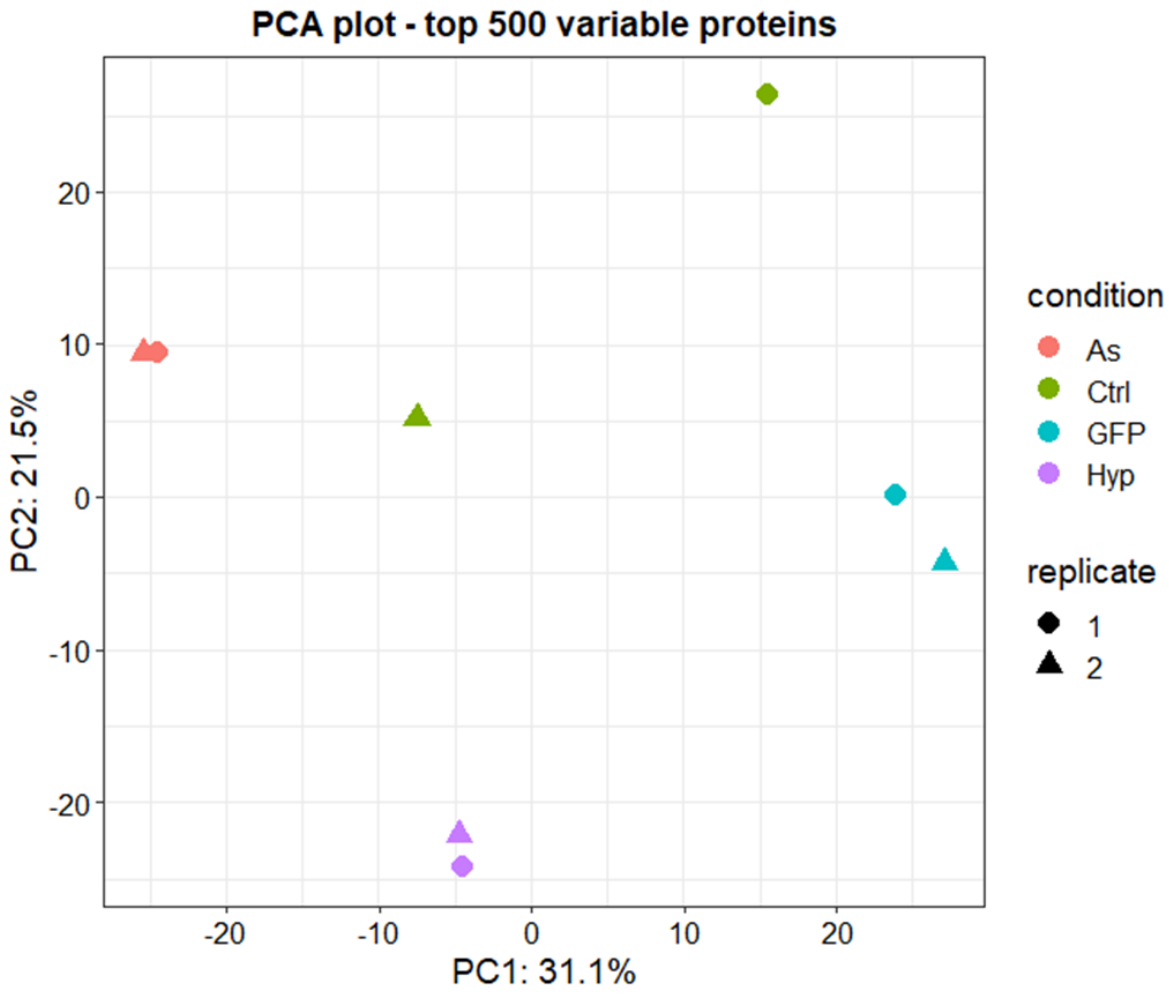


Figure 3.11 PCA analysis of the top 500 variable proteins shows high degree of clustering in stress granule proteomics data set.
A high degree of clustering is seen for arsenite stressed (As), hypoxia stressed (Hyp) and GFP empty vector (GFP) control samples. The non-stressed control appears more variable. This may be a consequence of imputation or higher background in replicate 2.

A correlation matrix was also plotted to assess the level of Pearson correlation between samples (figure 3.12). Reassuringly, each condition had the highest correlation between its own replicates, with arsenite and hypoxia showing the lowest level of correlation. It is important to note that there are relatively high levels of correlation between all conditions. This indicates that a large number of the peptides identified were identified with similar intensity levels in all samples. A possible explanation for this is an underappreciation of the levels of background in the control conditions when using less sensitive methods as in figures 3.7 and 3.8.

A heatmap visualising log transformed centered intensity, with proteins as rows and conditions as columns was then plotted (figure 3.13). This heatmap shows that despite the high degree of correlation observed, there are proteins that appear differentially enriched in the hypoxia and arsenite samples specifically. In the arsenite condition a group of proteins in the bottom left corner of the heatmap and in hypoxia a group of proteins in the top right corner have higher intensity values when compared to all other samples.

As a final quality control measure, the identified proteins were subjected to pathway enrichment analysis using the MetaCore Process Networks one click analysis method (figure 3.14). This enrichment analysis showed an enrichment in terms relating to regulation of translation and mRNA processing which are known functions of stress granule proteins (Jain et al., 2016).

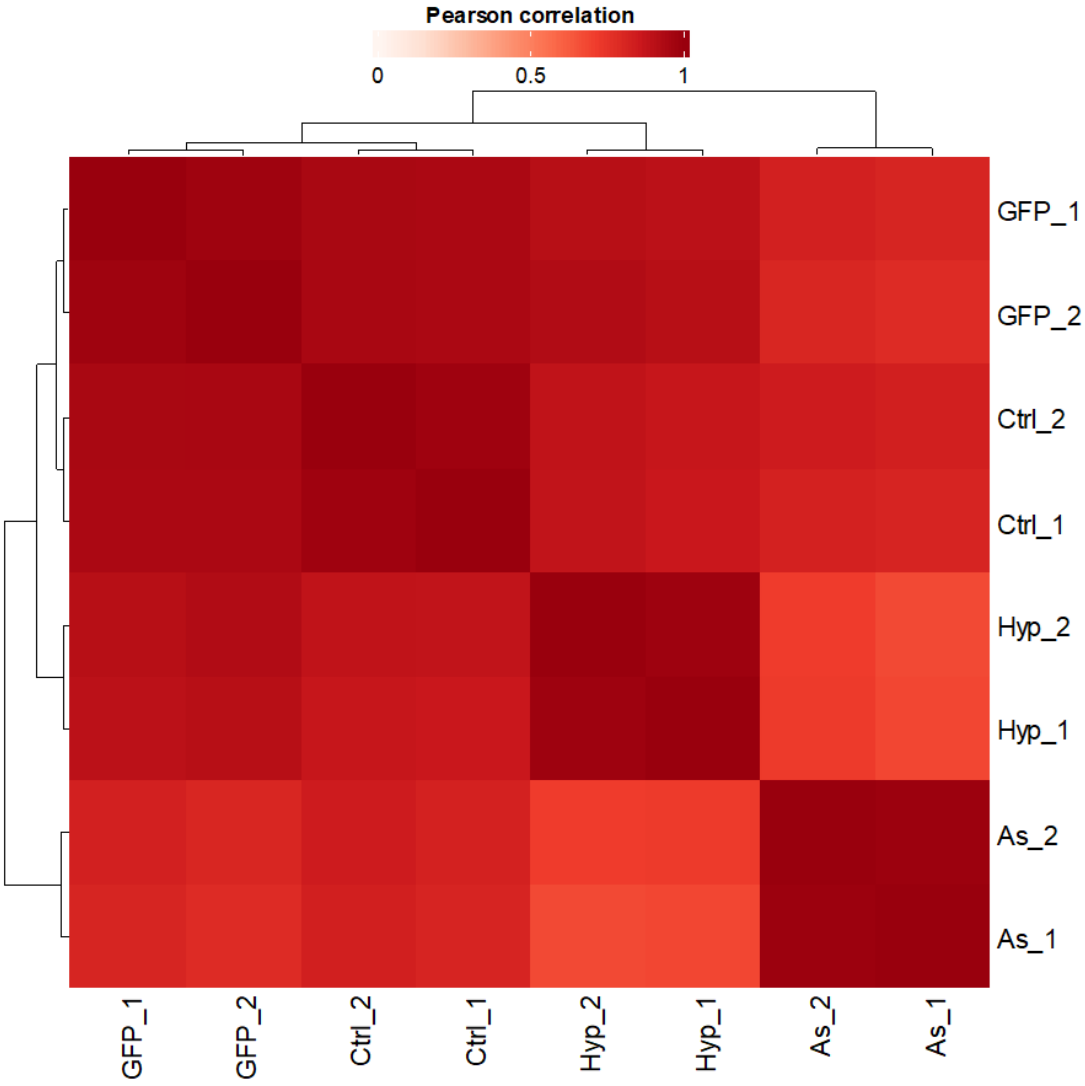


Figure 3.12: Pearson correlation matrix of centered intensity values shows high correlation between all conditions.
Each condition had the highest correlation between its on replicates. The conditions with the lowest correlation are hypoxia stressed (Hyp) and arsenite stressed (As) samples. GFP and Ctrl denote GFP empty vector control and non-stressed control respectively.

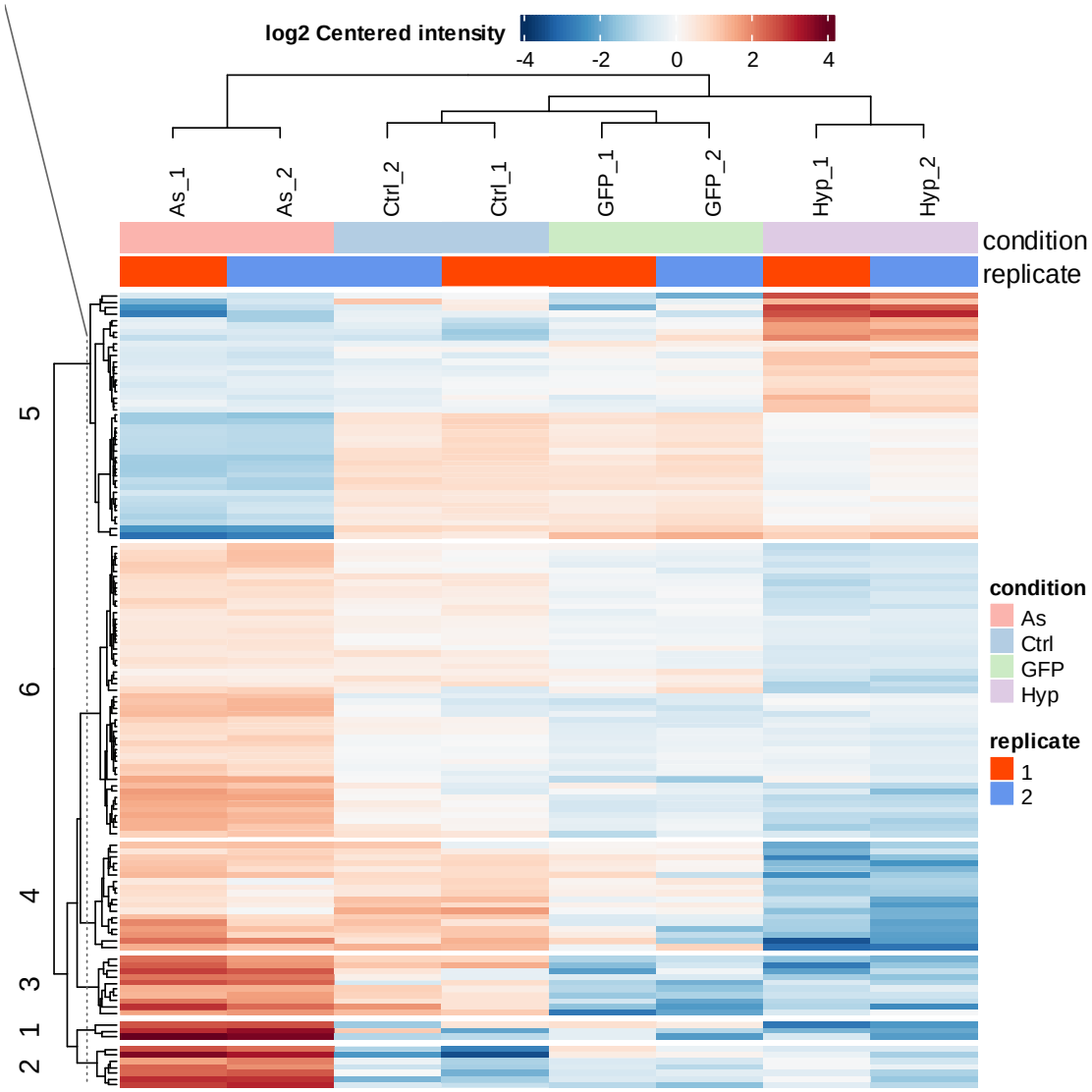


Figure 3.13: Arsenite and hypoxia stress granule enrichments contain uniquely enriched proteins.
Heatmap of log2 centered intensity values calculated by limma with proteins as rows and conditions as columns. Uniquely enriched proteins are identifiable in arsenite (As, bottom left corner) and hypoxia (Hyp, top right corner). Ctrl and GFP denote non-stressed and GFP empty vector controls respectively.

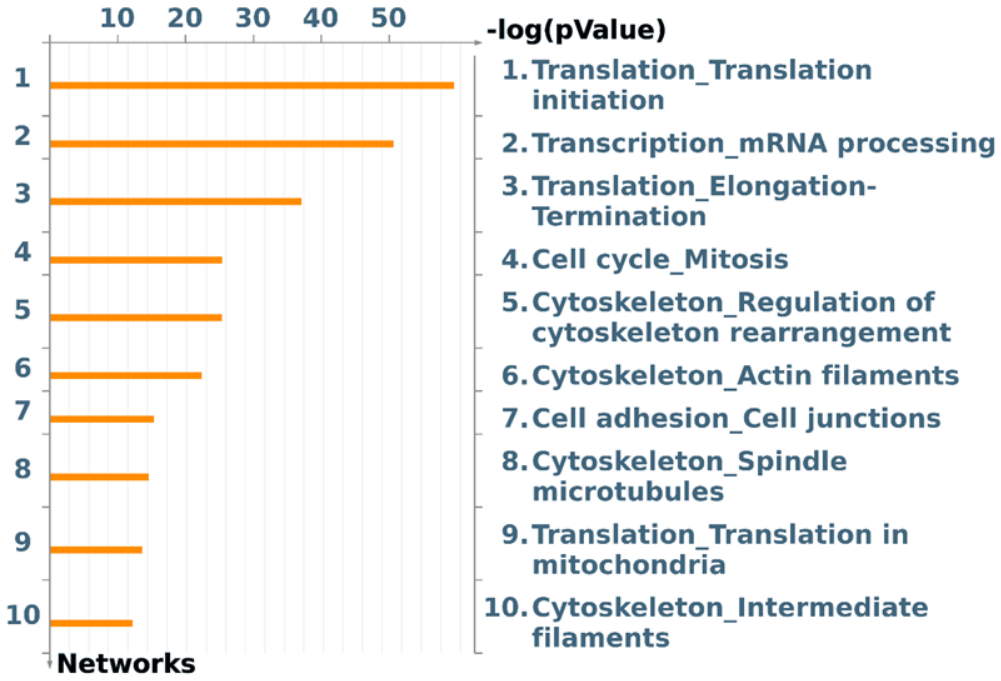


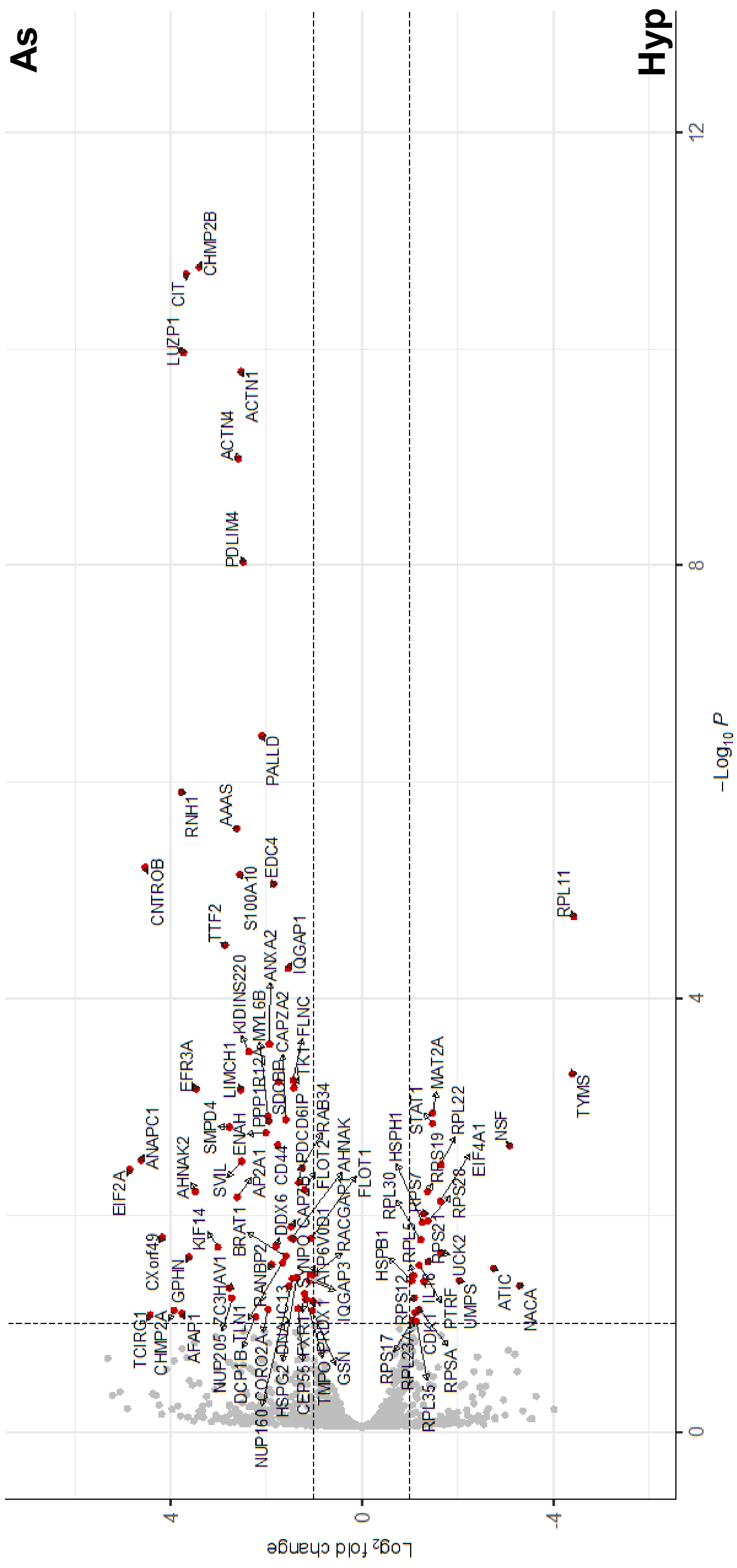
Figure 3.14 MetaCore pathway enrichment analysis of stress granule proteomics shows an enrichment of stress granule associated terms.
Process networks enrichment analysis of proteins identified in stress granule isolation proteomics.

3.5 Arsenite and hypoxia induced stress granules contain differentially enriched proteins.

Differential enrichment analysis identified 63 enriched proteins in arsenite, and 26 enriched proteins in hypoxia induced stress granules. Significance thresholds of an adjusted p value of 0.1 and a $\log_2$ fold change of 0.5 were used. These findings are summarised in a volcano plot in figure 3.15.

Three of the differentially enriched proteins in the arsenite condition, DDX6, EDC4 and DCP1B are predicted to localise to both P-Bodies and stress granules (Youn et al., 2019). Enrichment of these proteins corroborates the phenotype observed in figure 3.1, that hypoxia induced stress granules appear deficient in P-Body docking. Bar plots of the log2 centered intensity values are shown in figure 3.16. This further highlights that the enrichment is specific to the arsenite sample versus hypoxia and both the non-stressed and empty vector controls.

6 proteins (RPL11, RPL22 RPL23A, RPL30, RPL35 and RPL5) which are components of the large ribosomal subunit were identified as enriched in hypoxia. The original papers that characterised the large ribosomal subunit as being absent from stress granules used immunofluorescence approaches against large ribosomal subunits. (Kedersha et al., 2002, Kimball et al., 2003). These studies showed that large subunits are depleted comparatively to the cytoplasm but that they are absent entirely is an overstatement of the findings as discussed in (Mateju et al., 2020). Two subsequent more recent studies have shown that large ribosomal subunits can localise to stress granules (Seguin et al., 2014) and that active translation can occur within stress granules using a reporter mRNA with the ATF4 5'UTR (Mateju et al., 2020). A bar plot of the centered intensity value for RPL11 shows that it is indeed depleted in the Arsenite condition whereas similar levels are observed in hypoxia and the two control conditions (figure 3.17). Western blotting shows recovery of RPL11 specifically in hypoxia and neither of the control conditions (figure 3.18). These data taken together suggest that large ribosomal subunit proteins are not depleted to the same level in hypoxia induced stress granules comparatively to arsenite.



As
Hyp
Log2 fold change
−Log10 P
4
0
-4
0
4
8
12
CHMP2B
CIT
LUZP1
ACTN1
ACTN4
PDLIM4
PALLD
RNH1
AAAS
CNTROB
EDC4
S100A10
TTF2
RPL11
IQGAP1
ANXA2
KIDINS220
MYL6B
CAPZA2
FLNC
TKT
EFR3A
LIMCH1
PPP1R12A
SDCBP
PDCD6IP
RAB34
TYMS
MAT2A
STAT1
RPL22
RPS19
NSF
EIF4A1
ANAPC1
EIF2A
AHNAK2
SMPD4
ENAH
SVIL
AP2A1
CD44
DDX6
AHNAK
FLOT2
FLOT1
HSPH1
RPL30
RPS7
RPS28
RPS2
UCK2
RPL5
HSPB1
KIF14
CXorf49
GPHN
TCIRG1
CHMP2A
AFAP1
ZC3HAV1
BRAT1
RANBP2
NUP205
TLN1
DCP1B
CORO2A
NUP160
DNAJC13
HSPG2
CEP55
TXRT
TMPO
PRDX1
GSN
IQGAP3
RACGAP1
ATP6V0D1
SYNPO
CAPZB
RPS12
RPS17
RPL23A
RPL35
CDK1
RPSA
PTRF
UMPS
IL18
ATIC
NACA

Figure 3.15 Volcano plot of differentially enriched proteins in Arsenite and Hypoxia stress granule isolations.
Significance cut offs of an adjusted p-value of 0.1 and a log2 fold change of 0.5 were used.

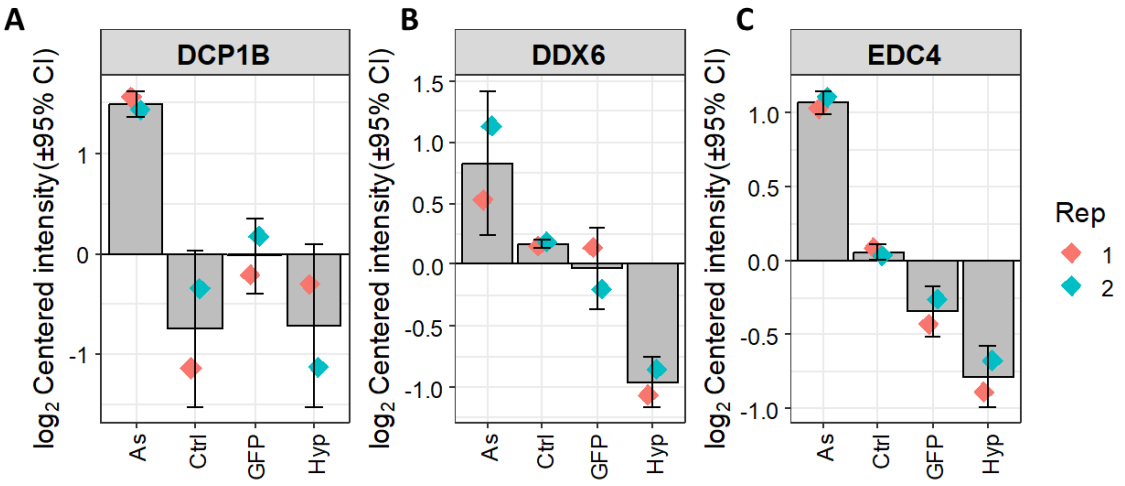


Figure 3.16 Bar plots showing log transformed centered intensity values of representative proteins show specific enrichment of proteins in arsenite stress granule proteomics.
Error bars represent 95% confidence intervals. A: DCP1B B: DDX6 C: EDC4.

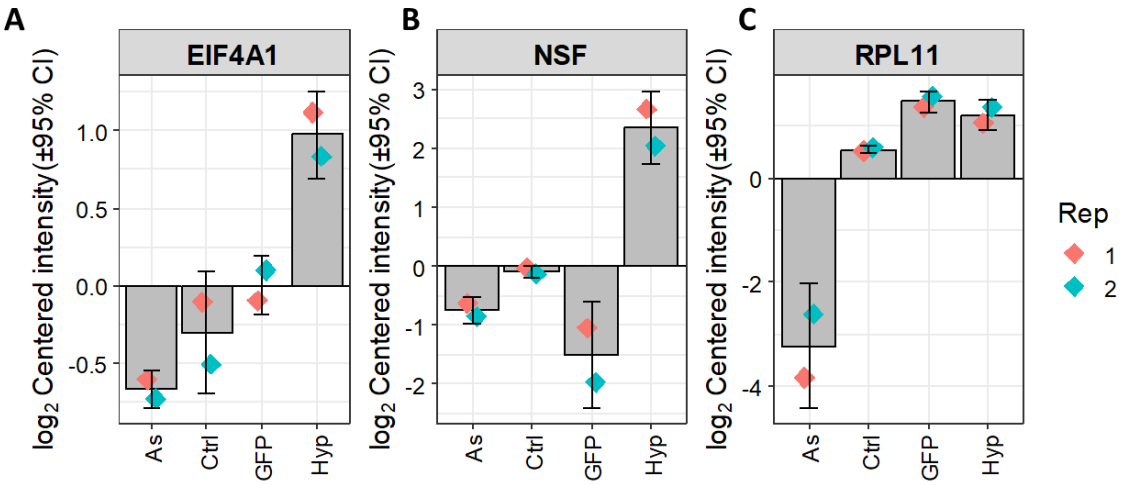


Figure 3.17 Bar plots showing log transformed centered intensity values of representative proteins show specific enrichment of proteins in hypoxia stress granule proteomics.
Error bars represent 95% confidence intervals. A: eIF4A1 B: NSF C: RPL11 is depleted in arsenite (As). Same level of depletion is not observed in hypoxia (Hyp) induced stress granules.

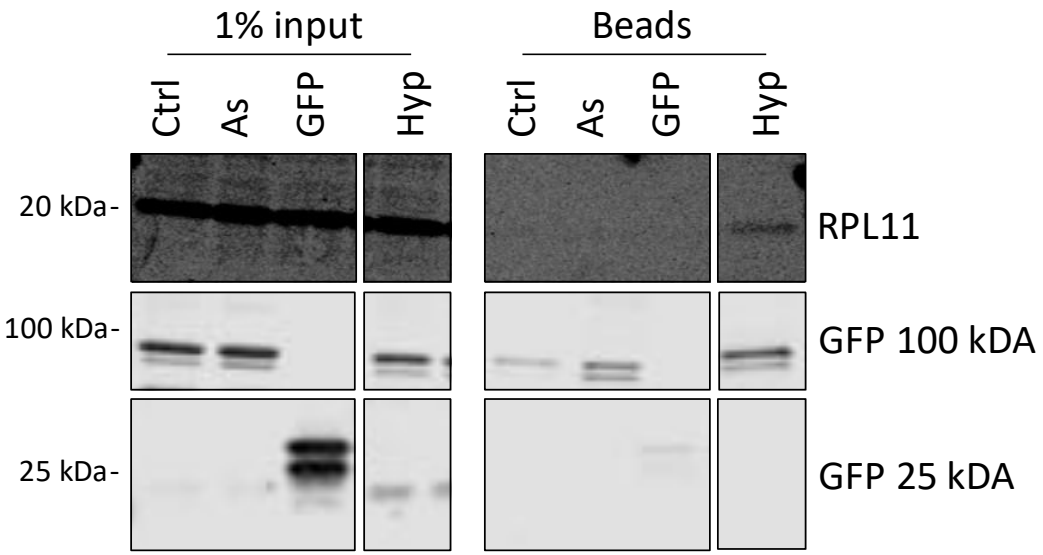


Figure 3.18 RPL11 is enriched in hypoxia induced stress granules by western blotting: RPL11 is recovered in hypoxia stress granule enrichment but is not detected in the nonstressed control (Ctrl) arsenite stressed (As) or GFP empty vector control (GFP). GFP-G3BP1 (GFP 100 kDa) and empty vector GFP (GFP 25 kDa) are shown as controls. Equal volumes of cell lysate and purified fractions loaded. Representative image of 3 biological replicates.

The DEAD box helicase and member of the eIF4F complex eIF4A1 was identified as enriched in hypoxia induced stress granules figures 3.15 and 3.17. This validates the western blot result in figure 3.8. DEAD box helicases have been implicated as master regulators of phase separation (Hondele et al., 2019). eIF4A1 has been shown to regulate stress granule dynamics and P-body docking (Tauber et al., 2020). In this study, inhibition of eIF4A1 with hippuristanol increased stress granule P-body docking. As hypoxia induced stress granules were deficient in body docking and had an enrichment of eIF4A1, it was hypothesised that hippuristanol could revert this phenotype.

U2OS cells expressing GFP-G3BP1 were stressed with arsenite and hypoxia as in section 3.3. 30 minutes prior to the completion of the stress inductions, cells were treated with 300 nM hippuristanol. Stress granules and P-bodies were visualised through GFP-G3BP1 and EDC3 respectively (figure 3.19). Treatment with hippuristanol resulted in stress granule formation in the non-stressed control condition as previously reported in (Mazroui et al., 2006). Stress granules with docking P-bodies were present in the arsenite treated samples with or without the addition of hippuristanol (figure 3.19). The addition of hippuristanol in the hypoxia stressed conditions reduced stress granule size and docking P-bodies could also be seen (figure 3.19). eIF4A1 and its relationship with G3BP1 will be the main focus for the rest of this chapter.

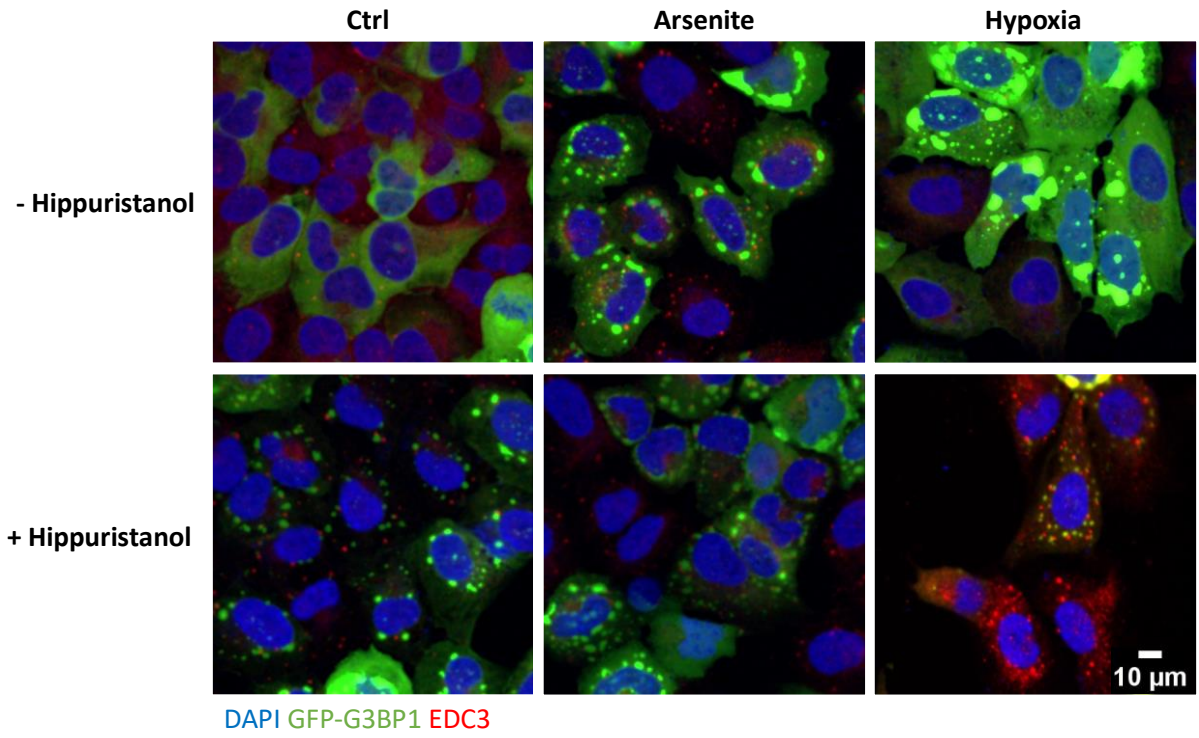


Figure 3.19 Hippuristanol treatment reverts P-Body docking and reduces stress granule size in Hypoxia.
Immunofluorescence images showing stress granules using GFP-G3BP1, co-stained with EDC3 to visualise P-bodies. Nuclei are stained with DAPI. Scale bar represents 10 µM.

3.6 Understanding the relationship between eIF4A1 and G3BP1.

The primary aim of this thesis is to try and understand the role of stress granules in regulating translation in cancer. Cancer is a vastly complex disease, arising in multiple tissues with huge disease heterogeneity (Hanahan, 2022). In order to refine the scope of this project, a MetaCore disease biomarker enrichment analysis, using the proteomics data generated in sections 3.4 and 3.5 was carried out (figure 3.20). This enrichment analysis showed a number of terms related to lung cancers. Namely, lung-, respiratory tract- and thoracic- neoplasms which were the top three enriched terms by P-value.

Using UALCAN (Chandrashekar et al., 2017), a portal for exploring large cancer datasets such as those generated by TCGA (Weinstein et al., 2013) and CPTAC consortia, analysis of these datasets showed that both eIF4A1 and G3BP1 are overexpressed in lung adenocarcinoma comparatively to normal lung tissue in terms of RNA and protein levels (figure 3.21). In addition to both of these stress granule proteins being overexpressed, they show a strong positive correlation at the level of RNA (figure 3.22 A) and protein (figure 3.22 B). CPTAC protein expression data was accessed through the online repository cBioPortal (Cerami et al., 2012, Gao et al., 2013), with a Pearson correlation of R=.37 and R=.49 for RNA and protein expression respectively.

In order to determine if these two proteins could be working cooperatively to promote tumorigenesis, TCGA data was accessed through Xena (Goldman et al., 2020). TCGA Lung adenocarcinoma data was stratified by median RNA expression values and a Kaplan Meir survival plot was generated (figure 3.23). The strata with high G3BP1 and eIF4A1 expression represented the worst prognostic outcome whereas low expression of both of these genes represented the best prognostic outcome.

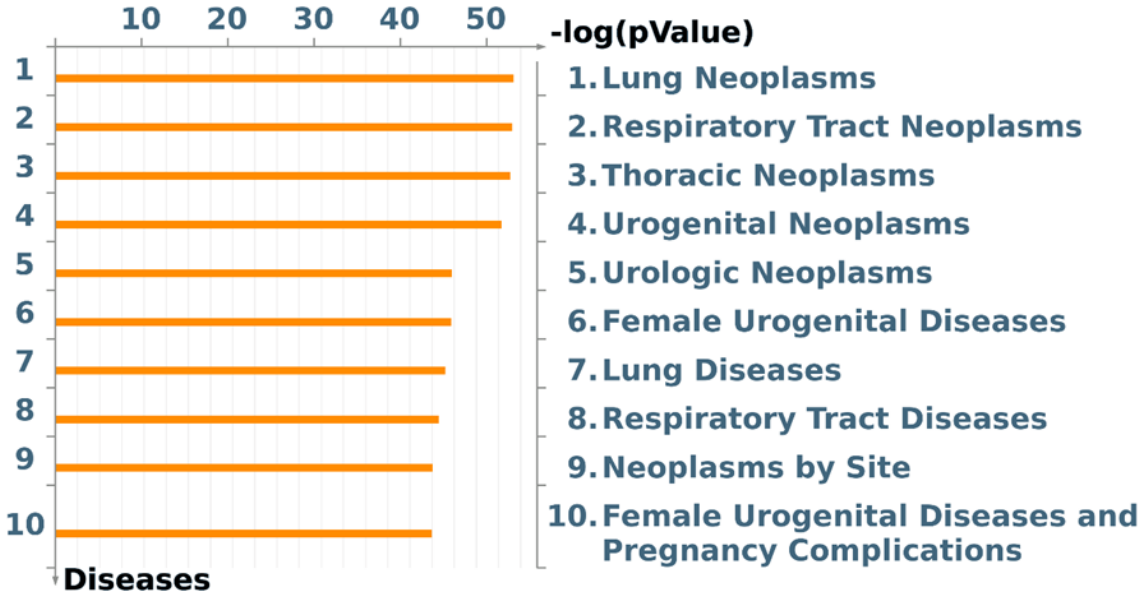


Figure 3.20: Proteins identified in stress granule proteomics show an enrichment for terms associated with lung cancers:
MetaCore disease biomarker enrichment analysis of proteins identified in stress granule proteomics.

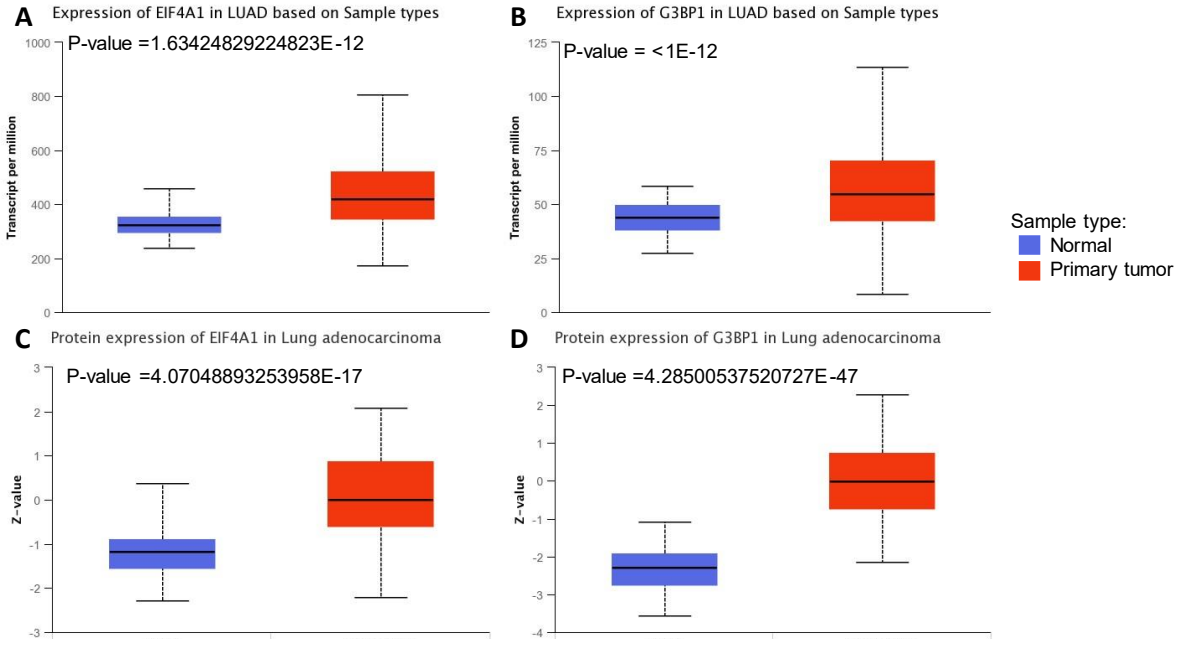


Figure 3.21: eIF4A1 and G3BP1 are overexpressed at the level of mRNA and protein in lung adenocarcinoma.
Boxplots represent the median, maximum, minimum and interquartile range. Reported p-values are generated using a Student's T-test assuming unequal variance. A and B show normalised RNAseq reads (transcripts per million) from TCGA in normal tissue (n=59) versus primary lung adenocarcinoma tumour (n= 515) for eIF4A1 and G3BP1 respectively. C and D show z-values representing the standard deviation from the median across all samples for lung adenocarcinoma for normal tissue (n=111) and primary tumour (n=111) for eIF4A1 and G3BP1 respectively. Log_2 spectral count ratios generated by CPTAC were first normalised within each sample and subsequently across all samples.

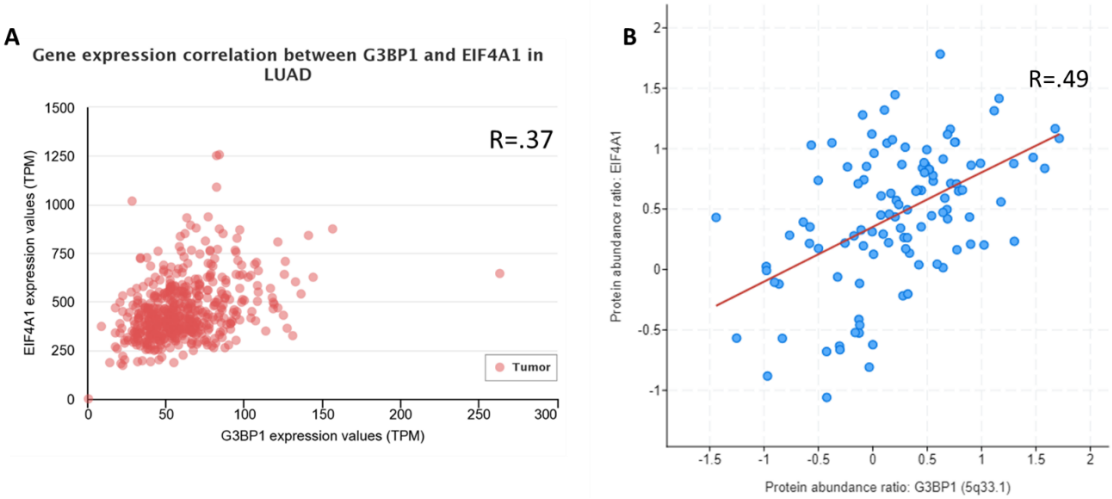


Figure 3.22: G3BP1 and eIF4A1 are co-expressed in lung adenocarcinoma.
A: Scatter plot of normalised TCGA RNAseq reads (transcripts per million) for G3BP1 and eIF4A1 with Pearson correlation coefficient (R). **B:** Scatter plot of normalised CPTAC protein abundance ratios for eIF4A1 and G3BP1 with Pearson correlation coefficient (R).

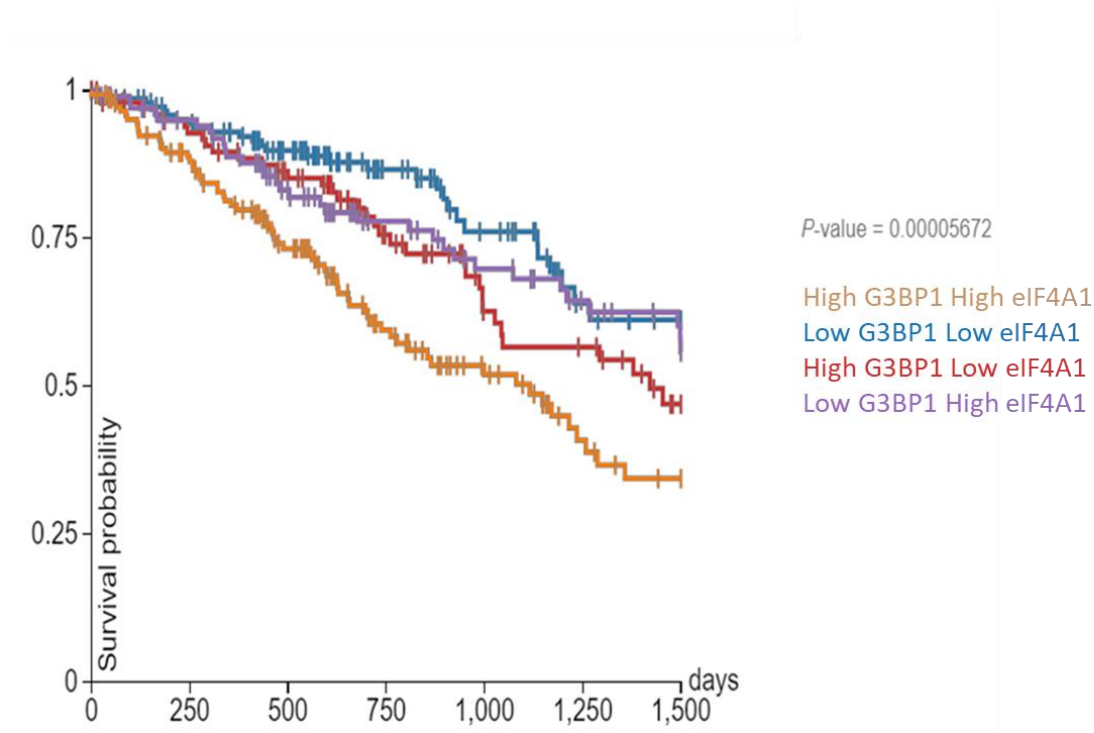


Figure 3.23 High expression of G3BP1 and eIF4A1 corresponds with the worst survival outcomes in lung adenocarcinoma:
Kaplan Meier survival curve of TCGA data stratified by median RNAseq expression values. P-value reported is generated by the log-rank test. Censored at 1500 days.

A selective TCP seq (Wagner et al., 2022) dataset generated in the Bushell lab during the course of this PhD project showed that eIF4A1 and G3BP1 may be assembling into complex co-translationally (figure 3.24). Selective TCP seq combines ribosome profiling with immunoprecipitation which allows for identification of proteins assembling into complex co-translationally. There is an increase in ribosome occupancy towards the end of the G3BP1 mRNA CDS in the eIFA1 immunoprecipitated sample that is not present in the control conditions. Control conditions are purified 80S fractions with no immunoprecipitation plus and minus hippuristanol. An eIF4A2 immunoprecipitation was used as a control to show that the interaction is eIF4A1 specific.

The first 20 amino acids of eIF4A1 are a predicted intrinsically disordered region (IDR) with a net acidic charge. G3BP1 harbours four IDRs, of which IDR 3 and IDR 4 are basic. Using HADDOCK 2.4 (van Zundert et al., 2016), a protein-protein interaction modelling system, and AlphaFold predicted structures of both proteins, a number of potential structures showing binding of eIF4A1 to G3BP1 were predicted *in silico*. An example of one of these structures showing the N-terminal IDR of eIF4A1 bound to the C-terminal IDR of G3BP1 is shown in figure 3.25.

Whilst *in silico* protein modelling has advanced rapidly over the past number of years, these findings still needed to be validated *in cellulo.* An eIF4A1 immunoprecipitation in A549 cells showed that G3BP1 was recovered and enriched versus the IGG control when analysed by western blot (figure 3.26 A). Vinculin and eIF4G were used as positive and negative controls respectively. As both eIF4A1 and G3BP1 are known to interact with the 40S ribosome, this result may not be as a consequence of direct binding. In order to elucidate whether these proteins can directly bind each other, an immunoprecipitation using in vitro reconstituted proteins was performed. Equimolar ratios of reconstituted eIF4A1 and 6xHis tagged G3BP1 were incubated for 15 minutes followed by a 30-minute incubation with protein G Dynabeads pre conjugated with a 6xHis antibody or rabbit IGG (figure 3.26 B). Western blot analysis of this experiment showed an enrichment in the recovery of both G3BP1 and eIF4A1, thus indicating that these two proteins can directly bind each other.

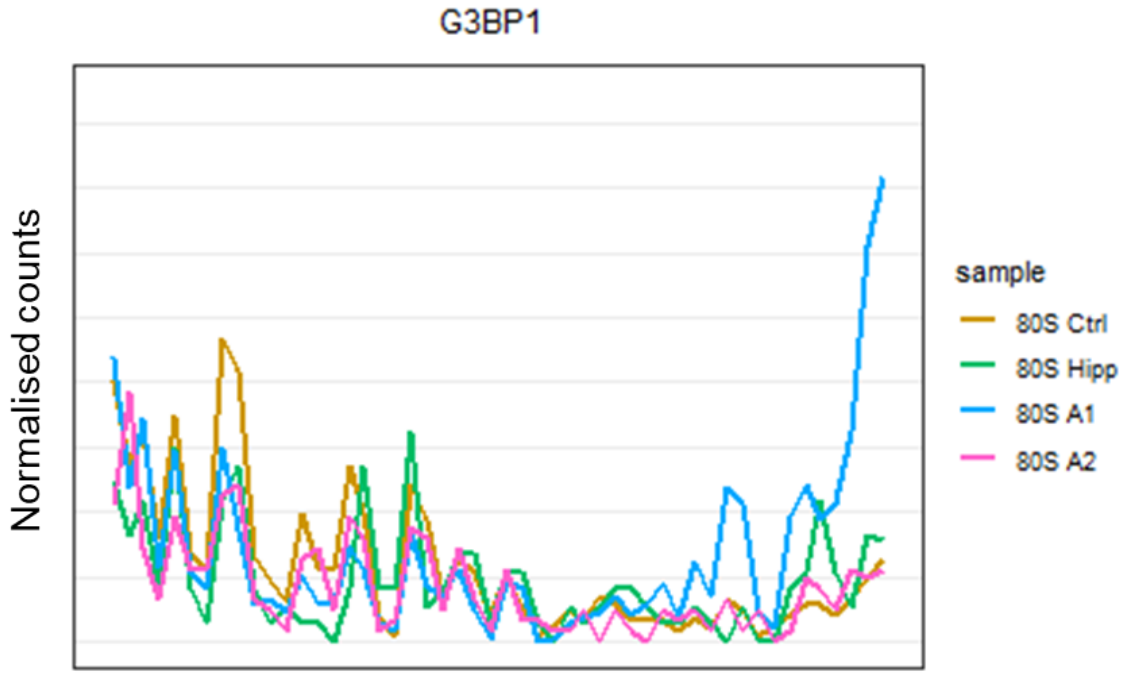


Figure 3.24 eIF4A1 and G3BP1 assemble co-translationally
Ribosome occupancy plot of G3BP1 CDS binned into 100 bins. showing increased reads of 80S fragments that had been additionally purified by eIF4A1 (80S A1) immunoprecipitation in comparison to 80S purified fragments with no immunoprecipitation (80S Ctrl), following treatment with hippuristanol (80S Hipp) or eIF4A2 immunoprecipitation (80S A2).

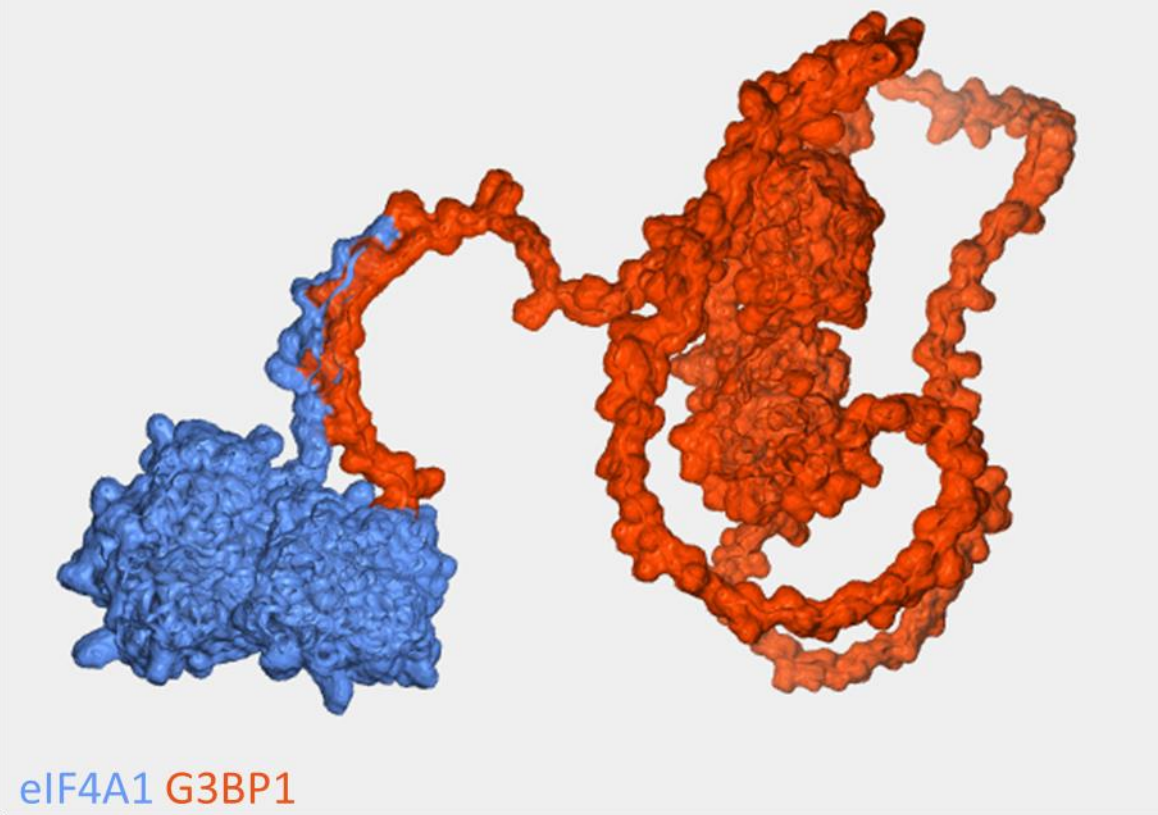


Figure 3.25 eIF4A1 and G3BP1 are predicted to interact *in silico*:
Predicted structure of the N-terminal intrinsically disordered region of eIF4A1 interacting with the C-terminal intrinsically disordered region of G3BP1 generated by HADDOCK.

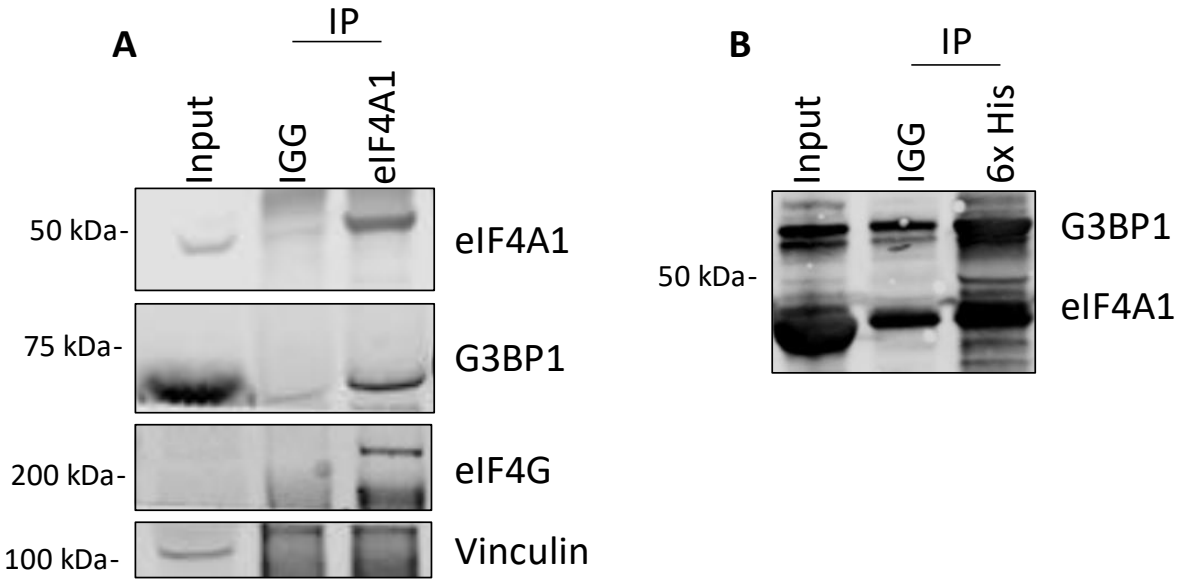


Figure 3.26 eIF4A1 and G3BP1 co-purify *in cellulo* and *in vitro*.
A: Western blot of eIF4A1 immunoprecipitation in A549 cells showing co-purification of G3BP1 and eIF4G. Vinculin is used as a negative control. Equal volumes of cell lysate and purified fractions loaded. Representative image of 3 biological replicates. **B:** Western blot of anti 6xHis immunoprecipitation of reconstituted eIF4A1 and 6xHis tagged G3BP1 showing enrichment over the IGG control. Equal volumes loaded. Representative image of 3 biological replicates.

As the commercially bought reconstituted G3BP1 was of poor quality, with a large proportion of degraded or truncated protein, the G3BP1 coding sequence was PCR amplified from the GFP-G3BP1 expression construct and cloned into a pET-SUMO vector for heterologous production in E. coli DE3 CodonPlus RP as N-terminal SUMO fusion proteins, as in (Schmidt et al., 2023, Wilczynska et al., 2019). Protein was then HisTrap purified followed by SUMO protease cleavage. Protein was then further purified using a ResourceQ anion exchange resin followed by sepherose size exclusion chromatography. Purified protein was run on SDS PAGE gels and stained with Coomassie (figure 3.27). Protein of superior quality was obtained and can now be used in future in vitro biochemical assays.

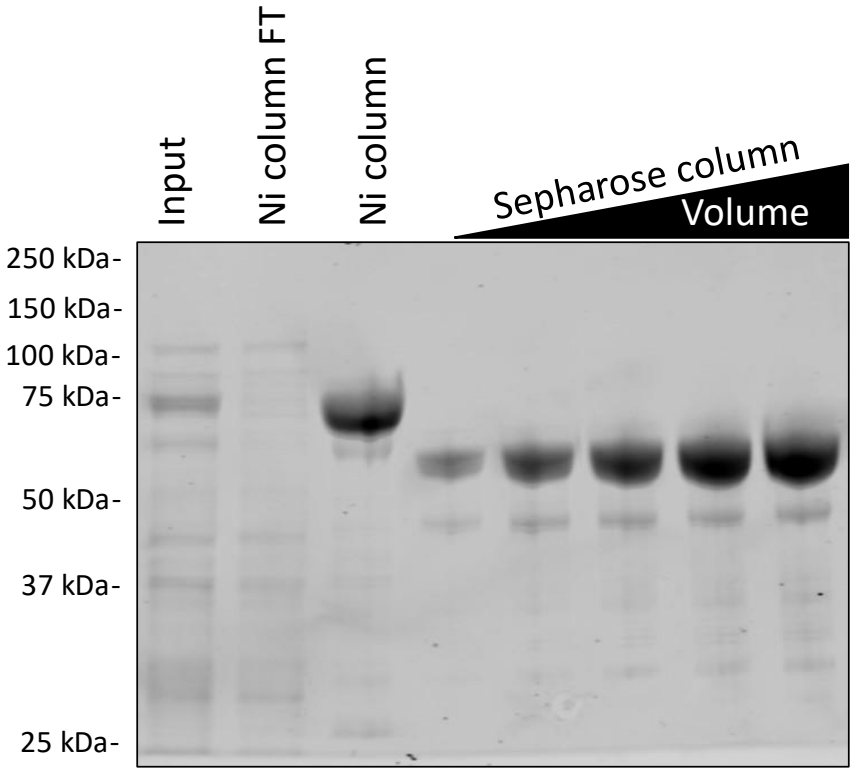


Figure 3.27: High quality reconstituted G3BP1 can be purified.
Coomassie stained SDS PAGE gel of G3BP1 purification from *E. coli.* Nickel column eluate (Ni column) shows enrichment of SUMO tagged G3BP1 versus input and flow through (Ni column FT). Increasing volume of the eluate from the Sepharose size exclusion chromatography was loaded showing considerably less truncated or degraded protein than in the commercially bought preparation.

3.7 Chapter discussion

Stress granules with distinct morphologies and upstream signalling cascades were identified in figures 3.1 and 3.2. In response to hypoxia stress, larger, poorly defined stress granules with no evident docking P-bodies formed. A reduction in mTOR activity led to hypophosphorylation of 4E-BP1 and inhibition of eIF4F complex assembly, thus resulting in stress granule formation in response to hypoxia (figures 3.3 and 3.4). Immunofluorescent imaging revealed co-localisation of eIF3B (figure 3.5). This result coupled with the signalling events observed suggests that the hypoxia induced stress granules are indeed cytoprotective and do not represent the pro-apoptotic type II stress granules identified in (Aulas et al., 2018).

Stress granule structure and dynamics is primarily regulated through their recruitment of helicases and ATPases (Jain et al., 2016). The DEAD box helicases/ATPases DDX6 and eIF4A1 were identified as differentially enriched in arsenite and hypoxia induced stress granules respectively (figures 3.8, 3.15-17). Both of these proteins have been heavily implicated in the regulation of biomolecular condensates. DDX6 is known to promote liquid-liquid phase separation *in vitro* (Hondele et al., 2019), and is required for P-body biogenesis in both yeast and human cell lines (Ayache et al., 2015, Andrei et al., 2005, Hondele et al., 2019). DDX6 is known to be a member of multiple protein complexes, including the mRNA decapping complex (Coller et al., 2001, Ayache et al., 2015). Two other members of this complex were also identified as enriched in arsenite induced stress granules, EDC4 and DCP1B. Very little is known mechanistically about the stress granule P-body docking interaction. It is theorised to be facilitated by RNA-RNA mediated interactions (Tauber et al., 2020). In response to arsenite stress, the enrichment of DDX6 with EDC4 and DCP1B suggests that the decapping complex may also be a regulator of stress granule P-body docking interactions through promoting RNA-RNA mediated interactions on the surface of these two closely related liquid-liquid phase separated biomolecular condensates.

Conversely, eIF4A1 has been shown to limit RNA condensation *in vitro* and overexpression prevents stress granule formation in the context of arsenite stress (Tauber et al., 2020). A limitation of the *in vitro* work presented in this study is that eIF4A1 was the only protein

present in the assays. In a cellular context, proteins rarely work as lone agents, their structure and function is often regulated by recruitment to multi-protein complexes. Figures 3.24-26 show that eIF4A1 and G3BP1 can directly bind, potentially through their opposingly charged IDRs. What is unknown is how this interaction will influence eIF4A1's propensity to limit RNA condensation. Nevertheless, in the context of hypoxia induced stress granules, eIF4A1 may still be limiting RNA condensation particularly the formation of stress granule cores and limiting stress granule P-body docking interactions. To this end, one way of interpreting the larger morphology observed (figure 3.1) is that that these granules are more liquid like and diffuse. The presence of the large ribosomal subunit in hypoxia induced stress granules (figures 3.15, 3.17, 3.18) suggests that these granules may be a more liquid like, dynamic and translationally active subcellular compartment than what has previously been described in response to arsenite stress (Mateju et al., 2020). DDX6 also contains a predicted N-terminal IDR. Unlike eIF4A1, this N-terminal region is basic. As G3BP1 contains both basic and acidic IDRs it may also be able to bind to DDX6 in a similar manner as proposed for eIF4A1.

Mining of publicly available cancer databases TCGA and CPTAC further emphasized that the binding of eIF4A1 and G3BP1 may not just be critical regulator of phase separation *in vitro* and in cell culture experiments but also of lung adenocarcinoma disease progression. Both G3BP1 and eIF4A1 are overexpressed in primary tumours comparatively to normal lung tissue (figure 3.21). This overexpression is coupled with a strong positive correlation between eIF4A1 and G3BP1 expression at the level of both RNA and protein (figure 3.22). Survival analysis revealed that having high expression of both eIF4A1 and G3BP1 corresponds with the worst patient outcomes in lung adenocarcinoma (figure 3.23). A potential limitation of many of the studies characterising stress granules is the use of high dose arsenite. For example, one of the original studies that concluded large ribosomal subunits are absent from stress granules used 1mM arsenite (Kedersha et al., 2002). In more recent work that showed active translation occurring within stress granules (Mateju et al., 2020) and selective portioning of G3BP1 bound pro-survival transcripts into polysomes, thus increasing their expression (Somasekharan et al., 2020), the concentration of arsenite used was 5 and 10 times lower respectively. These studies have undoubtedly propelled the field forward and changed the dogma of stress granules being exclusively composed of repressed mRNAs. It also proposes another question, of how representative extreme arsenite stress is

of disease. It is therefore possible that not only does hypoxia represent a more disease relevant stress but also a more disease relevant stress granule state.

In conclusion, the two primary aims of this chapter have been fulfilled. Robust methodologies for inducing stress granule formation have been developed. A methodology for isolation of stress granules was also successfully developed which enabled proteomic characterisation in response to different stresses. Stress granule proteomics further implicated the DEAD box helicases DDX6 and eIF4A1 in regulating liquid-liquid phase separation dynamics, with potential opposing roles in promoting and limiting RNA-RNA mediated interactions. The next chapter of thesis will use the methodologies developed in this chapter to characterise the mRNAs present in arsenite and hypoxia induced stressed granules.

4 Chapter 4: Isolation and RNAseq analysis of stress granule cores

4.1 Chapter Introduction

In addition to stress specific proteomes, stress granules have been shown to have stress specific mRNA localisation (Namkoong et al., 2018, Padrón et al., 2019). Therefore, to understand how stress granules are regulating translation in response to sodium arsenite treatment or hypoxia the stress specific mRNA localisation must be determined. In order to achieve this the stress granule core isolation methodology developed in the previous chapter will be used to generate libraries suitable for RNA sequencing.

Previous studies have identified characteristic sequence features of mRNAs. Stress granule mRNAs tend to be longer (Khong et al., 2017, Namkoong et al., 2018, Padrón et al., 2019). mRNA length has been proposed to contribute to stress granule assembly as longer RNAs provide a platform for multivalent protein-RNA and RNA-RNA interactions. Stress granule mRNAs in response to sodium arsenite have also been described as having lower GC content than stress granule depleted mRNAs (Khong et al., 2017). In addition to this, previous work has shown that stress granule mRNAs contain AU rich elements or motifs (Namkoong et al., 2018). Understanding how RNA sequence features regulate stress granule localisation is still not fully understood. In addition to mRNAs, non-coding RNAs were identified in stress granules (Khong et al., 2017). Non-coding RNAs contribute to biomolecular condensate assembly in paraspeckles (section 1.3) however their role in stress granules is still not known.

The aims of this chapter are:

1. Adapt the stress granule core isolation methodology to enable RNAseq analysis of stress granule mRNAs.

2. Characterise stress granule mRNAs and their sequence features in sodium arsenite and hypoxia induced stress granules.

4.2 Isolation of stress granule cores and generation of RNAseq libraries

In order to determine the mRNAs present in stress granules U20S cells transiently expressing GFP-G3BP1 were stressed using 200µM arsenite for 3 hours or hypoxia (0.1% O_2 for six hours). During the work of this thesis two studies using 200µM sodium arsenite were published to study translation as 500µM arsenite was shown to be inhibitory of the activation of the integrated stress response (Mateju et al., 2020, Somasekharan et al., 2020). Therefore the treatment regimen from the previous chapter was changed. Non-stressed cells expressing GFP-G3BP1 or a GFP empty vector control were used as a control. The stress granule isolation methodology developed in the previous chapter was then used to isolate stress granule cores by means of differential centrifugation followed by a GFP immunoprecipitation. Stress granule isolation was then assessed by a Coomassie stained SDS PAGE gel (figure 4.1).

There was an enrichment of total level of protein recovered after stress granule isolation in both the arsenite and hypoxia stressed samples compared to the non-stressed and GFP only empty vector controls. This enrichment in response to stress indicates that the stress granule isolation had worked (figure 4.1). 50 µl total RNA was collected after the first spin in the enrichment protocol and isolated using TRIzol according to manufacturer's protocol. After the stress granule isolation, GFP-trap beads were ProteinaseK treated at 55°C for 20 minutes with shaking at 1100 rpm. RNA was isolated using TRIzol LS. After the isopropanol precipitation step, both total and stress granule RNA samples were subjected to an additional overnight lithium chloride precipitation. Final pellets were resuspended in nuclease free water. RNA concentration was determined using a NanoDrop One spectrophotometer.

Isolated RNA was then ribosomal RNA depleted using a RiboCop v2 kit, using 1 µg of RNA as input for the total samples and the entirety of the remaining RNA from the stress granule isolations. RNAseq libraries were then generated using a CORALL library prep kit. As the quantity of input RNA was significantly lower in the samples from stress granule isolations compared to the total input RNA, an additional qPCR optimisation step was added to the

protocol to determine the appropriate number of PCR cycles to use for final library PCR amplification (figure 4.2).

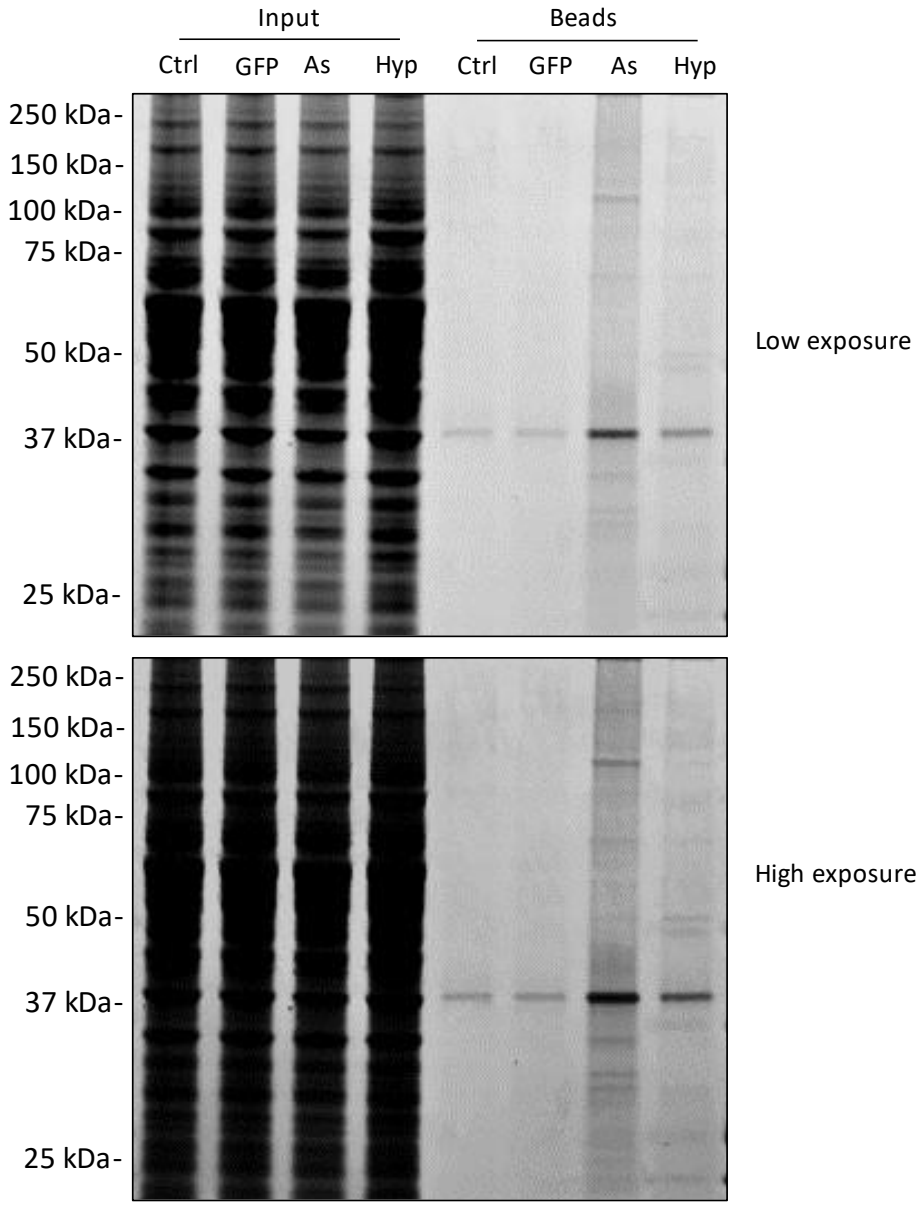


Figure 4.1 Coomassie validation of stress granule isolations for RNAseq:
Coomassie stained SDS PAGE gel of proteins recovered after stress granule isolation methodology in non-stressed control (Ctrl), GFP empty vector control (GFP) and arsenite (As) and hypoxia (Hyp) stressed U2OS cells. Induction of stress with either arsenite or hypoxia results in an enrichment of total protein recovered compared to the controls Equal volumes of cell lysate and purified fractions loaded. Representative image of 3 biological replicates.

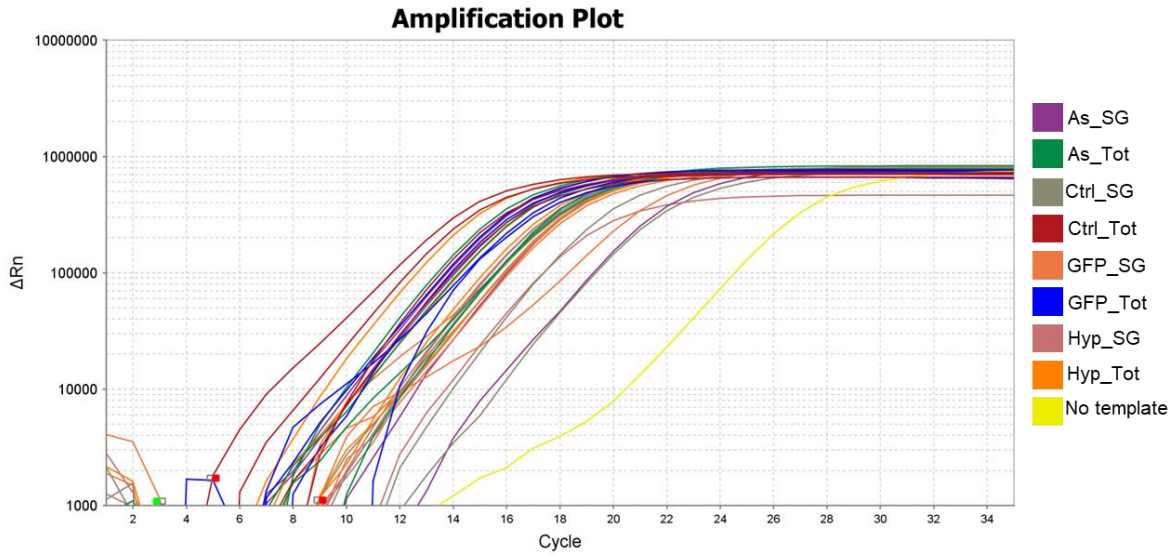


Figure 4.2 qPCR amplification plot of RNAseq libraries.
Amplification plot of all stress granule isolated (SG) and total (Tot) RNAseq libraries for non-stressed control (Ctrl), GFP empty vector control (GFP) arsenite (As) and hypoxia (Hyp) stressed U20S cells. All libraries amplified with a lower cycle number than the no template control.

All libraries amplified with lower cycle numbers than the no template control. Total RNA samples were PCR amplified using 13 PCR cycles. Stress granule RNA libraries were amplified using 15 PCR cycles due to their lower average CT values. PCR amplicon size was then analysed using a TapeStation D1000 high sensitivity screen tape. A representative electropherogram with the samples from replicate 4 is shown in figure 4.3. Libraries across all samples had a mean fragment length of ~300 bp which is the expected fragment size from the CORALL kit.

Library concentrations were determined using a Qubit 2.0 fluorometer with high sensitivity dsDNA reagents. Libraries were then pooled to a final concentration 2nM. Libraries were then sequenced on an Illumina NextSeq 500 using 75 cycles by the Beatson Institute Molecular Tecnologies Services.

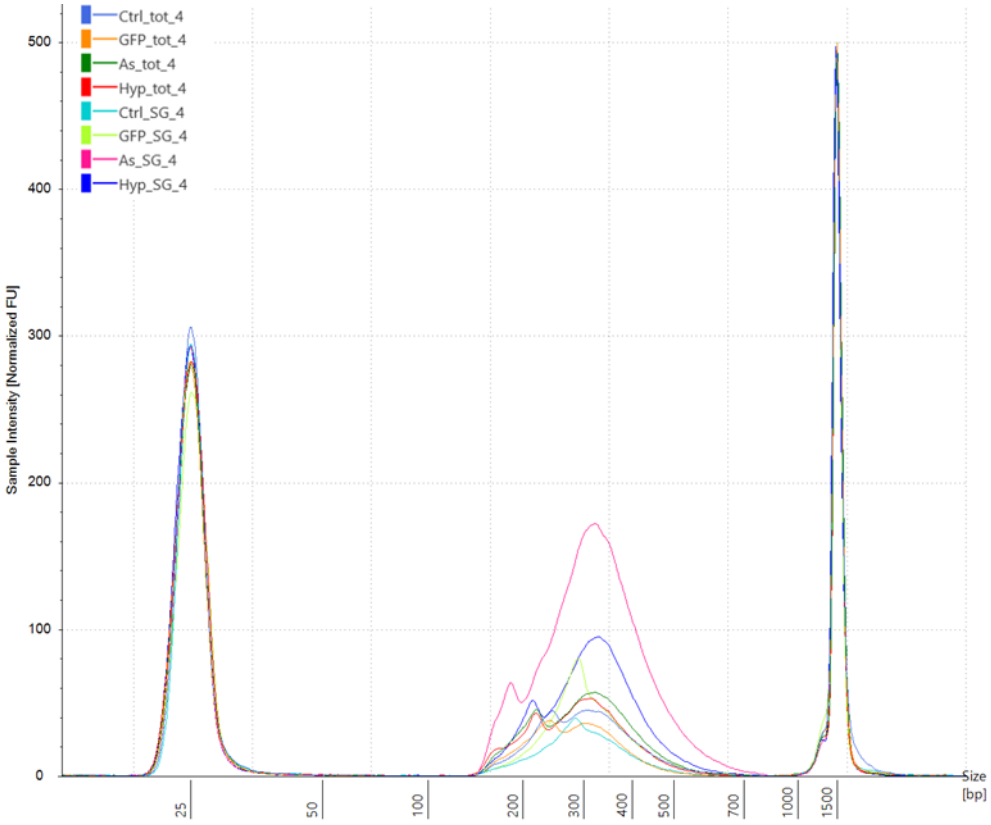


Figure 4.3 Stress granule RNAseq libraries are of expected size:
Representative electropherogram using samples generated in replicate 4 of stress granule isolations from TapeStation D1000. All libraries generated are of the expected size of ~300 bp. Peaks at 25- and 1500 bp represent the markers from the TapeStation D1000 assay.

4.3 Arsenite and hypoxia induced stress granules contain differentially enriched mRNAs.

Raw sequencing data was converted from bcl format to fastq format and demultiplexed using the Illumina sofatware package bcl2fastq. Sequencing data quality was assessed using the fastqc software package developed by Babraham Bioinfomatics. Adaptor sequences were removed using cutadapt (Martin, 2011). UMI-tools was used for sample deduplication to remove overrepresented PCR artefacts (Smith et al., 2017). UMIs were first extracted followed by transcriptome alignment to the GENCODE v38 human transcriptome, which had been pre-filtered to contain only protein coding transcripts, using bowtie2 (Langmead and Salzberg, 2012). SAM files were converted to BAM format and indexed using SAMtools (Li et al., 2009). Files were then deduplicated using UMI-tools. Calculation of isoform and gene level expression was conducted using RSEM (Li and Dewey, 2011).

RSEM output was loaded into R using tximport (Soneson et al., 2015) followed by differential expression analysis with DESeq2, using an experimental design accounting for condition and batch correction by replicate (Love et al., 2014). Low abundance genes were removed by filtering genes that had less than an average of 10 counts across all samples. Log fold change shrinkage using the apeglm method (Zhu et al., 2018) was used to reduce noise within the data. PCA plots were generated for total- and stress granule RNA individually to visualise the degree of clustering between samples (figure 4.4).

There was a high level of clustering observed in the total RNA samples. GFP empty vector control and the non-stressed GFP-G3BP1 expressing control showed a high degree of clustering. This indicates that the changes observed in the total RNAseq samples are a consequence of stress induction and not an artefact of transfection. In the hypoxia total RNAseq samples, a high degree of clustering was observed for all samples. Arsenite stressed samples showed a high degree of clustering for replicates 1-3. Replicate four clusters with the hypoxia stressed samples indicating that there may have been an error during library preparation. Removal of this sample did not influence the number of differentially expressed genes identified in the arsenite total RNA versus stress granule RNA comparison, therefore it was not removed from the analysis. This library preparation should be repeated to ensure

no error occurred during the processing (figure 4.4 A). For the stress granule RNA libraries, the two control conditions showed high levels of clustering, similarly to what was seen with the total RNA samples. There is a high degree of clustering in both the arsenite and hypoxia samples respectively (figure 4.4 B).

Each stress granule condition was compared to its own respective total RNA sample in a DESeq2 differential expression analysis. All genes with an adjusted p-value <0.1 were deemed significant. 4031 genes were identified as significantly enriched and 3925 genes were identified as significantly depleted in the arsenite treated stress granule isolated fraction versus its total RNA. 3799 genes were identified as significantly enriched and 3891 genes were identified as significantly depleted in the hypoxia treated stress granule isolated fraction versus its total RNA. 4149 genes were identified as significantly enriched and 3906 genes were identified as significantly depleted in the non-stressed control stress granule isolated fraction versus its total RNA. 3492 genes were identified as significantly enriched and 3542 as significantly depleted in the GFP empty vector control stress granule isolated fraction versus its total RNA. These findings are summarised as volcano plots in figure 4.5.

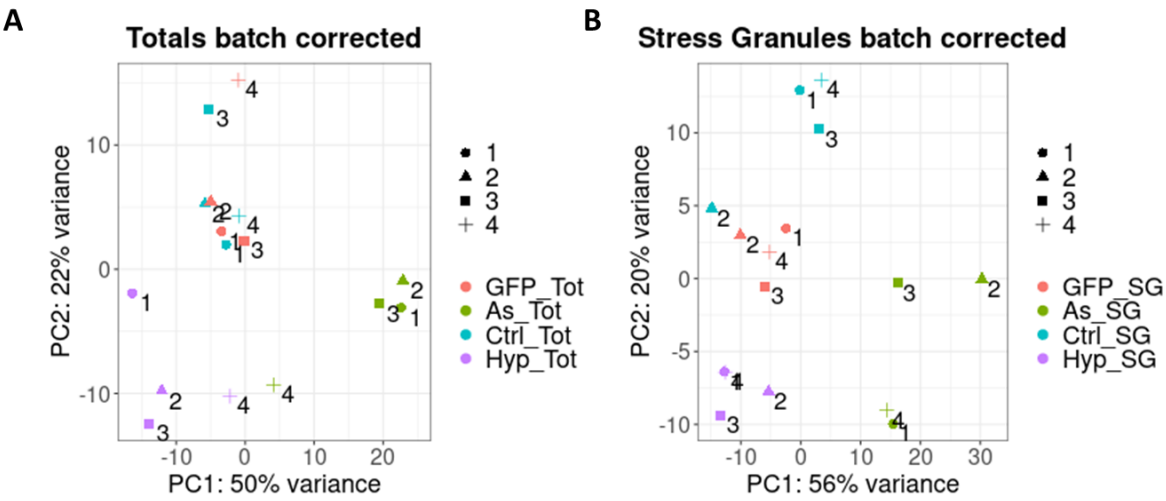


Figure 4.4 PCA analysis of RNAseq data shows high degree of clustering for both total- and stress granule isolated RNA samples.
PCA plot of total RNA (A) and stress granule isolated RNA (B) High degree of clustering is observed for all samples in non-stressed control (Ctrl), GFP empty vector control (GFP) and hypoxia stressed (Hyp) samples. Arsenite stressed (As) total replicate 4 clusters with hypoxia total RNA.

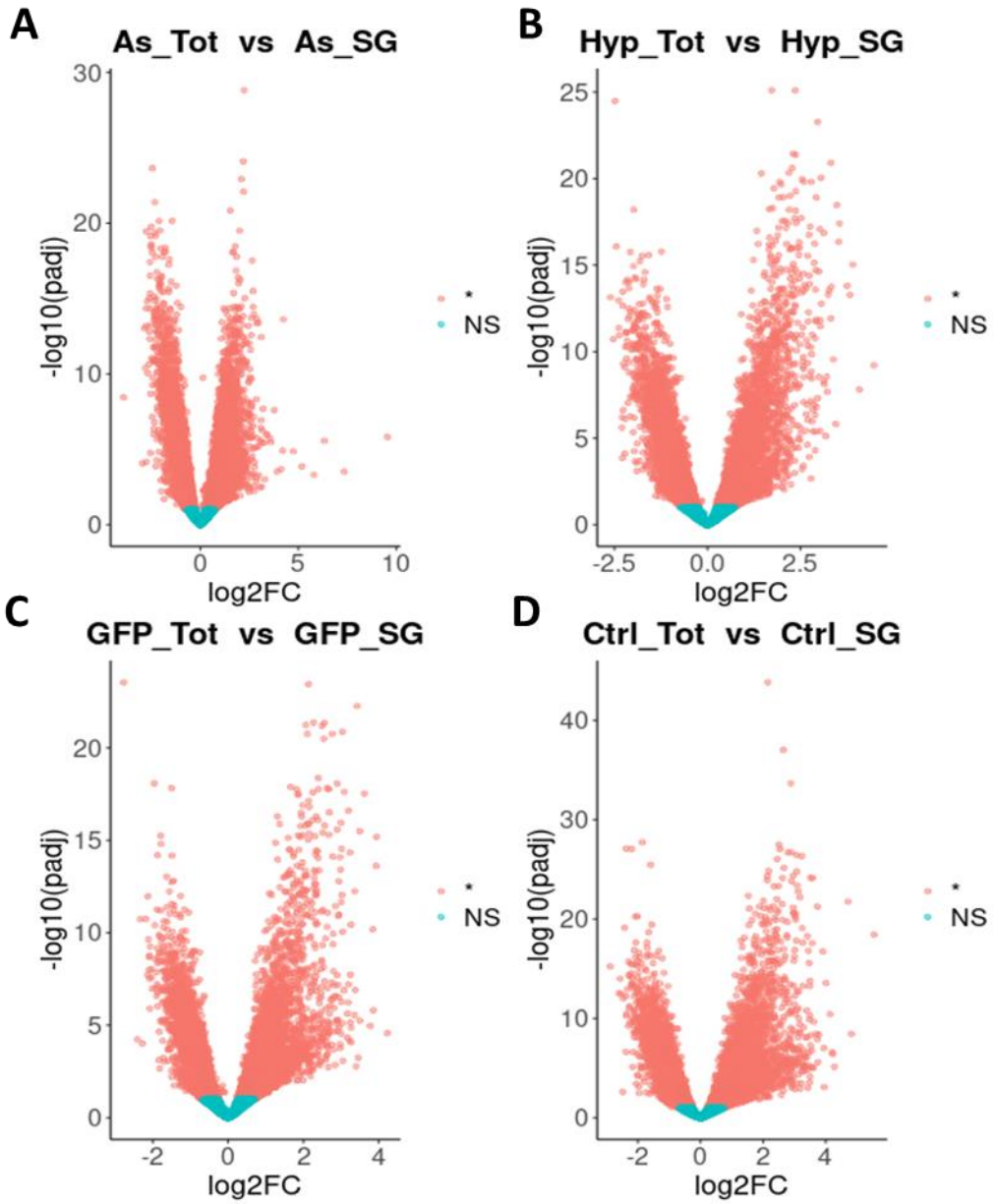


Figure 4.5 Volcano plots of differentially enriched genes for stress granule isolated RNA versus total RNA.
A: Arsenite stressed **B:** Hypoxia stressed **C:** GFP empty vector control **D** non stressed control. Significance cut-off of an adjusted p-value of 0.1 was used. There are a large number of differentially enriched and depleted genes in all conditions, including the controls.

A higher number of differentially enriched genes were present in the non-stressed and GFP only empty vector than expected. In order to determine the overlap of these gene sets they were plotted as Venn diagrams (figure 4.6). 2229 genes were identified as significantly enriched in the stress granule isolated fraction in all conditions (figure 4.6 A). 2242 were identified as significantly depleted in the stress granule isolated fraction in all conditions (figure 4.6 B). The high level of overlap suggests a large proportion of the mRNAs identified as enriched or depleted in all conditions is owing to technique bias as opposed to stress induction, similar to the high correlation of proteins identified in the previous chapter. This dataset was therefore stratified to remove all mRNAs that were identified as enriched or depleted in either of the control conditions. Stress granule enriched or depleted mRNAs were defined as mRNAs that were identified as enriched or depleted in either the arsenite or hypoxia treated samples but not in either of the controls. This resulted in 920 mRNAs being assigned as enriched in the arsenite stress granule fraction and 451 in hypoxia with an overlap of 162 (figure 4.6 A). 920 mRNAs were identified as depleted in the arsenite stress granule fraction and 509 in hypoxia with an overlap of 229 (figure 4.6 B). All subsequent downstream analyses used these stratified groups of mRNAs. The top ten enriched and depleted stress granule mRNAs for sodium arsenite and hypoxia are summarised in table 1 and table 2 respectively. Top ten enriched mRNAs for sodium arsenite had a range of a $\log_2$ fold change of 9.534 to 3.91. Tope depleted mRNAs had a range of a $\log_2$ fold change of -3.915 to -2.67. In hypoxia the top ten enriched mRNAs had a range of a $\log_2$ fold change of 4.477 to 3.386. The top depleted mRNAs had a range of a $\log_2$ fold change of -2.611 to -2.29. Whilst relative quantitation from RNA sequencing a high degree of enrichment or depletion from stress granules in both conditions, smFISH would allow for an absolute quantitation of the proportion of an mRNAs in a stress granule versus cytosol.

gene_sym	padj	log2FoldChange
BICDL2	1.52E-06	9.535502873
NFE2L3	3.08E-04	7.341216552
RADIL	2.72E-06	6.323961645
HRNR	4.92E-04	5.793119207
C22orf23	1.36E-04	5.179835324
PDZK1	1.36E-05	4.743205281
COL7A1	2.40E-14	4.236359275
TMEM184A	1.22E-05	4.206375311
GLUD1	2.08E-04	4.124088965
IL15	3.05E-04	3.910003462
CLEC2B	3.56E-09	-3.915945478
STC1	8.49E-05	-2.959021251
NDUFA4	9.55E-14	-2.854578231
RIDA	4.15E-11	-2.816164737
B2M	6.49E-14	-2.763673008
C1orf53	6.82E-05	-2.75588877
RPL39	4.27E-15	-2.755783176
SMIM30	3.52E-20	-2.752027074
GLRX	1.51E-15	-2.717429781
MORN2	8.75E-11	-2.672100303

Table 1: Table of top ten stress granule enriched and depleted mRNAs in response to sodium arsenite.
Gene_sym denotes gene symbol, padj denotes adjusted p-value, log2FoldChange denotes $\log_2$ fold change of SG enriched fraction vs total RNA. Positive values represent stress granule enriched mRNAs, negative values represent stress granule depleted mRNAs

gene_sym	padj	log2FoldChange
HDAC10	6.26E-10	4.477686767
MAMDC4	1.55E-08	4.084846604
PLEC	9.34E-16	3.908325373
CRTC1	5.29E-14	3.831133375
PRR12	1.55E-14	3.758021507
SYNGAP1	3.98E-18	3.548787362
CGN	4.42E-17	3.52879917
LENG8	3.45E-19	3.478621928
SH3D21	1.50E-06	3.456095141
COL7A1	2.83E-10	3.386733059
RPL39	7.28E-14	-2.611515551
NDUFA4	1.88E-11	-2.536361575
TPT1	3.27E-25	-2.481167908
RPL12	8.36E-17	-2.451593817
RPL22L1	1.10E-13	-2.389926998
PRDX4	4.25E-12	-2.382410173
PFDN4	1.16E-11	-2.37591223
MTRNR2L11	6.29E-07	-2.298663181
RP4-541C22.8	2.45E-06	-2.298605817
CAPSL	1.41E-04	-2.290113975

Table 2 : Table of top ten stress granule enriched and depleted mRNAs in response to hypoxia.
Gene_sym denotes gene symbol, padj denotes adjusted p-value, log2FoldChange denotes $\log_2$ fold change of SG enriched fraction vs total RNA. Positive values represent stress granule enriched mRNAs, negative values represent stress granule depleted mRNAs

Enriched and depleted mRNAs from both stress conditions were then compared to the previously published stress granule transcriptome, generated in (Khong et al., 2017), to assess overlap of datasets (figure 4.7). There are low levels of overlap between the previously published stress granule transcriptome and the data generated for this thesis. Only 57 mRNAs were shared between all three of the stress granule enriched datasets. Hypoxia shared a further 50 mRNAs and arsenite shared a further 150 mRNAs with the published stress granule transcriptome (figure 4.7 A). This represents an overlap of 15.8% of the 1625 mRNAs identified as enriched from (Khong et al., 2017). 67 mRNAs were shared amongst all three stress granule depleted mRNA datasets. Arsenite shared a further 110 and hypoxia shared a further 39. This represents an overlap of 12.1% of the 1779 stress granule depleted mRNAs identified in (Khong et al., 2017).

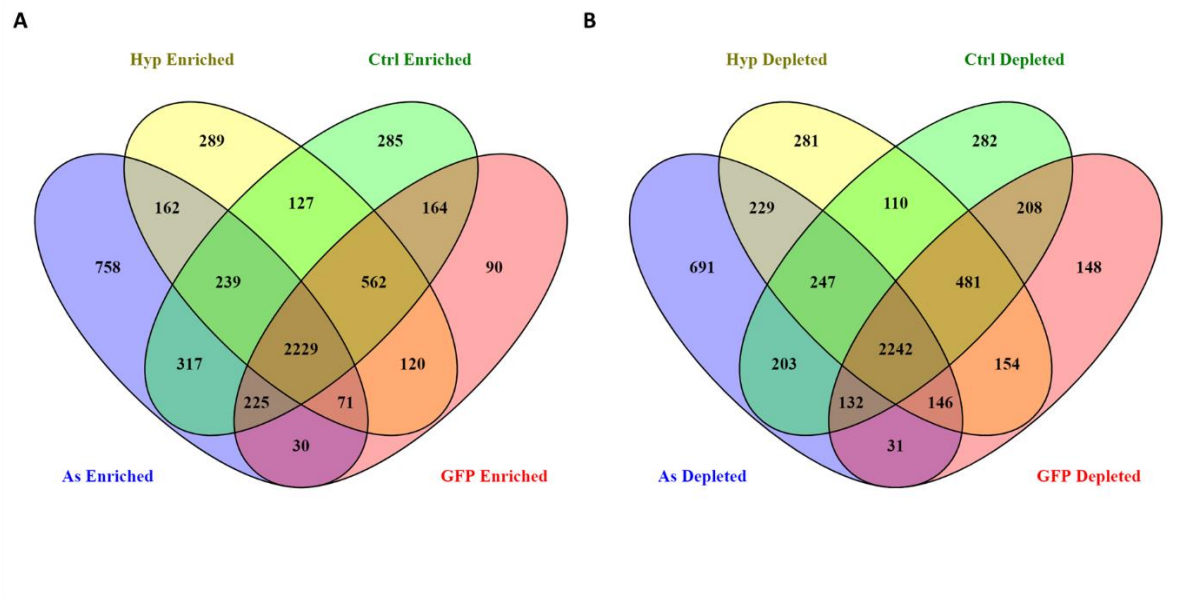


Figure 4.6 High number of differentially enriched and depleted genes in stress granule isolations are common to all conditions.
Venn diagrams showing overlap of genes identifed as differentially enriched **(A)** or depleted **(B)** in arsenite stressed (As), hypoxia stressed (Hyp), non-stressed control (Ctrl) or GFP empty vector control conditions. All genes that were identified in control conditions were removed from subsequent downstream analyses.

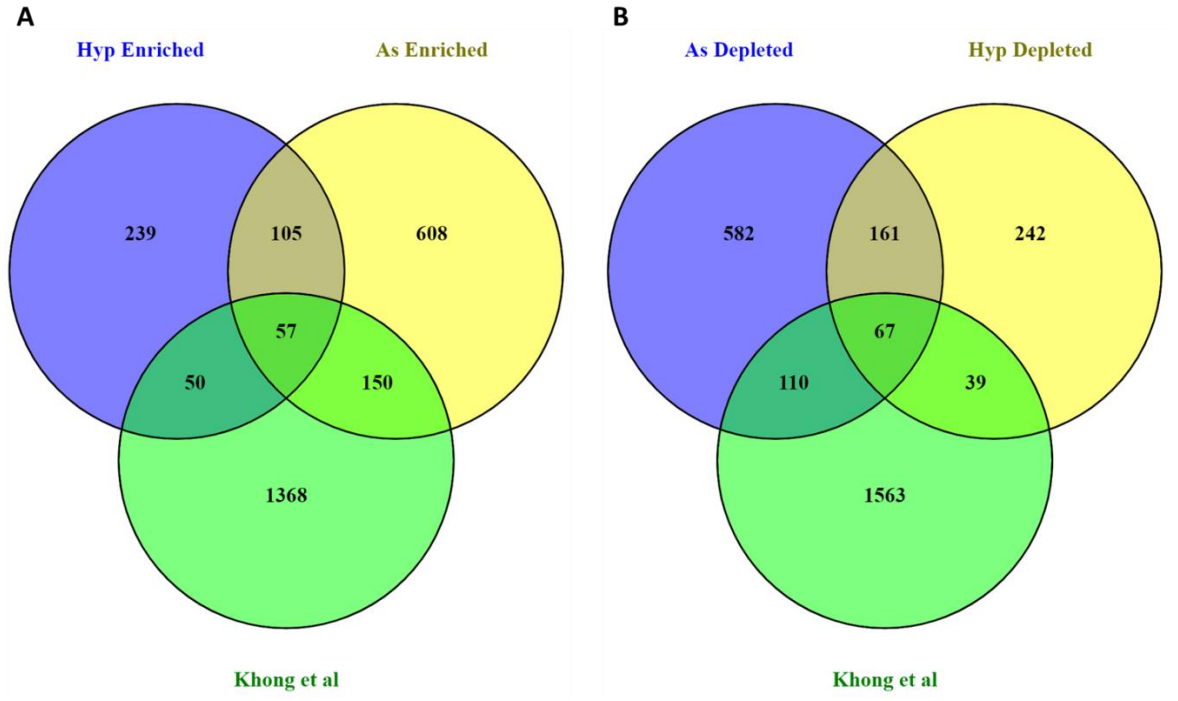


Figure 4.7 Hypoxia and arsenite induced stress granule enriched and depleted genes show low level of overlap with published stress granule transcriptome.
Venn diagrams showing overlap of genes identified as differentially enriched **(A)** or depleted **(B**) in arsenite stressed (As), hypoxia stressed (Hyp) and the published stress granule transcriptome (Khong et al).

As there was a large number of mRNAs identified as enriched or depleted in all conditions, including the controls, (figure 4.6) these groups were also compared to the mRNAs identified in (Khong et al., 2017) (figure 4.8). A much greater overlap is present between these datasets than those analysed in figure 4.7. In the stress granule enriched mRNAs there is an overlap of 626, representing 38.5% of the those originally identified in (Khong et al., 2017). In the stress granule depleted datasets there is an overlap of 694 mRNAs, representing 39% of the mRNAs identified in (Khong et al., 2017).

In order to determine if the overlaps observed in these gene sets was greater than what would be expected by chance, the GeneOverlap R package was used to calculate p-values from a Fisher's exact test (Fisher, 1992). Size of the background genome was set to 13390 as this was the sum total number of genes used for DESeq2 analysis and in the results table from (Khong et al., 2017). The results are represented as a bar plot of $-\log_{10}$(p-value) in figure 4.9. The mRNAs identified as enriched (p-value = 4e-124) or depleted (p-value = 5e-142) in all experimental conditions, including the two negative controls, had the most statistically significant overlap with those identified in (Khong et al., 2017). mRNAs identified as enriched or depleted in arsenite (enriched p-value = 4e-22, depleted p-value = 6e-9) and hypoxia (enriched p-value = 2e-13, depleted p-value = 2e-7) induced stress granules did however also have statistically significant overlap with those identified in (Khong et al., 2017) as reported for heat shock stress granule mRNAs identified in (Padrón et al., 2019).

These results indicate that a large proportion of mRNAs originally identified as either stress granule enriched or depleted may be due to technique artefacts as opposed to bona fide stress granule association. Additionally, these results show the absolute necessity for the inclusion of adequate controls in high-throughput methodologies to account for technique biases.

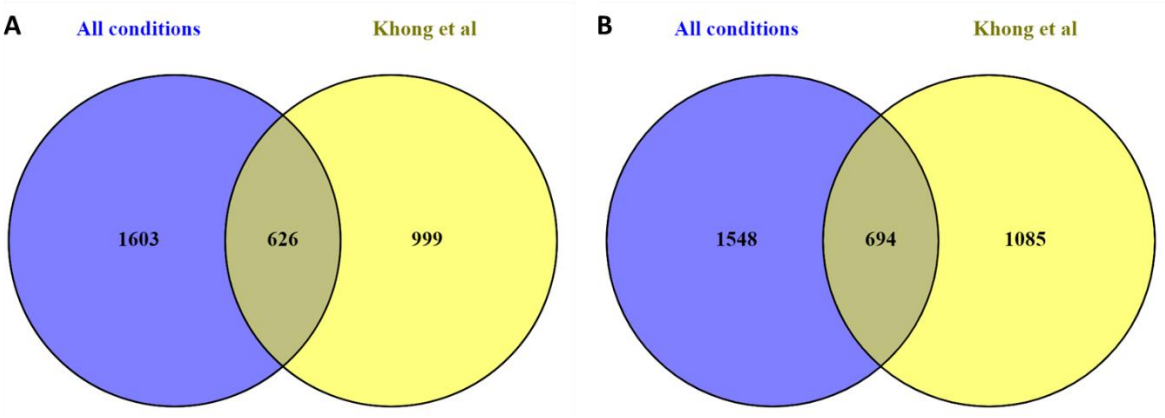


Figure 4.8 Genes identified as enriched or depleted in all conditions show a high level of overlap with published stress granule transcriptome.
Venn diagrams showing overlap of genes identified as enriched **(A)** or depleted **(B)** in all experimental conditions with the published stress granule transcriptome.

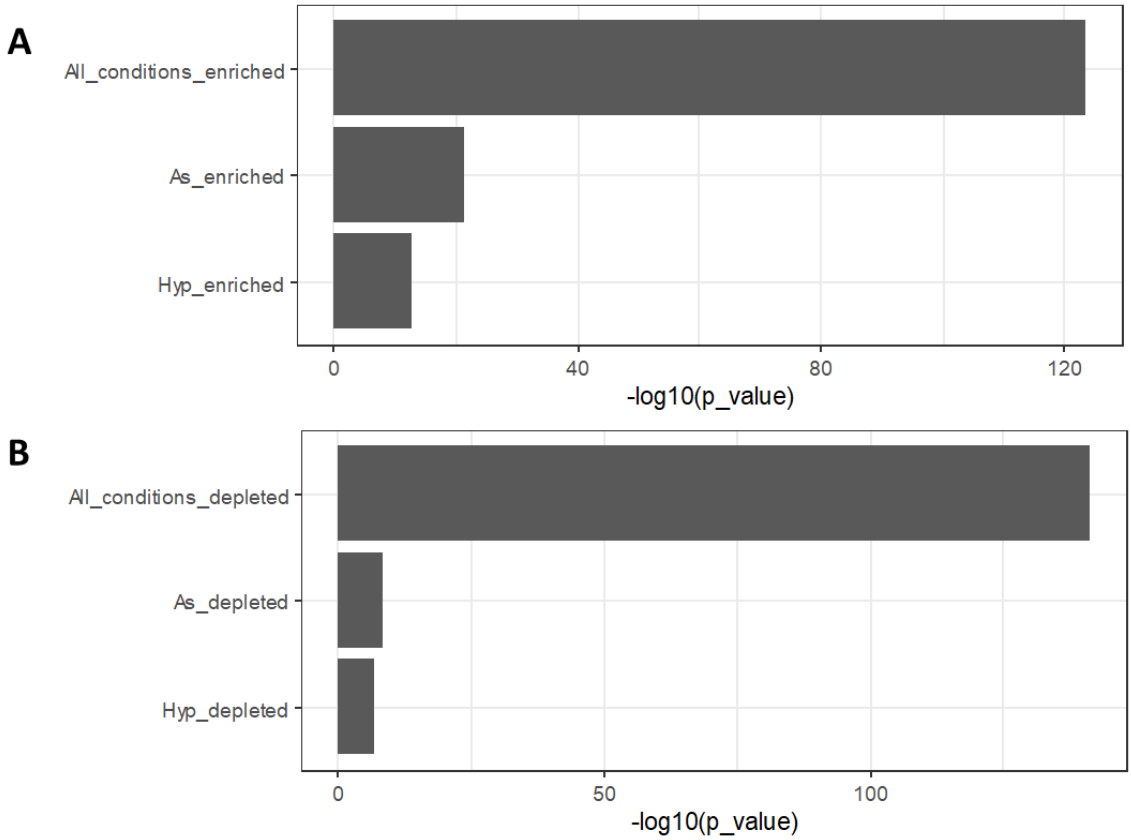


Figure 4.9: Genes identified as enriched or depleted in all conditions have the most statistically significant overlap with the published stress granule transcriptome.
Bar plots of $-\log_{10}$(p-value) from Fishers exact test, comparing overlap with published stress granule transcriptome for genes identified as enriched **(A)** or depleted **(B)** in arsenite induced stress granules (As), hypoxia induced stress granules (Hyp) or in all experimental conditions.

In order to elucidate the pathways that are being regulated by inclusion or exclusion of mRNAs in stress granules, an enrichment analysis using the fgsea R package was utilised (Sergushichev, 2016). Pathway enrichment was calculated using the Gene Set Enrichment Analysis (GSEA) hallmark gene sets as a reference. An adjusted p-value <0.05 was used as the significance threshold (figure 4.10). In both the arsenite and hypoxia conditions there was an intriguing convergence on pathways synonymous with cancer. These included, but are not limited to, KRAS signalling down (Uprety and Adjei, 2020), apical junction (González-Mariscal et al., 2020), and mitotic spindle (Stumpff et al., 2014) pathways being identified as enriched pathways in both the arsenite and hypoxia stress granule enriched mRNAs. This convergence was also evident in the stress granule depleted mRNAs in both conditions. Myc targets v1 (Dang, 2012), unfolded protein response (Madden et al., 2019) and DNA repair pathways (Helleday et al., 2008) were identified as enriched pathways in the stress granule depleted mRNAs in both arsenite and hypoxia. The KRAS signalling up pathway was also identified as enriched in the stress granule depleted mRNAs in both conditions. Stress granules have previously been implicated as a mechanism of chemotherapeutic resistance in KRAS mutant cell lines (Grabocka and Bar-Sagi, 2016). The inclusion and exclusion of KRAS signalling associated mRNAs further implicates stress granules as a potential critical, yet underappreciated, regulator of KRAS mediated oncogenesis and disease progression.

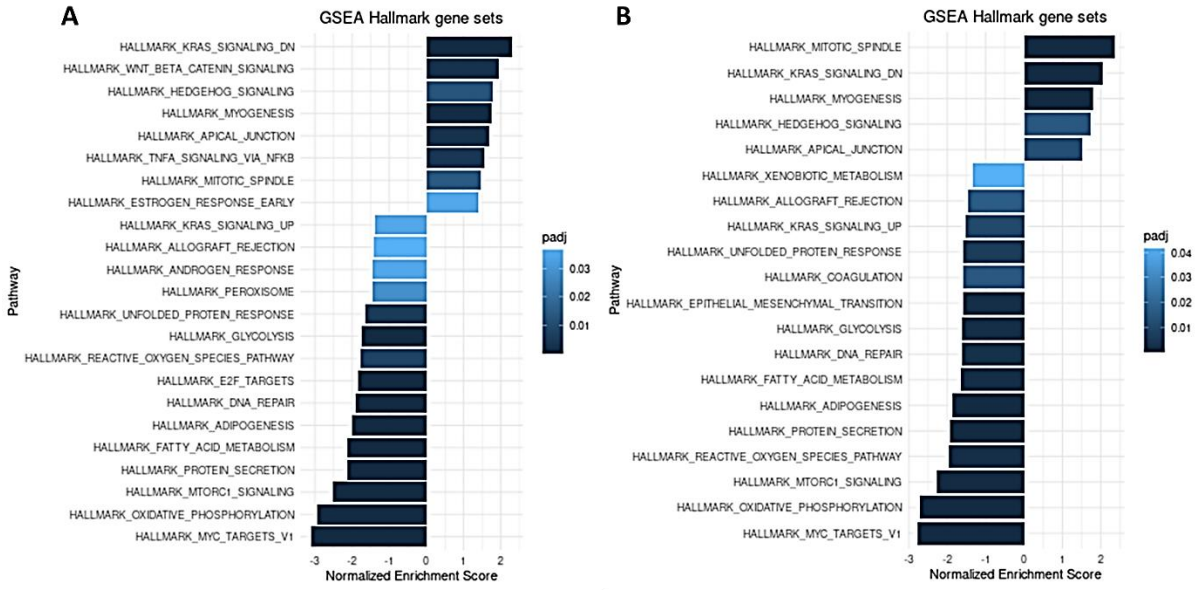


Figure 4.10 Pathway enrichment analysis of stress granule enriched and depleted genes in response to arsenite or hypoxia stress shows convergence on cancer associated pathways. Bar plots of normalised enrichment score from fgsea analysis for arsenite **(A)** and hypoxia **(B)** stress granule enriched and depleted mRNAs. Positive normalised enrichment score represents stress granule enriched mRNAs. Negative normalised enrichment score represents stress granule depleted mRNAs.

4.4 Sequence feature properties of stress granule mRNAs

mRNA length has previously been identified as a determining factor of stress granule localisation. Stress granule enriched mRNAs were identified as longer than depleted mRNAs by three different stress granule isolation techniques, in both mouse and human cell lines (Namkoong et al., 2018, Khong et al., 2018, Padrón et al., 2019). In order to determine whether the mRNAs identified in arsenite and hypoxia induced stress granules were indeed longer, 5'UTR, CDS and 3'UTR lengths were log transformed to reduce any potential outlier effects. Mean length was then compared in the enriched, depleted and non-significantly enriched or depleted subsets of mRNAs. Arsenite stress granule enriched mRNAs on average contain longer 5- and 3'UTRs in addition to longer CDSs when compared to both stress granule depleted or non-significantly enriched or depleted subsets (figure 4.11). mRNAs depleted from arsenite stress granules contained shorter 5- and 3'UTRs compared to the non-significantly enriched or depleted subset, but shorter length was not conserved for the CDS. Hypoxia induced stress granule mRNAs displayed similar length characteristics (figure 4.12). Hypoxia induced stress granule mRNAs also on average have longer 5'UTRs, CDSs and 3'UTRs compared to the depleted subset or the non-significantly enriched or depleted subset. Stress granule depleted mRNAs in hypoxia were significantly shorter across their 5'UTRs, CDSs and 3'UTR when compared to the non-significantly enriched or depleted subset.

In addition to sequence length, mRNA nucleotide sequence compositional biases have also been identified as features of stress granule mRNAs. Arsenite induced stress granule mRNAs were characterised as having lower GC content in (Khong et al., 2018). P-body mRNAs have also previously been shown to have an AU rich sequence bias (Courel et al., 2019). The arsenite induced stress granule mRNAs identified in this thesis, on average, had a higher GC content across their 5'UTR, CDS and 3'UTR in comparison to the depleted subset and the non-significantly enriched or depleted subset (figure 4.13). Arsenite stress granule depleted mRNAs had lower GC content across their 5'UTR, CDS and 3'UTR compared to the non-significantly enriched or depleted subset. Hypoxia induced stress granule mRNAs did not have the same GC bias (figure 4.14). This subset of mRNAs had a lower GC content in their CDS and 3'UTR compared to the depleted and non-significantly enriched or depleted subsets, with no statistically significant differences in their 5'UTR composition. Depleted

mRNAS had lower GC content in their 5’UTR, with higher GC content in their CDS and 3’UTR when compared to the non-significantly enriched or depleted subset.

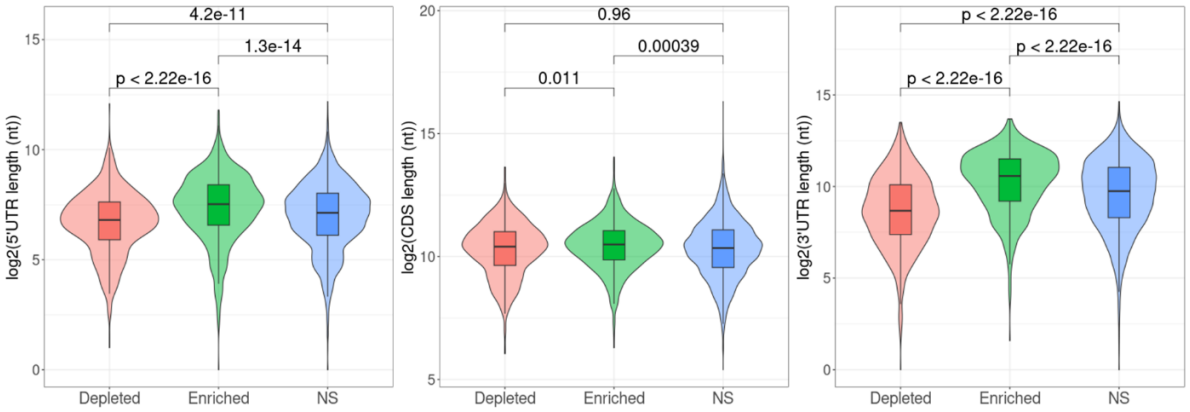


Figure 4.11 Arsenite stress granule enriched mRNAs contain longer 5'UTRs, CDSs and 3'UTRs.
Violin plots showing distributions with boxplots represnting minimum, maximum, interquartile range and median of log_2 transformed lenghths of 5'UTR, CDS and 3'UTR. Reported p-values are from a Wilcoxon rank-sum test. Groups represent mRNAs significantly depleted, enriched or non significantly enriched or depleted (NS) in arsenite stress granule isolated fraction

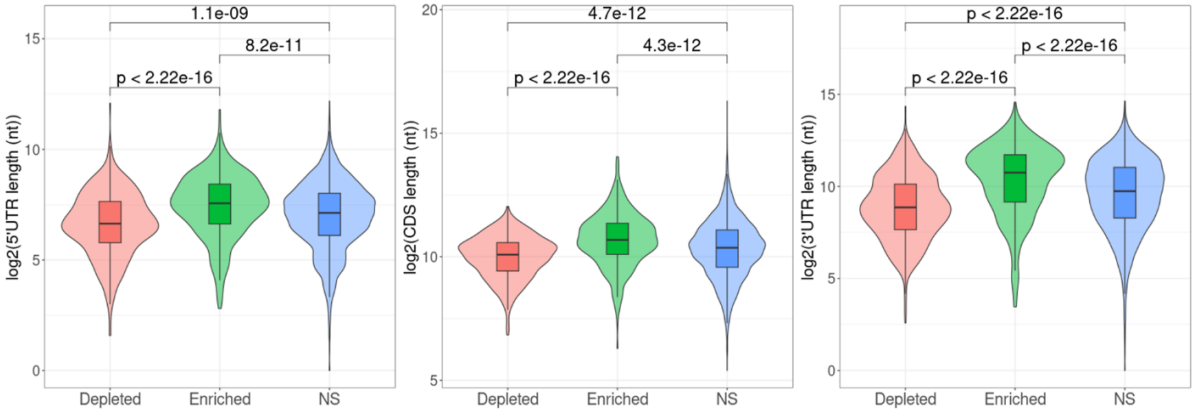


Figure 4.12 Hypoxia stress granule enriched mRNAs contain longer 5'UTRs, CDSs and 3'UTRs.
Violin plots showing distributions with boxplots represnting minimum, maximum, interquartile range and median of log_2 transformed lenghths of 5'UTR, CDS and 3'UTR. Reported p-values are from a Wilcoxon rank-sum test. Groups represent mRNAs significantly depleted, enriched or non significantly enriched or depleted (NS) in hypoxia stress granule isolated fraction

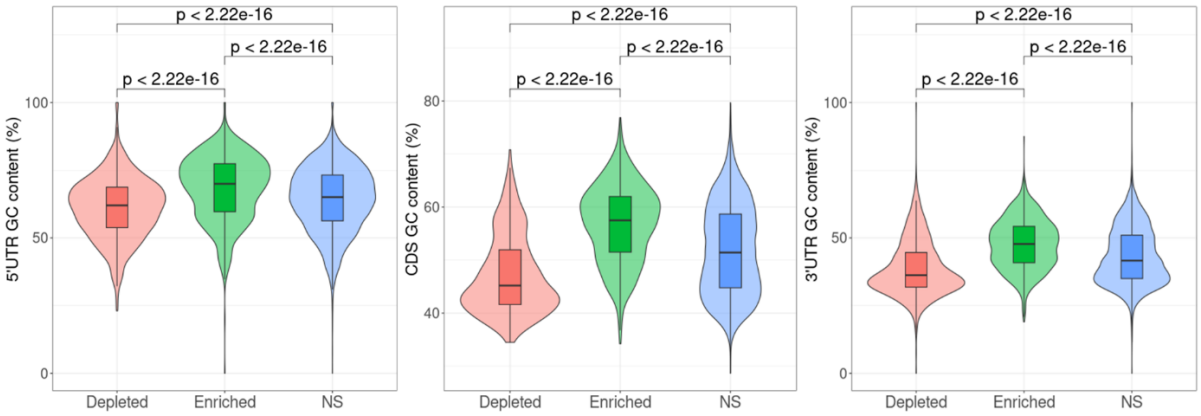


Figure 4.13 Arsenite stress granule enriched mRNAs have a GC rich sequence bias in their 5'UTR, CDS and 3'UTR.
Violin plots showing distributions with boxplots represnting minimum, maximum, interquartile range and median of percentage GC content of 5'UTR, CDS and 3'UTR. Reported p-values are from a Wilcoxon rank-sum test. Groups represent mRNAs significantly depleted, enriched or non significantly enriched or depleted (NS) in arsenite stress granule isolated fraction

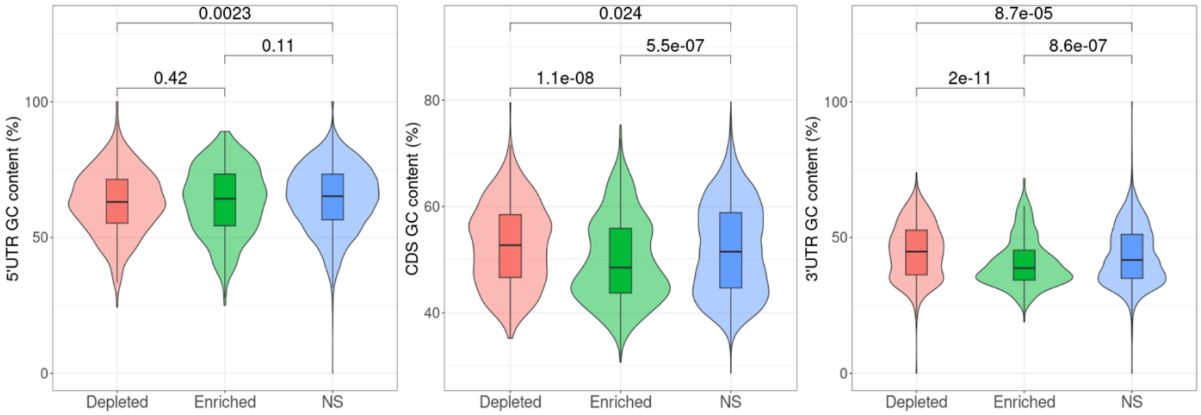


Figure 4.14 Hypoxia stress granule enriched mRNAs have an AU rich sequence bias in their CDS and 3'UTR.
Violin plots showing distributions with boxplots represnting minimum, maximum, interquartile range and median of percentage GC content of 5'UTR, CDS and 3'UTR. Reported p-values are from a Wilcoxon rank-sum test. Groups represent mRNAs significantly depleted, enriched or non significantly enriched or depleted (NS) in hypoxia stress granule isolated fraction

In addition to sequence length and nucleotide compositional bias, stress granule associated mRNAs have been shown to have an enrichment of specific motifs. In response to thapsigargin treatment AU rich motifs were identified as enriched in stress granule mRNAs using *de novo* motif enrichment analysis (Namkoong et al., 2018). There was a positive correlation between the AU rich elements identified in response to thapsigargin treatment and granule enrichment in response to two additional stresses, heat shock and arsenite. In response to heat shock, GC rich motifs were also identified as differentially enriched. In order to determine whether hypoxia or arsenite enriched mRNAs also contained differentially enriched motifs, a *de novo* motif enrichment analysis using STREME (Bailey, 2021) was implemented, using the stress granule depleted subsets as background (figure 4.15). 100 differentially enriched motifs were identified for arsenite stress granule enriched mRNAs in the STREME analysis, however, all motifs were GC rich with no AU rich elements identified as statistically significant. The top five enriched motifs are represented in figure 4.15 A. One differentially enriched motif was identified in the hypoxia stress granule enriched mRNAs. Similarly to what was described in (Namkoong et al., 2018), this is a predominantly AU rich sequence (figure 4.15 B).

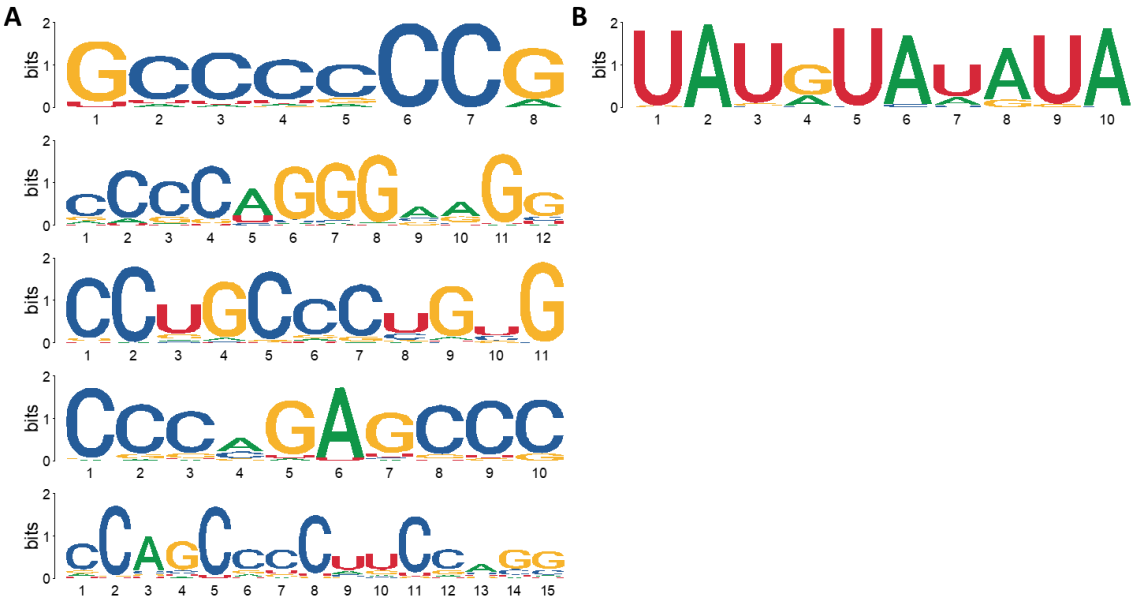


Figure 4.15 STREME *de novo* motif enrichment analysis reveals differentially enriched mRNA motifs in arsenite and hypoxia induced stress granule mRNAs.
Sequence logos of top five enriched motifs in arsenite induced stress granule mRNAs **(A)** and the only significant enriched motif in hypoxia induced stress granule mRNAs **(B).**

4.5 Generation of endogenously tagged APEX2-GFP-G3BP1 cell line.

Although transient expression of GFP-G3BP1 has enabled the elucidation of stress dependent protein and mRNA localisation to stress granules, high levels of background were a persistent feature of the work presented thus far in this thesis. Moreover, optimisation of transfections for every plasmid preparation is a time consuming and laborious process. Endogenously tagging G3BP1 may provide a cleaner purification system. A CRISPR approach was used, as developed in (Markmiller et al., 2018), to add a dual APEX2-GFP tag to endogenous G3BP1 (figure 4.16). This will also allow affinity purification by biotinylation, in addition to GFP, of both RNA and protein (Somasekharan et al., 2020, Padrón et al., 2019). This dual purification strategy may enable identification of high confidence stress granule RNAs and proteins.

In brief, U2OS cells were dual transfected with a Cas9 expression plasmid, which also contained the sequences for expression of the short guide RNA and puromycin resistance, with a linearised homologous recombination (HR) template. The HR template contained two G3BP1 homology arms flanking the desired APEX2-GFP insertion sequence. 24 hours following transfection, cells were placed in puromycin selection for an additional 72 hours then allowed to recover for 48-72 hours. GFP expression was assessed visually with a fluorescent microscope. Cells were then placed into monoclonal selection using a 96 well plate serial dilution method (John A. Ryan). Clones were expanded and fusion protein expression was validated by western blotting (figure 4.17).

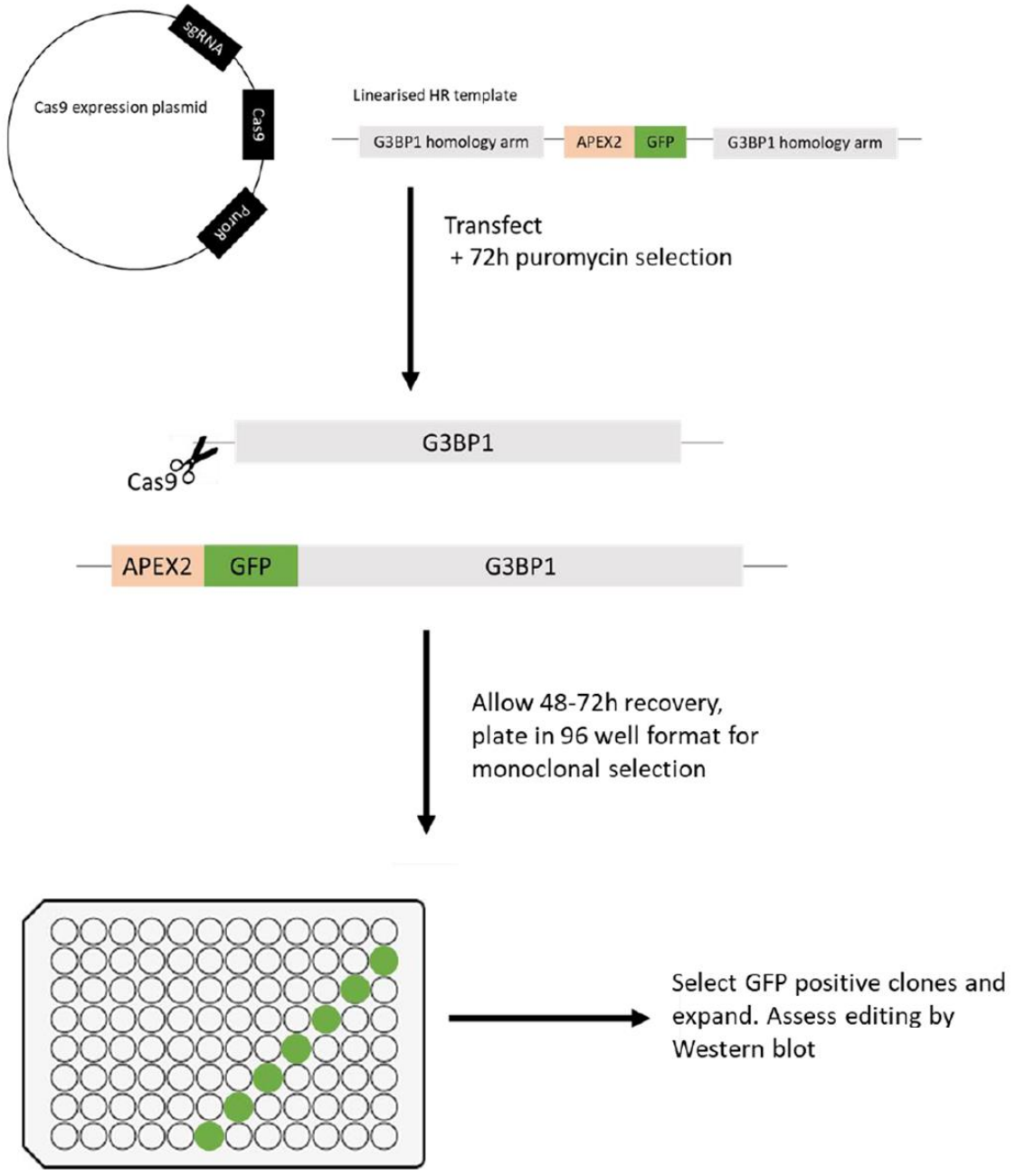


Figure 4.16 Schematic of CRISPR workflow used to generated endogenously tagged APEX2-GFP-G3BP1 in U2OS cells.
Endogenous tagging of G3BP1 using a Cas9 expression plasmid which also encodes the short guide RNA (sgRNA) and puromycin resistance (PuroR), with a linearised G3BP1 homology template that contains the APEX2-GFP insertion sequence. After puromycin selection cells were monoclonally selected by serial dilution.

There was a reduction in the intensity of the endogenous G3BP1 band compared to the wild type control cell line, indicating a heterozygous knock-in. Additionally, there was a band at the desired molecular weight of the APEX2-GFP-fusion protein that could be detected by G3BP1 and TurboGFP antibodies (figure 4.17A). For further validation of the knock-in, an siRNA approach was used. siRNAs targeting G3BP1 and a non-targeting scramble sequence control were transfected at a concentration of 20nM for 48 hours. G3BP1 expression was then analysed by western blotting (figure 4.17 B). siRNA knockdown of *G3BP1* resulted in a reduction of band intensity at both the endogenous and fusion protein molecular weights, indicating that the knock-in is indeed a consequence of editing at the G3BP1 genomic locus.

To assess if the fusion protein was functional and could be recruited to stress granules, wild type- and APEX2-GFP-G3BP1 U2OS cells were treated with 500µM arsenite for 1 hour followed by immunofluorescent imaging. Cells were co-stained with the canonical stress granule marker TIA1 (figure 4.18). In response to arsenite treatment, stress granules formed in both the wild type and APEX2-GFP-G3BP1 cell lines, as visualised by TIA1 staining. GFP expression was specific to the CRISPR edited cell line. In response to arsenite treatment, GFP positive cytoplasmic foci formed that co-localised with TIA1, indicating that the fusion protein can be recruited to stress granules.

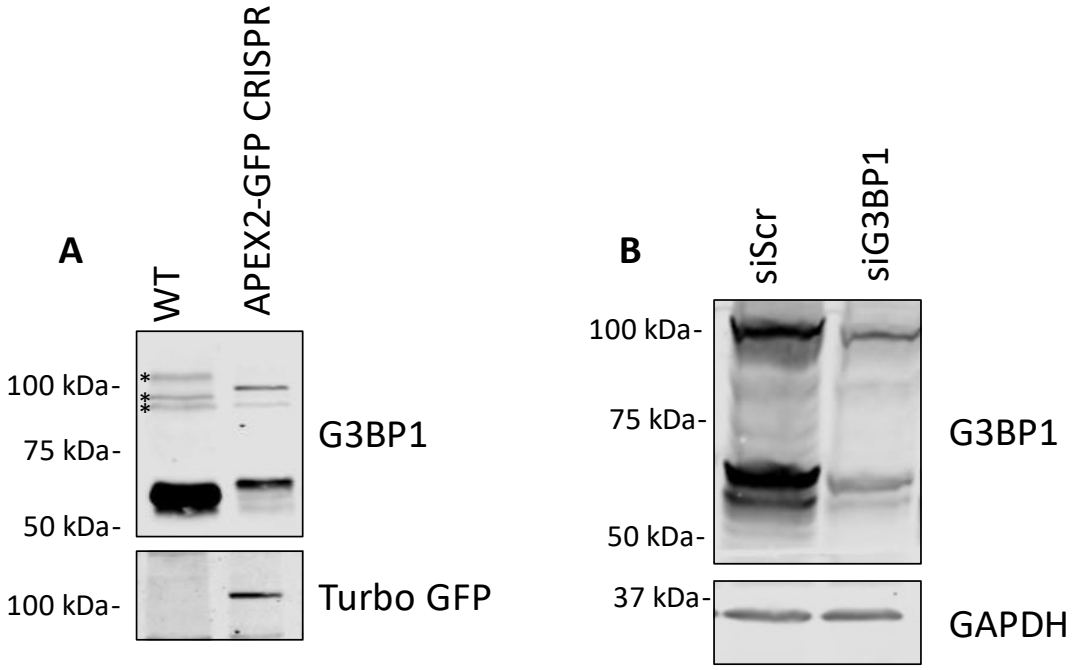


Figure 4.17 APEX2-GFP-G3BP1 fusion protein can be detected by western blot and is responsive to siRNA knockdown.

A) APEX2-GFP-G3BP1 can be detected by G3BP1 and Turbo GFP antibodies but not in wild type (WT) U2OS cells. There is a reduction in signal of the endogenous G3BP1 band indicating a heterzygous knock-in. Equal volumes of cell lysate loaded. Representative image of 3 biological replicates. * denotes suspected non-specific band

B) Treatment with an siRNA targetting G3BP1 results in a reduction in band intensity for both endogenous- and APEX2-GFP-G3BP1. Equal volumes of cell lysate loaded. Representative image of 3 biological replicates.

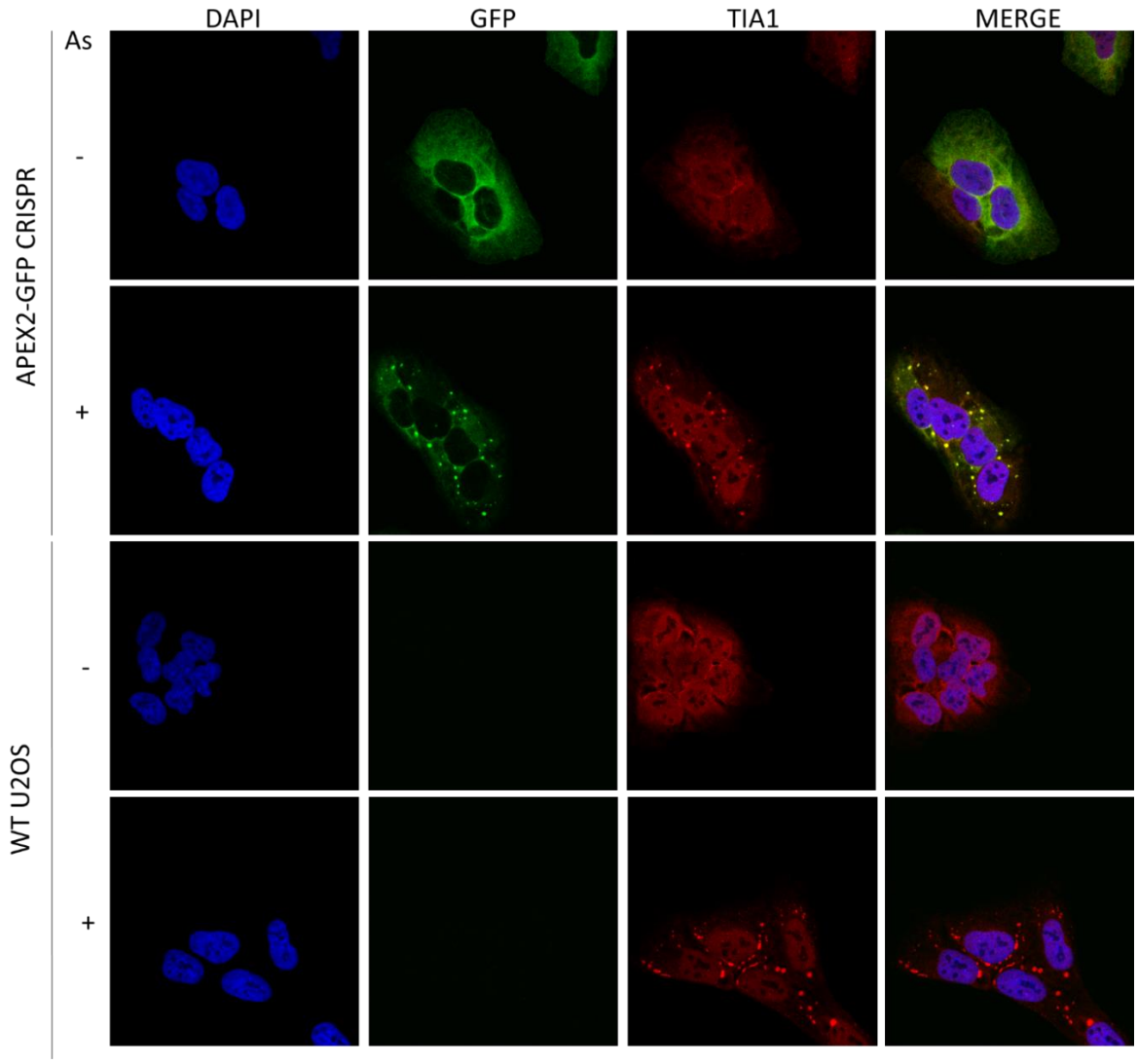


Figure 4.18 APEX2-GFP-G3BP1 localises to stress granules in response to arsenite stress. Immunofluorescent images of wild type (WT) U2OS cells or CRISPR edited APEX2-GFP-G3BP1 in response to arsenite stress. GFP expression is specific to APEX2-GFP-G3BP1 cell line. APEX2-GFP-G3BP1 fusion protein localises to sress granules in response to arsenite stress as visualised by co-localisation with TIA1. Nuclei are visualised by DAPI staining. Representative image of 3 biological replicates.

4.6 Chapter discussion

The stress granule isolation methodology developed in the previous chapter was successfully used to generate RNAseq libraries (figures 4.1-4.3). This method enabled the identification of stress granule enriched and depleted mRNAs in response to both arsenite and hypoxia stress (figures 4.5-4.7). For arsenite stress, 920 mRNAs were enriched in the stress granule isolated fraction with 920 mRNAs being depleted. 713 stress granule enriched and 743 stress granule depleted mRNAs, that had not been identified in (Khong et al., 2017), were identified and represent a novel dataset of stress granule mRNAs in response to arsenite. For hypoxia stress, 451 mRNAs were enriched in the stress granule isolated fraction with 510 being depleted. 344 stress granule enriched and 403 stress granule depleted mRNAs were novel to this thesis and not identified in (Khong et al., 2017). Gene set enrichment analysis highlighted a convergence on critical pathways in cancer, such as KRAS and MYC, for both stress granule enriched and depleted mRNAs (figure 4.10). These findings further implicate stress granules as critical mediators of cancer progression. A better understanding of how stress granules regulate gene expression networks may highlight mechanisms of chemotherapeutic resistance (Grabocka and Bar-Sagi, 2016).

Although statistically significant, the overlap observed between arsenite and hypoxia enriched/depleted mRNAs in this thesis and the previously published stress granule transcriptome is low in terms of absolute numbers (figures 4.6-7, 4.9). There are a number of experimental differences that may account for this variation. These include, but are in no way limited to, different stress induction methods, including different concentration and length of arsenite treatment, immunoprecipitation method, library preparation and bioinformatic analyses. Stress granules are known to have temporally variable protein and RNA content (Padrón et al., 2019), which may account for some of the variability between the arsenite treatment used in this thesis and the data from (Khong et al., 2017). Stress granule RNA content is also known vary depending on the stress applied (Padrón et al., 2019, Khong et al., 2017, Namkoong et al., 2018). Notably, the authors in this study observed no statistically significant overlap with (Khong et al., 2017) in response to hippuristanol treatment. This stress specific stress granule recruitment may account for some of the variability between hypoxia stress granule associated mRNAs and those published in (Khong et al., 2017).

The statistical overlap for the groups of mRNAs identified as enriched or depleted in all conditions, including the two negative controls was orders of magnitude greater than for the stress specific groups (figures 4.8-9). These groups also shared close to 40% identity with the mRNAs identified in (Khong et al., 2017). These results show that, as with many isolation or purification techniques, there will be inherent biases that need to be accounted for by inclusion of robust controls. It is however important to note at this point that mRNAs whose stress granule localisation that was validated by means of smFISH (AHNAK, PEG3 and DYNC1H1) were present in the subset that was enriched in all conditions, including the controls. One limitation of the analysis used in this chapter is that it does not account for potential enrichment of an mRNA in the stress granule isolated fraction in response to either arsenite or hypoxia versus the control conditions. It is, therefore, reasonable to hypothesise that a more nuanced analysis method that accounts for this may increase both the overlap with (Khong et al., 2017) and enable the identification of more novel stress granule mRNAs.

mRNA length has been identified as a determining factor in stress granule localisation in response to numerous stresses (Namkoong et al., 2018, Khong et al., 2017, Padrón et al., 2019). The results presented in figures 4.11-12 are in agreement with these findings as both arsenite and hypoxia stress granule mRNAs are longer than either stress granule depleted mRNAs or non-significantly enriched or depleted mRNAs. It has been proposed that length contributes to stress granule localisation for two reasons (Khong et al., 2017). Firstly, longer mRNAs may harbour multiple binding sites for RNA binding proteins that could promote partitioning into stress granules. Secondly, longer mRNAs could facilitate more RNA-RNA mediated interactions. Which of these factors is rate-limiting or whether both these mechanisms work simultaneously or synergistically is still yet to be determined.

In addition to unique stress dependent subsets of stress granule mRNAs being identified in response to arsenite and hypoxia, stress specific sequence compositional biases were also identified (figures 4.13-14). Arsenite stress granule mRNAs had a GC sequence bias across their 5’UTR, CDS and 3’UTR. Hypoxia stress granule mRNAs had an AU sequence bias in their CDS and 3’UTR. AU sequence bias specifically was identified as a feature of stress granule mRNAs in (Khong et al., 2017). Differential motif enrichment analysis identified GC rich motifs in arsenite stress granule mRNAs (figure 4.15 A). One motif was identified

in hypoxia stress granule mRNAs, this was a predominantly AU rich sequence (figure 4.15 B). AU rich elements were previously identified as a feature of stress granule mRNAs (Namkoong et al., 2018). It is important to note that in Namkoong et al, the correlation of AU elements and stress granule fraction enrichment was lower in response to arsenite stress than for thapsigargin treatment or heat shock. Furthermore, AU rich elements were identified by *de novo* motif enrichment analysis for thapsigargin and heat shock stress enriched mRNAs but not arsenite. AU rich elements identified in response to thapsigargin treatment were used for the further analyses on the arsenite stress granule enriched mRNAs. In addition to AU rich elements, heat shock stress granule fraction enriched mRNAs contained GC rich motifs. Furthermore this study was conducted in a mouse (MEF) cell line, indicating that species variation of stress granule mRNA features may also be playing a role. Identification of alternate sequence feature properties for arsenite and hypoxia stress granule mRNAs is a fascinating observation for two reasons. It suggests that stress granule mRNAs will be exposed to different regulatory elements, such as RNA binding proteins. Utilisation of publicly available eCLIP datasets may shed light on whether different RNA-protein interaction networks are present in stress granules in response to different stresses. Secondly these findings further highlight that stress granule formation and recruitment of mRNAs is not a consequence of bulk translation inhibition or a biophysical phenomenon limited to mRNAs with particular features. It is a bespoke, stress dependent cytoprotective response.

In order to try combat the high levels of background observed in stress granule isolations using transient GFP-G3BP1 expression, a CRISPR approach was used to generate an endogenous APEX2-GFP-G3BP1 fusion protein (figures 4.16-18). The fusion protein appears functional, as it colocalises to stress granules with TIA1 in response to arsenite stress. This will allow a dual purification strategy, namely anti-GFP immunoprecipitation and biotin proximity-ligation, to determine stress granule mRNAs and proteins with high confidence.

In conclusion, the two aims of this chapter were successfully fulfilled. The methodology of stress granule isolation developed in the previous chapter was adapted to allow RNAseq analysis of stress granule mRNAs in response to arsenite and hypoxia stress. Additionally, this allowed the identification of unique stress specific stress granule mRNAs that have differing sequence feature properties. What is now required is a deeper understanding of how stress granules regulate gene expression. The following chapter will use ribosome profiling, to try to understand the gene expression landscape in response to arsenite and hypoxia stress.

5 Chapter 5: Ribosome profiling in arsenite and hypoxia induced stress.

5.1 Chapter Introduction

In the previous two results chapters stress granule proteomics and RNAseq were conducted. Stress granule proteomics highlighted the localisation of proteins of the large ribosomal subunit were present in hypoxia induced stressed granules. This led to the development of the hypothesis that in response to hypoxia, stress granules may represent a more translationally active compartment than what has previously been described for eIF2α phosphorylation driven stress granules. RNAseq revealed that stress granule mRNA localisation is largely stress dependent and stress granule mRNAs have distinct feature properties in response to different stresses. This suggests that stress granules are a stress specific response.

In order to understand how mRNA localisation to stress granules influences translation rates, a ribosome profiling approach will be used. Ribosome profiling is a technique that was originally developed in (Ingolia et al., 2009). Ribosome profiling combines the well-established methods of sucrose density gradients with enzymatic mRNA digestion. The ~30 nt stretch that is bound by the ribosome is protected from digestion. Ribosome protected fragments can then be purified and used for the generation of RNAseq libraries. Deep sequencing combined with bioinformatic methods allows the study of translation on a single codon scale. Ribosome profiling can be used for differential expression analysis but the positional element of the data generated has enabled the characterisation of many novel open reading frames (Mudge et al., 2022) and provided great insights into the mechanisms of translation regulation (Jackson and Standart, 2015).

The aims of this chapter are:

1. Generate high quality ribosome profiling libraries in conditions of sodium arsenite and hypoxia induced stress.

2. Integrate ribosome profiling data with stress granule RNAseq data to understand how stress granule localisation regulates translation.

5.2 Generation of libraries for ribosome profiling

Ribosome profiling, originally developed in (Ingolia et al., 2009), was used to interrogate gene expression in response to stress. U2OS cells were stressed using 200 mM sodium arsenite for three hours or hypoxia (0.1% O_2 for six hours). Non-treated cells were used as a control. For harvesting, culture media was removed and stored in 50 ml tubes. Cells were washed in PBS. PBS was added to media in 50ml tubes and placed on ice. Cells were trypsinised and plates were placed on ice. Trypsinised cells were added to media/PBS and pelleted by centrifugation. Supernatant was aspirated and cell pellets were resuspended in ice cold PBS followed by pelleting by centrifugation. PBS was aspirated and cells were flash frozen in liquid nitrogen.

10-50% sucrose gradients were made using a Biocomp Gradient master and cooled to 4°C (figure 5.1 A). Sucrose density gradients allow fractionation of the mRNAs in lysates into sub-polysomal and polysomal fractions. Sub-polysomal fractions represent mRNAs with one or less ribosomes bound. These mRNAs are poorly translated. Polysomal fractions contain mRNAs bound by two or more ribosomes, representing actively translating, highly expressed mRNAs (Poria and Ray, 2017). Polysome gradient profile plots can be generated to compare global translation rates by measuring 260 nm absorbance during fractionation.

Cell pellets were thawed on ice for five minutes followed by lysis. Lysates were cleared of cellular debris and nuclei by centrifugation. Cleared lysate was loaded onto a sucrose density gradient followed by ultracentrifugation at 38000 rpm for 2 hours at 4°C. Samples were then fractionated and 260 nm absorbance was measured (figure 5.1 B). Treatment with sodium arsenite resulted in translational inhibition, in agreement with previously published literature (McEwen et al., 2005, Somasekharan et al., 2020, Tauber et al., 2020), when compared to the control. This is evident due to the increase in the proportion of sub-polysomal to polysomal RNA. Polysomal distribution in response to hypoxia were comparable to those in the control. In order to ensure a reduction in translation elongation rates was not retaining

mRNAs in the polysomes, a puromycin incorporation assay was used. Puromycin mimics a tRNA and enters the A-site of the ribosome. Puromycin is then incorporated into the peptide and translation is terminated (Pestka, 1971). Antibodies that recognise puromycilated peptides have been developed, which allows for elongation rates to be assayed. U2OS cells were subjected to sodium arsenite, or hypoxia induced stress for 1, 2, 3, 4 or 6 hours. 10 minutes prior to harvest cells were treated with puromycin. Two unstressed control samples were used, one treated with puromycin, and one not treated with puromycin (-puro) to account for background. Whole cell lysates were generated and samples were analysed by western blotting (figure 5.2). Low levels of background are present in the no puromycin control (-puro). Puromycin incorporation was drastically reduced in response to sodium arsenite treatment, indicating translational repression. Puromycin incorporation was comparable in all hypoxia treated samples compared to the control. This indicates that translation elongation is not repressed in response to hypoxia.

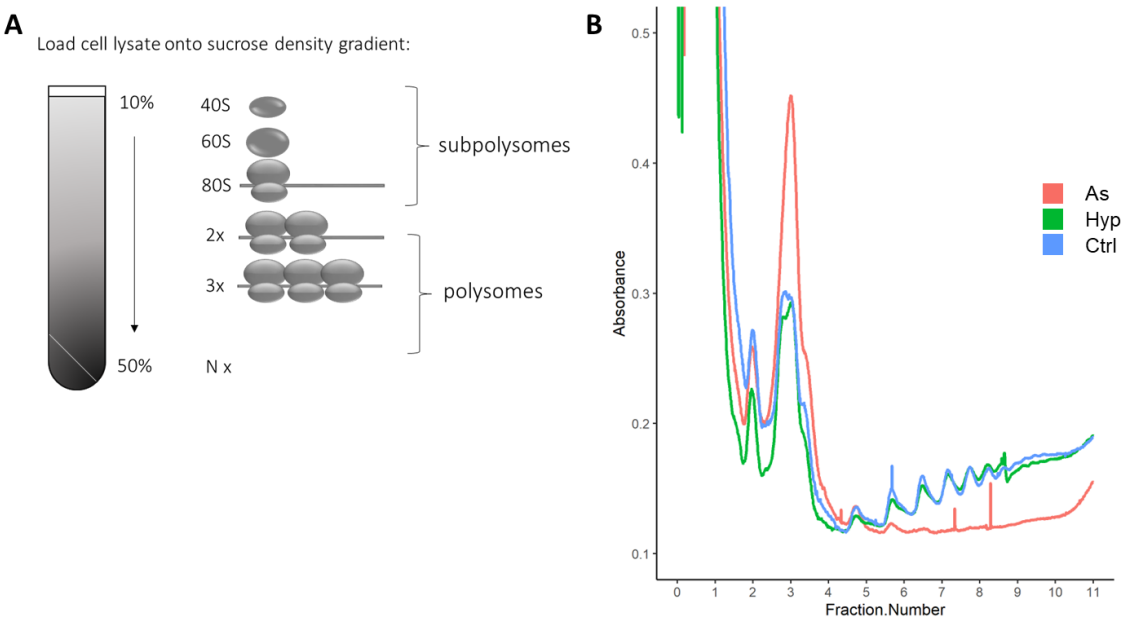


Figure 5.1 Sodium arsenite induces translational repression compared to control and hypoxia

A) Schematic of sucrose density gradient. Ultracentrifugation allows fractionation of cell lysates into subpolysomes and polysomes. mRNAs bound by multiple ribosomes (polysomes) sediment further into the gradient. This technique can be used to assess translation rates by measuring absorbance whilst fractionating.

B) Polysome gradient profile plot of U20S cells treated with sodium arsenite (As), hypoxia (Hyp) or non-stressed control (Ctrl). Y axis represents 260 nm absorbance. Sodium arsenite treatment results in an increased subpolysomal distribution, indicating translational repression. Hypoxia and Ctrl samples have similar distributions.

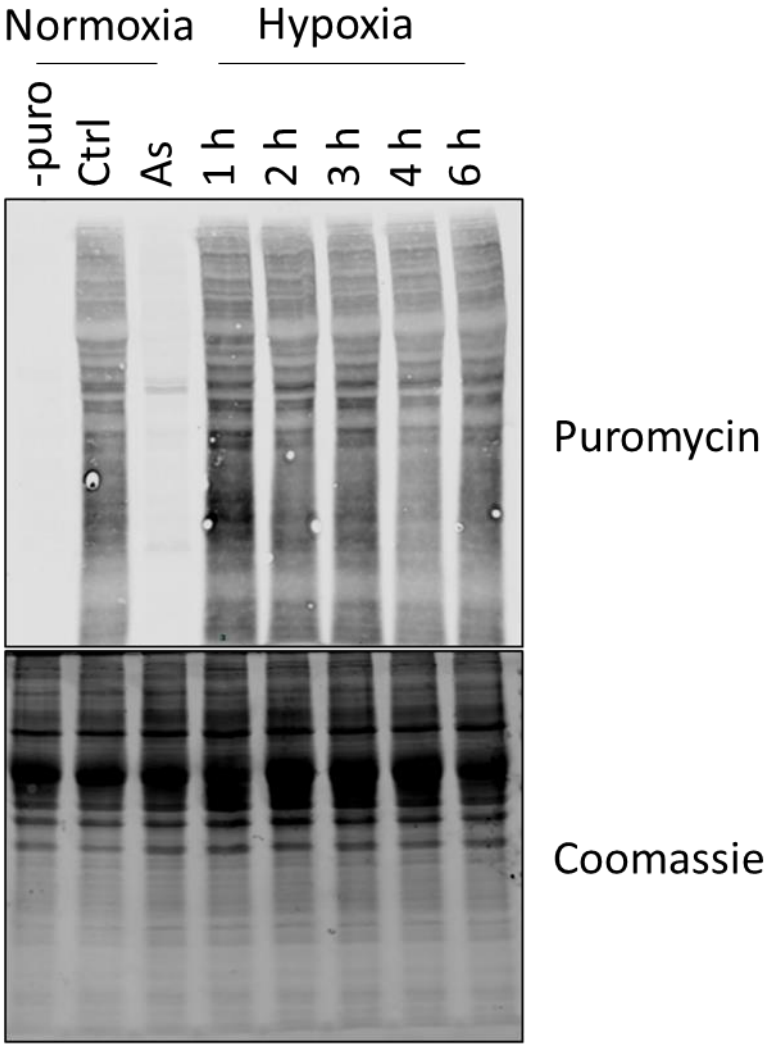


Figure 5.2 Translation elongation is not inhibited in hypoxia
U2OS cells were treated with sodium arsenite (As) or cultured in hypoxia for 1, 2, 3, 4 or 6 hours. Ctrl denotes non stressed control, -puro denotes non stressed control that was not treated with puromycin. Cells were treated with puromycin 10 minutes prior to harvest. Puromycin incorporation was analysed by western blotting. Low levels of background are present in the -puro control. Treatment with sodium arsenite induces translational repression. Puromycin incorporation rates are comparable in Ctrl and hypoxia conditions, indicating translation elongation is not inhibited in hypoxia. Coomassie gel is used as a loading control. Equal volumes of cell lysate loaded. Representative image of 3 biological replicates.

To isolate ribosome protected fragments (RPFs), lysates were digested using 7 µl RNase I for 40 minutes at 22°C with shaking at 650 rpm, prior to loading onto sucrose density gradients. RNase I digestion reaction was quenched using superaseIn. Digested lysates were then fractionated and fractions 4-6 were collected (figure 5.3). Large peaks representing the monosome were present in control and hypoxia stressed samples. A considerably lower monosome peak is observed in the sodium arsenite stressed sample. This is likely due to the lower translation rates in this condition (figure 5.1 B). RNA from the collected fractions was acid-phenol chloroform extracted followed by overnight sodium acetate ethanol precipitations at -20°C. RNA pellets were twice washed in 75% ethanol before being resuspended in 10 µl nuclease free water. Fractions from each condition were then pooled. Pooled fractions were run on a 15% TBE-urea gel to enable size selection. 28-, 34-, 53-, and 60 nt markers were also loaded on the gel to enable accurate size selection. Gels were stained using SYBR Gold and imaged using a Typhoon Phospho Imager. The area of the gel inclusive of the 28 nt marker but exclusive of the 34 nt marker was excised. A representative gel, from replicate 4 of the hypoxia stressed RPFs, is shown in figure 5.4 A.

RNA was extracted from the gel overnight at 16°C with shaking at 600 rpm. SpinX columns were used to remove residual gel pieces. RNA was precipitated overnight in sodium acetate isopropanol at -20°C, followed by two ethanol washes and resuspension in nuclease free water. RNA was then T4 PNK treated to enable adapter ligation in the library preparation protocol. Following T4 PNK treatment, RNA was precipitated overnight in sodium acetate and ethanol, then washed and resuspended as described above. RPF libraries compatible with NGS were generated using a Nextflex small RNA kit v3-19.01. RPF libraries were PCR amplified using 12 cycles. The final PCR product was run on a 6% TBE-PAGE gel (figure 5.4 B). Gel was Sybr Gold stained and imaged as described above. A band of expected size of > 150 bp was present in all conditions. This band was gel extracted and purified according to Nextflex small RNA kit v3-19.01.

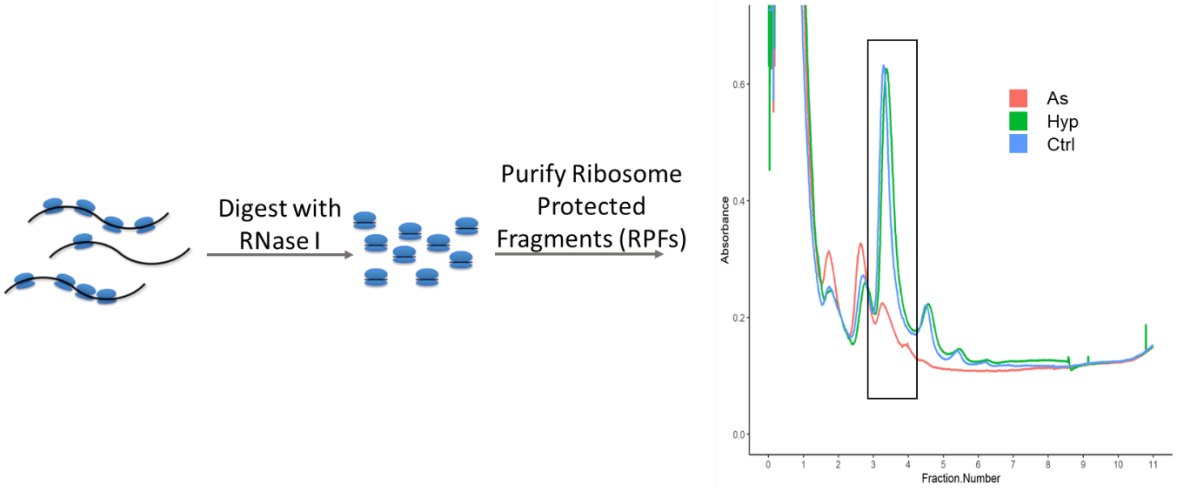


Figure 5.3 RNAse I digestion enables isolation of ribosome protected fragments
Treatment of lysates with RNAse I cleaves mRNAs. The portion of mRNA bound by the ribosome bound by the ribosome is protected from RNAse I cleavage, thus they are called ribosome protected fragments. Digested lysates are then loaded on a sucrose gradient and fractionated. Digestion yields in an increased peak representing the monosome in control (Ctrl) and hypoxia treated samples (Hyp), highlighted by the inset box. A similar increase in monosomes is not present in sodium arsenite treated sample (As). This is likely due to the bulk translational repression in this condition.

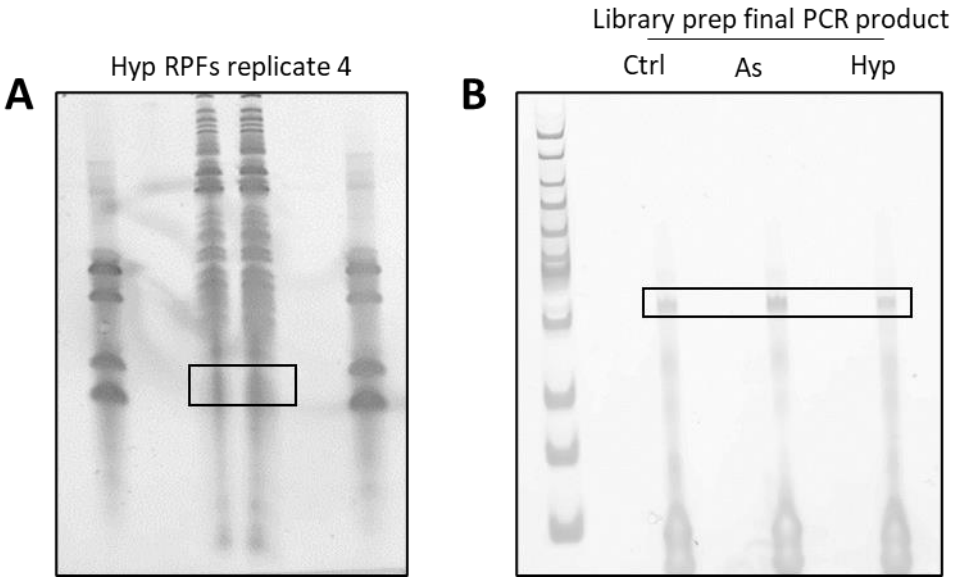


Figure 5.4 Ribsome protected fragments can be purified and final library preparation product is of expected size.

A) Representative Sybr Gold stained TBE Urea gel of purified fractions containing RPFs from digested gradient. Area inclusive of the 28 nt marker but exclusive of the 34 nt marker is excised from the gel and used for library preparations. Inset box represents the gel area excised.

B) Representative Sybr Gold stained TBE-PAGE gel of final small RNA library prep product for control (Ctrl), sodium arsenite treated (As), or hypoxia treated (Hyp) samples. A final PCR product >150 bp is present in all samples, highlighted by the inset box.

Gel extracted PCR products were analysed using a TapeStation D1000 high sensitivity screen tape. A representative electropherogram with RPF libraries generated in replicate 4 is shown in figure 5.5. Average PCR fragment length was ~170 nt, which is the expected size of RPF libraries (McGlincy and Ingolia, 2017).

Total RNA was extracted from undigested lysate using Trizol. Libraries for total RNAseq were generated in parallel by ribosomal RNA depletion using RiboCop v2 followed by library generation using a CORALL library prep kit. Library concentration was measured using a Qubit 2.0 fluorometer with high sensitivity dsDNA reagents. Total and RPF libraries were pooled to a final concentration of 2nM. Pooled libraries were then sequenced by the Beatson Institute Molecular Technologies Services, using a NextSeq 500, 75 cycles high output kit.

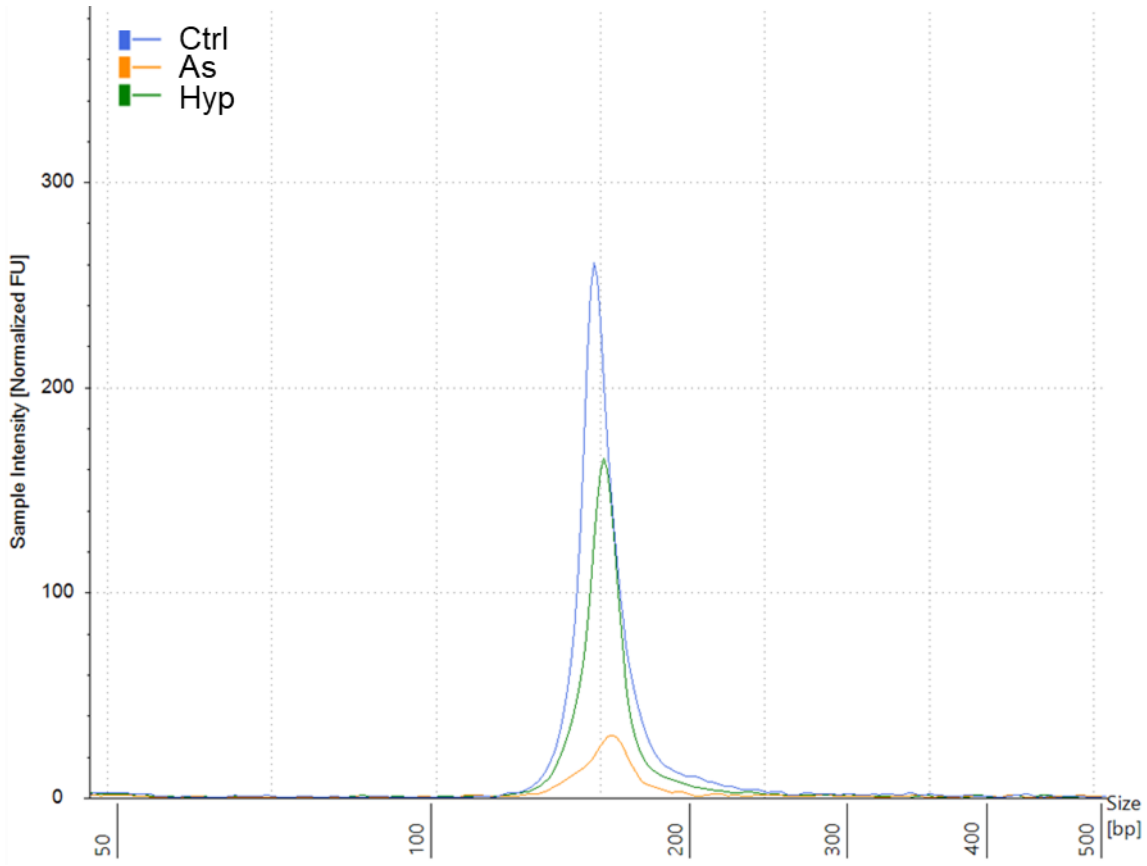


Figure 5.5 Ribosome profiling libraries are of expected size when run on Tapestation. Representative electropherogram from Tapesation D1000 screen tape f libraries generated from replicate 4. Tapestation allows for accurate calculation of fragment size. Fragments were ~170 bp long for non-stressed control (Ctrl), sodium arsenite (As) or hypoxia (Hyp) treated samples.

5.3 Ribosome profiling initial data processing and quality control

Raw sequencing data for total- and RPF RNAseq was demultiplexed and converted from bcl format to fastq using the Illumina software package bcl2fastq. Quality of the sequencing data was assessed using the fastqc software package from Babraham Bioinformatics. Adaptor sequences were trimmed using cutadapt (Martin, 2011). cd-hit-dup was used for deduplication to remove overrepresented PCR artefacts from the library generation methodologies (Fu et al., 2012). UMIs were the trimmed from sequencing reads using cutadapt (Martin, 2011).

The analysis pipeline used requires that RPFs are aligned to a transcriptome containing one transcript per gene. In order to overcome this issue, the total RNAseq data is used to calculate the most abundant transcript per gene using RSEM (Li and Dewey, 2011). BBMap was first used to remove all ribosomal RNA reads, mitochondrial mRNA and tRNA reads from RPF data. Sequence reads were aligned to the GENCODE v38 human transcriptome, which had been pre filtered to contain only protein coding transcripts. Custom python scripts were then used to generate counts files suitable for assessing RPF library quality and for calculation of the offset required to identify the ribosome P-site. High quality RPF libraries have distinct features when compared to normal RNAseq libraries. Firstly, RPFs are ~30 nt long (McGlincy and Ingolia, 2017, Jackson and Standart, 2015) in mammalian cells (figure 5.6 A). In addition to this, reads should be enriched for the CDS of mRNAs as they are derived from translating ribosomes. In figure 5.6 B-D representative plots from replicate 1 summarising read length and proportion of reads that arise from the CDS, 5- or 3’UTR are shown. For control, sodium arsenite and hypoxia stressed cells read lengths peak at ~30 nt, with the vast majority of reads located within the CDS of mRNAs. This indicates that they represent genuine RPFs from translating ribosomes.

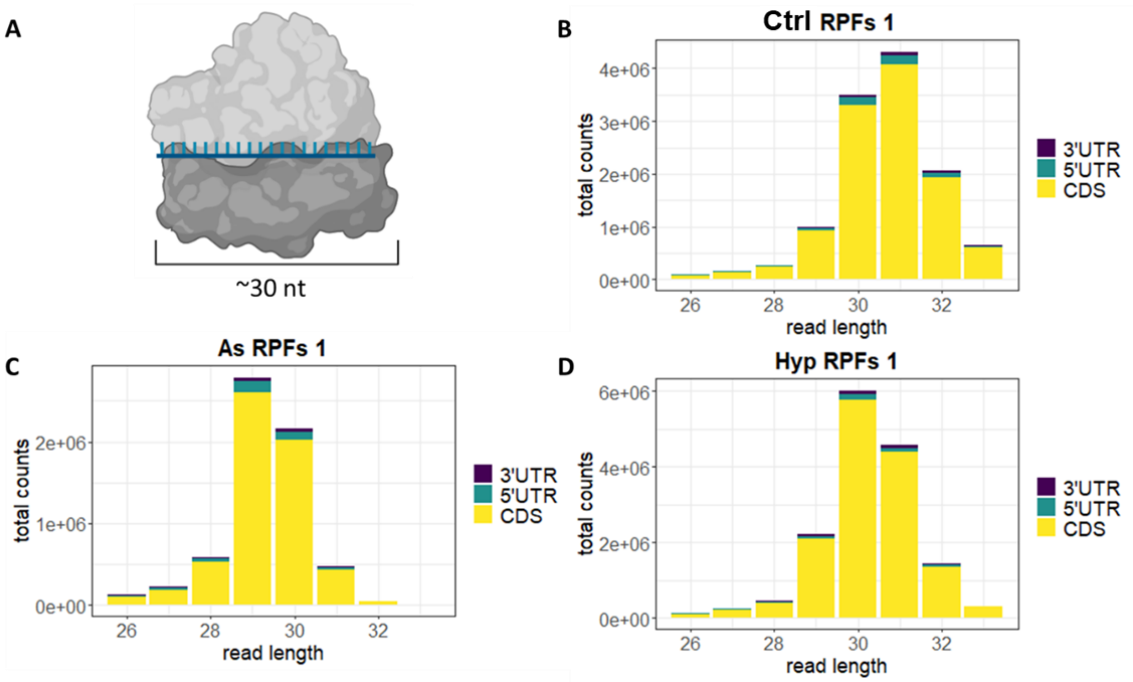


Figure 5.6 Sequenced ribosome profiling libraries are of expected size and show CDS enrichment

A) Schematic depicting a ribosome and ribosome protected fragment. Ribosome protected fragments are expected to be ~30 nt in mammalian cells.

B-D) Representative bar plots showing total counts from sequencing run by fragment size. Proportion of reads originating from the CDS, 5'- and 3' UTR is indicated. Reads in non-stressed control (Ctrl), sodium arsenite (As) and hypoxia (Hyp) treated samples. All samples show a peak in reads at ~30 nt and show CDS enrichment, indicating libraries are composed of genuine RPFs.

Another distinct feature of RPF libraries is trinucleotide periodicity (Ingolia et al., 2009). As a ribosome decodes an mRNA one codon (codon = 3nt) at a time (figure 5.7 A). In order to assess periodicity, counts were normalised to counts per million (cpm) for all conditions. Mean cpm for all conditions were then plotted for the last 25 nt of the 5'UTR, first and last 50 nt of the CDS, and the first 50 nt of the 3'UTR (figure 5.7 B). This plot further emphasises the enrichment of CDS counts versus either UTR. Additionally, strong periodicity is observed. These data presented in figures 5.5-6 indicate that high quality RPF libraries have been generated, that are suitable for downstream analyses.

Both total and RPF RNAseq datasets were loaded into R. Total RNAseq RSEM data was loaded using tximport (Soneson et al., 2015). DESeq2 was used for differential expression analysis using an experimental design that accounted for condition and batch correction by replicate (Love et al., 2014). Low abundance genes were removed by filtering genes that had an average of less than 10 counts across all samples. Log fold change shrinkage was applied using the apeglm method (Zhu et al., 2018). PCA plots were generated for total- and RPF RNAseq libraries to visualise the degree of clustering between samples (figure 5.8). All conditions formed distinct clusters in both total (figure 5.8 A) and RPF samples (figure 5.8 B). This indicates that distinct gene expression programmes are utilised in response to sodium arsenite or hypoxia stress at the level of both total RNA and translation.

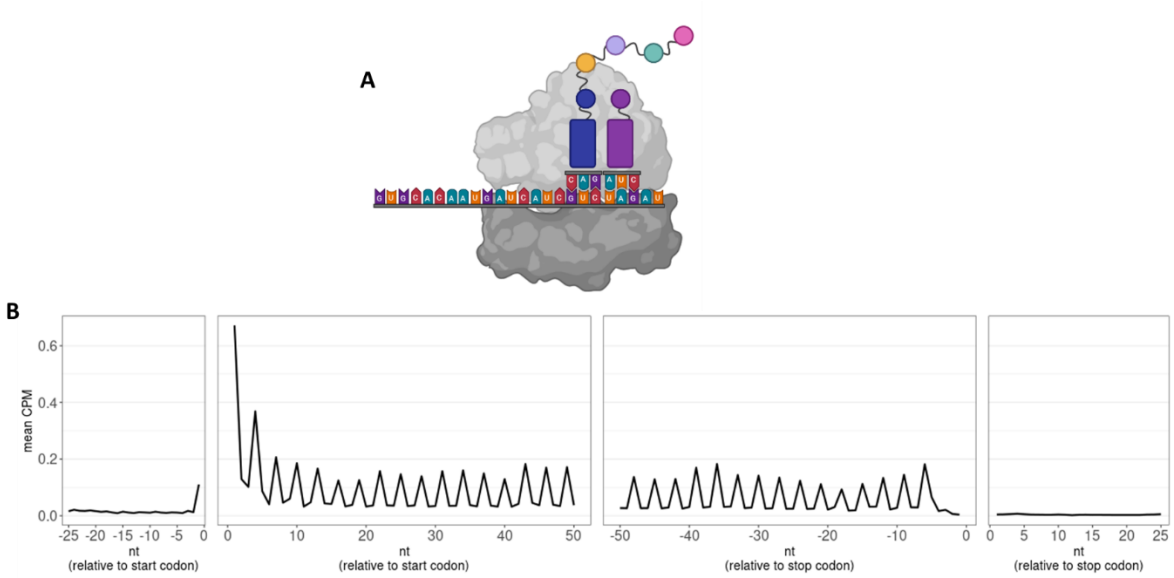


Figure 5.7 Ribosome profiling libraries show trinucleotide periodicity

A) Schematic of a translating ribosome decoding an mRNA 1 codon (3 nucleotides) at a time

B) Mean counts per million (CPM) plotted for all conditions for the last 25 nucleotides of the 5'UTR, first and last 50 nucleotides of the CDS and last 25 nucleotides of the 3'UTR. High levels of trinucleotide periodicity are observed, with the majority of reads originating from the CDS which are characteristics of high-quality ribosome profiling libraries.

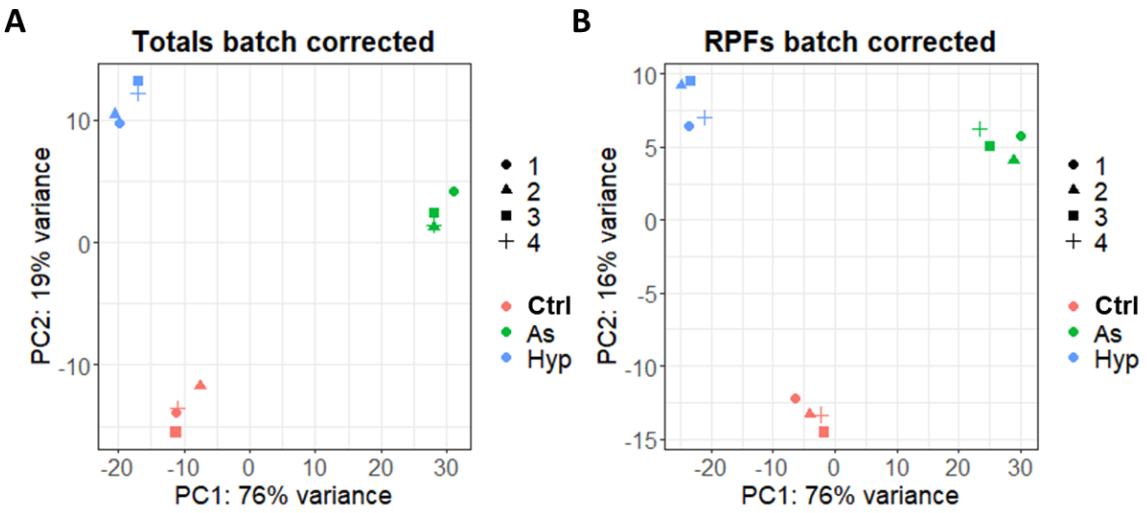


Figure 5.8 PCA analysis shows a high degree of clustering for all conditions for both total and RPF samples

A) PCA plot for batch corrected total RNAseq samples. A high degree of clustering is observed in the non-stressed control (Ctrl), sodium arsenite (As) and hypoxia treated samples. This indicates that distinct gene expression programmes are utilised at the level of total RNA in response to sodium arsenite and hypoxia.

B) PCA plot for batch corrected RPF samples. Similar to the total samples, a high degree of clustering is observed for all conditions. This indicates that distinct gene expression programmes are utilised at the level of translation in response to sodium arsenite and hypoxia.

5.4 Hypoxia stress granule mRNAs are translationally upregulated.

Stress granule mRNAs in response to eIF2α phosphorylation have previously been described as having low translation efficiency (TE) in response to various stresses (Khong et al., 2017, Namkoong et al., 2018, Padrón et al., 2019). Intriguingly, this was not conserved in response to eIF4A1 inhibition by hippuristanol treatment (Padrón et al., 2019). In response to hippuristanol treatment stress granule enriched mRNAs did not have statistically significantly lower TE values compared to stress granule depleted mRNAs. In order to determine the TE of mRNAs in response to sodium arsenite or hypoxia, DESeq2 was used as described in (Chothani et al., 2019). TE represents ribosome occupancy on an mRNA normalised to total RNA levels. As initiation is believed to be the rate limiting step of translation, this can be used as a surrogate measure for protein production sue to the fast decoding speeds of the ribosome (Chothani et al., 2019). This method integrates both RPF and total RNAseq data into a single DESeq2 analysis matrix and allows for a more accurate estimation of changes in TE than previously described methods. Both sodium arsenite and hypoxia stressed samples were compared to the control and $\log_2$ fold change in TE was calculated.

The primary objective of this chapter is trying to understand how mRNA compartmentalisation into stress granules regulates translation in response to sodium arsenite or hypoxia stress. The remainder of this section will focus on integrating the stress granule mRNA sequencing generated in the previous chapter with the ribosome profiling generated in this chapter.

Ribosome profiling data was subsetted into groups containing stress granule enriched, stress granule depleted, and non-significantly enriched or depleted mRNAs defined in the previous chapter. Average TE $\log_2$ fold change was then compared between these groups (figures 5.9-10). In response to sodium arsenite stress, stress granule enriched mRNAs had lower TE values than either the stress granule depleted or non-significantly enriched or depleted subsets. Stress granule depleted mRNAs on average had higher TE values than the non-significantly enriched or depleted subset (figure 5.9 A). This is in concordance with what has previously been described in the literature (Khong et al., 2018, Namkoong et al., 2018,

Padrón et al., 2019). These patterns of translational regulation for stress granule enriched or depleted mRNAs are depicted in the TE scatter plot (figure 5.9B). Although it is clear that in general stress granule enriched mRNAs are translationally repressed in response to sodium arsenite, this is not the case for all mRNAs. This supports the hypothesis from (Somasekharan et al., 2020), that selective partitioning of some G3BP1 associated mRNAs into polysomes occurs in response to arsenite stress. These findings suggest that selective partitioning into polysomes also occurs specifically for stress granule localised mRNAs.

Remarkably, hypoxia stress granule enriched mRNAs had higher average TE values than either the stress granule depleted or the non-significantly enriched or depleted subsets (figure 5.10 A). This result supports the hypothesis developed at the end of the first chapter, suggesting that hypoxia induced stress granules may be a more translationally active compartment than what has been previously described in response to sodium arsenite. A possible explanation for this is the localisation of constituent proteins of the large ribosomal subunit. Stress granule depleted mRNAs have lower TE values than either the enriched or non-significantly enriched or depleted subsets. Comparably to what was observed in the scatter plot for sodium arsenite stress, stress granule enrichment or depletion of an mRNA is not the sole determinant of its TE (figure 5.10 B). This is evident as some stress granule enriched mRNAs are translationally repressed whilst some stress granule depleted mRNAs are translationally upregulated.

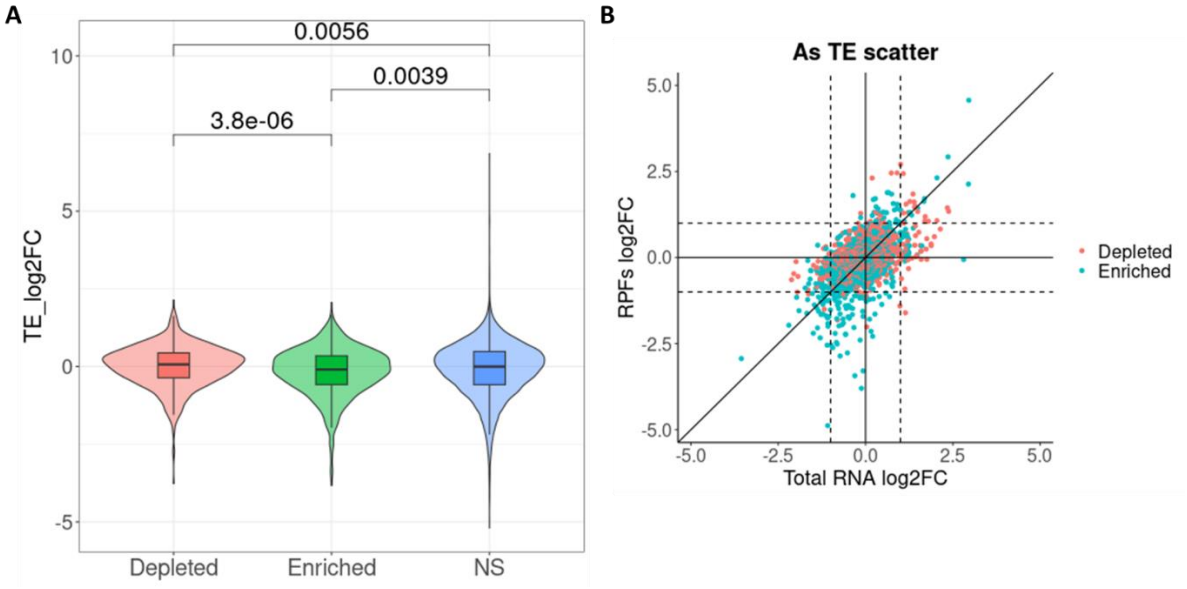


Figure 5.9 Sodium arsenite induced stress granule mRNAs are translationally repressed

A) Violin plots showing distributions with boxplots representing minimum, maximum, interquartile range and median of TE log2 fold change. Groups represent stress granule depleted, enriched or non-significantly enriched or depleted (NS) mRNAs in response to sodium arsenite. P-values are from a Wilcoxon rank-sum test.

B) Scatter plot representing translation efficiency of stress granule enriched or depleted mRNAs in response to sodium arsenite. In general, stress granule mRNAs are translationally repressed however some stress granule mRNAs are not repressed. This suggests dynamic partitioning between stress granules and the polysomes can occur, similar to what has been described for G3BP1 bound mRNAs previously.

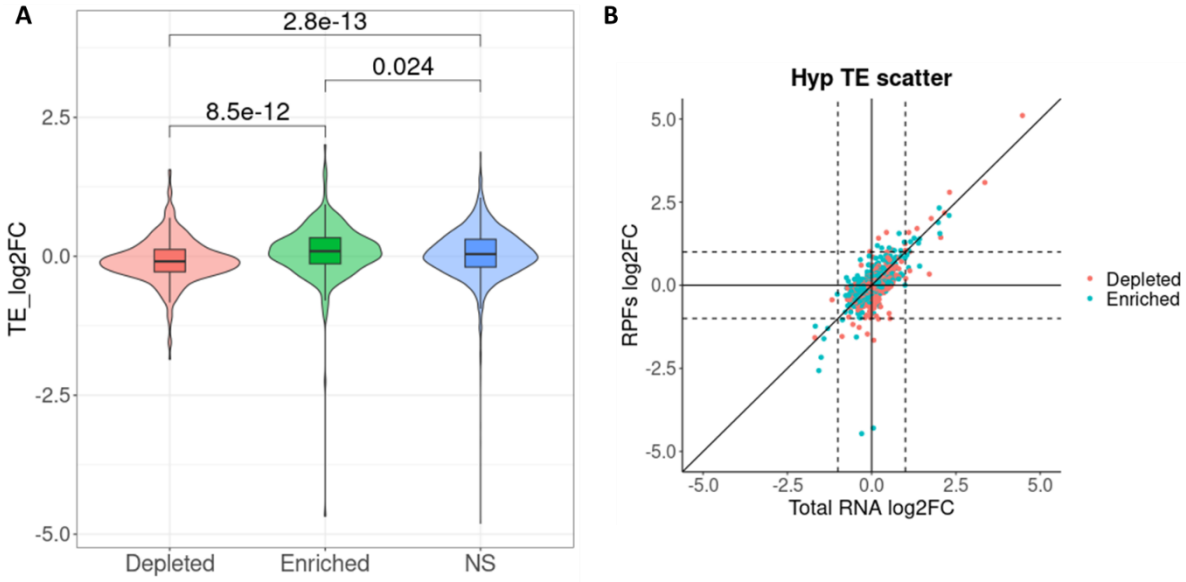


Figure 5.10 Hypoxia stress granule mRNAs are translationally upregulated

A) Violin plots showing distributions with boxplots representing minimum, maximum, interquartile range and median of TE log2 fold change. Groups represent stress granule depleted, enriched or non-significantly enriched or depleted (NS) mRNAs in response to hypoxia. P-values are from a Wilcoxon rank-sum test.

B) Scatter plot representing translation efficiency of stress granule enriched or depleted mRNAs in response to hypoxia. Similar to what was observed for sodium arsenite, stress granule localisation is not the sole determinant of TE as some stress granule enriched mRNAs are translationally repressed.

Stress granule enriched mRNAs were subsetted into two groups based on TE values for both sodium arsenite and hypoxia. These groups represent mRNAs that are actively translated (TE > 0, figure 5.11) and mRNAs that are translationally repressed (TE < 0, figure 5.12). Overlap between datasets was visualised by Venn diagrams (figures 5.11 A, 5.12 A). Pathway enrichment analysis was conducted using the enrichR R package (Kuleshov et al., 2016) against the GSEA hallmark pathways (Liberzon et al., 2015), using a p-value < 0.05 as a significance threshold.

327 mRNAs were identified as uniquely sodium arsenite stress granule enriched with positive TE values with 201 uniquely identified in hypoxia. There was an overlap of 29 mRNAs. A number of physiologically relevant genes implicated in cancer were identified as stress granule enriched and having positive TE values. NUAK family SNF1-like kinase 1 (NUAK1) was identified as stress granule enriched in response to sodium arsenite, but not translationally repressed. Previous work has shown NUAK1 to be required for colorectal cancer development by means of protection from oxidative stress (Port et al., 2018). NUAK1 also regulates cell cycle progression through governance of centrosome biogenesis in human pancreatic cancer cell lines (Whyte et al., 2023). Glycogen Synthase Kinase 3 Beta (GSK3β) was identified as stress granule enriched with positive TE values in response to hypoxia. GSK3β is known to be a key regulator of cancer invasion and chemotherapeutic resistance in numerous human cancers, including cancers of the lung, breast and glioblastomas (Domoto et al., 2016). VEGFA was identified as stress granule enriched in response to both sodium arsenite and hypoxia and had positive TE values in both conditions. VEGFA is a critical regulator of angiogenesis in cancer (Carmeliet, 2005). VEGFA expression is also known to be increased in response to both eIF2α phosphorylation (Ghosh et al., 2010) and hypoxia (Carmeliet, 2005).

Pathway enrichment analysis for sodium arsenite (figure 5.11 B) and hypoxia (figure 5.11 C) induced stress granule mRNAs with positive TE revealed distinct sets of pathways enriched in a stress dependent manner. Despite alternatively enriched pathways being discovered, there is still a convergence on pathways which are known critical regulators of cancer. In response to sodium arsenite stress, enriched pathways included TNF-alpha signalling via NF-κB (Hoesel and Schmid, 2013), p53 pathway (Levine, 2022) and KRAS

signalling (Uprety and Adjei, 2020). In response to hypoxia the enriched pathways identified were PI3K/AKT/ mTOR signalling (He et al., 2021) and Hedgehog signalling (Gupta et al., 2010).

415 mRNAs were uniquely identified as stress granule enriched and had negative TE values in response to sodium arsenite with 123 being identified in response to hypoxia. There was an overlap of 22 genes (figure 5.12 A). Positive regulators of cell death were identified as stress granule enriched and translationally repressed. BRI3BP was identified as stress granule enriched and translationally repressed in response to sodium arsenite stress. BRI3BP is known to augment the induction of apoptosis in response to the chemotherapeutic agent etoposide (Yamazaki et al., 2007). TP53INP1 was identified as stress granule enriched and translationally repressed in response to hypoxia. TP53INP1 is a known promoter of apoptosis through p53 dependent and independent mechanisms and exhibits supressed expression in multiple human tumours including pancreatic and colon (Shahbazi et al., 2013). Stress granules have previously been implicated in arresting cell cycle in response to sodium arsenite (Somasekharan et al., 2020) and UVC irradiation (Moutaoufik et al., 2014), however the direct mechanism remains uncharacterised. The E2F transcription factor 3 (E2F3) was stress granule enriched and translationally repressed in response to both sodium arsenite and hypoxia. E2F3 is known to regulate G1/S cell cycle transition in mammalian cells (Bertoli et al., 2013).

Pathway enrichment analysis for sodium arsenite (figure 5.12 B) and hypoxia (figure 5.12 C) enriched stress granule mRNAs with negative TE values revealed distinct sets of pathways being negatively regulated in a stress dependent manner. Similar to what was observed for positively regulated genes, a number of pivotal pathways in cancer progression were identified. In response to sodium arsenite these included TNF-alpha signalling via NF-κB (Hoesel and Schmid, 2013), epithelial mesenchymal transition (Ribatti et al., 2020) TGF-beta signalling (Syed, 2016) and Notch signalling (Nowell and Radtke, 2017). In response to hypoxia, enriched pathways included Inflammatory response (Hanahan, 2022) and the G2-M checkpoint (Löbrich and Jeggo, 2007).

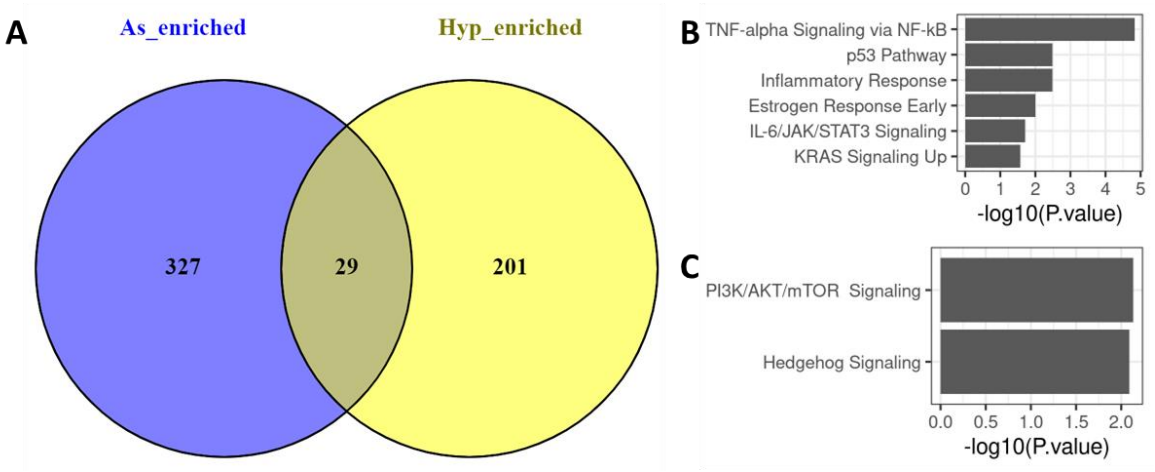


Figure 5.11 Overlap and pathway analysis of Sodium arsenite and hypoxia induced stress granule mRNAs with TE values > 0

A) Venn diagram of hypoxia enriched (Hyp) and sodium arsenite enriched (As) stress granule mRNAs with a TE value > 0.

B) EnrichR analysis of GSEA hallmark pathways of stress granule enriched mRNAs with positive TE values in response to sodium arsenite.

C) EnrichR analysis of GSEA hallmark pathways of stress granule enriched mRNAs with positive TE values in response to hypoxia.

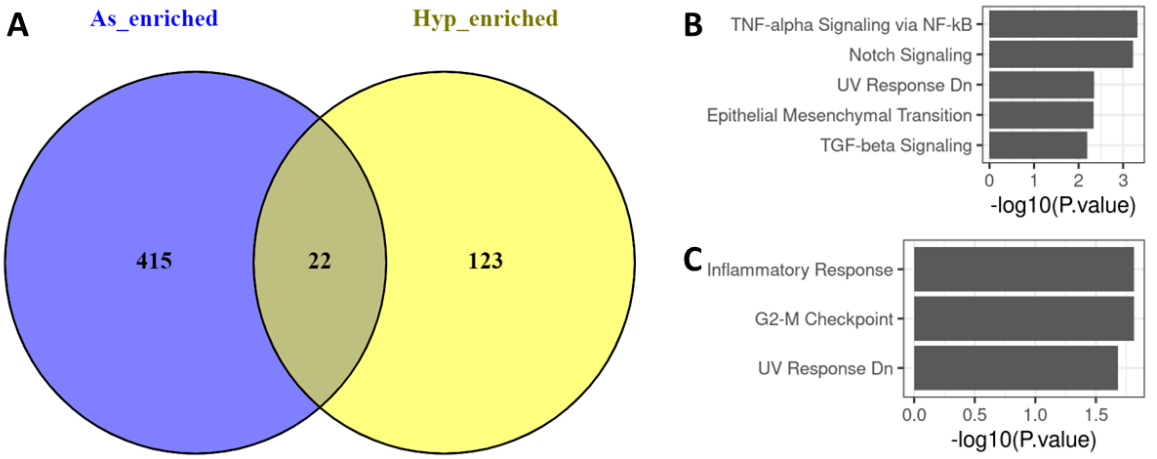


Figure 5.12 Overlap and pathway analysis of Sodium arsenite and hypoxia induced stress granule mRNAs with TE values < 0

A) Venn diagram of hypoxia enriched (Hyp) and sodium arsenite enriched (As) stress granule mRNAs with a TE value < 0.

B) EnrichR analysis of GSEA hallmark pathways of stress granule enriched mRNAs with negative TE values in response to sodium arsenite

C) EnrichR analysis of GSEA hallmark pathways of stress granule enriched mRNAs with negative TE values in response to hypoxia

As VEGFA is a stress granule enriched mRNA with positive TE values, that is of particular interest in the context of cancer, ribosome occupancy plots were generated to validate increased TE in both conditions. Counts from RPF libraries were first normalised to total RNA levels and converted to counts per million. CDS was then binned into 50 windows and UTRs were binned into 25 windows. In response to sodium arsenite and hypoxia there was an increased ribosome occupancy in the CDS, indicating that VEGFA is being translationally upregulated (figure 5.13). In addition to the CDS, there is also an increase in ribosome occupancy in the 5'UTR in both conditions. VEGFA has a complex 5'UTR with two IRESs and a uORF and the exact mechanisms of translation regulation of this mRNA are still being elucidated (Arcondéguy et al., 2013). In Mateju et al the authors showed that a reporter mRNA with the ATF4 5' UTR which contains uORFs was translated in stress granules (Mateju et al., 2020). These data therefore may suggest that uORF translation may be involved in regulating stress granule translation however further validation is required to understand the exact mechanisms.

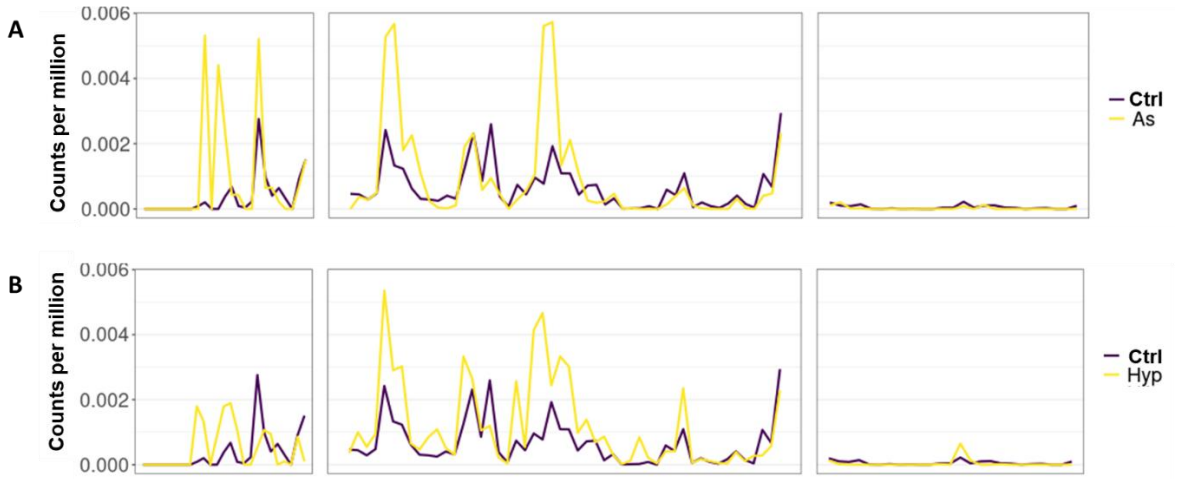


Figure 5.13 VEGFA has increased ribosome occupancy in response to both sodium arsenite and hypoxia

A) Ribosome occupancy plots of VEGFA mRNA in non-stressed control and sodium arsenite treated conditions. CDS is binned into 100 windows and UTRs are binned into 50 windows. An increase in ribosome occupancy is observed in the CDS and 5'UTR.

B) Ribosome occupancy plots of VEGFA mRNA in non-stressed control and hypoxia treated conditions. CDS is binned into 100 windows and UTRs are binned into 50 windows. An increase in ribosome occupancy is observed in the CDS and 5'UTR.

5.5 Alternative mechanisms of translation initiation.

Although stress granules play a pivotal role in regulating translation in response to stress, it is important to keep in mind that they are not likely to be the only regulatory mechanism. As stress granules contain approximately 10% of cellular mRNA in response to sodium arsenite stress (Khong et al., 2017), it is likely that they only regulate specific subsets of stress responsive mRNAs. Previous studies identified a hypoxia specific alternative mRNA cap binding complex in response to reduced mTOR activity (Uniacke et al., 2012, Melanson et al., 2017). This complex includes eIF4A1, in addition to eIF4E2 and eIF4G3. Another recent study characterised the ability of eIF3D to bind the mRNA cap and stimulate the translation of pro-survival genes in response to glucose deprivation (Lamper et al., 2020). The authors of this study showed regulation of mTOR kinase activity through selective eIF3D regulated translation in response to nutrient deprivation. eIF3D cap binding was stimulated by reduced CK2 kinase and loss of eIF3D phosphorylation. CK2 has also previously been shown to phosphorylate G3BP1 and regulate stress granule dynamics (Reineke et al., 2017). This poses another interesting question as to whether eIF3D may also be capable of regulating translation in hypoxia, as similar signalling cascades were observed in response to nutrient deprivation.

In order to assess CK2 kinase activity, U2OS cells were subjected to sodium arsenite or hypoxia stress as described above. Non stressed cells were used as a control. CK2 kinase activity was measured by western blot, using a canonical readout of CK2 kinase activity, phosphorylation of AKT at serine 129 (figure 5.14). There was a reduction of AKT phosphorylation specifically in response to hypoxia stress when compared to the control or sodium arsenite treated cells. 191 eIF3D targets in response to glucose deprivation were identified in (Lamper et al., 2020). TE $\log_2$ fold change of these eIF3D targets was compared to the TE of all other mRNAs in response to sodium arsenite (figure 5.15) and hypoxia stress (figure 5.16).

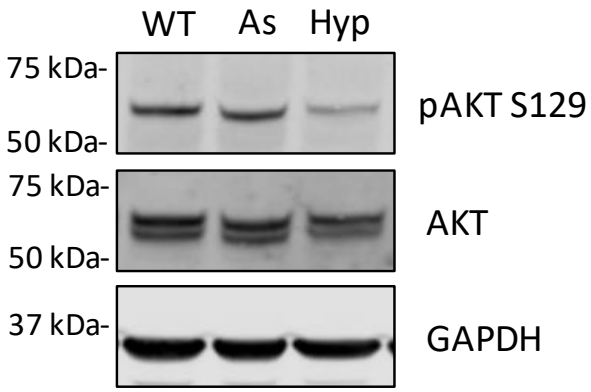


Figure 5.14 CK2 kinase activity is reduced in hypoxia
Western blot of U2OS lysates from non-stressed control (Ctrl), sodium arsenite (As) or hypoxia (Hyp) treated samples. Reduction of AKT phosphorylation at S129 is indicative of reduced CK2 kinase activity. Equal volumes of cell lysate loaded. Representative image of 3 biological replicates.

In response to sodium arsenite stress, eIF3D targets had a lower TE values compared to all other mRNAs (figure 5.14). This is in agreement with the CK2 kinase activity in figure 5.13. In contrast to this, eIF3D targets had higher TE values compared to all other mRNAs in hypoxia (figure 5.15). These findings implicate eIF3D selective translation as being upregulated in hypoxia in addition to what has previously been described for nutrient deprivation. This is of particular interest as hypoxia is a common physiological stress in numerous human diseases, including various cancers (Chen et al., 2020).

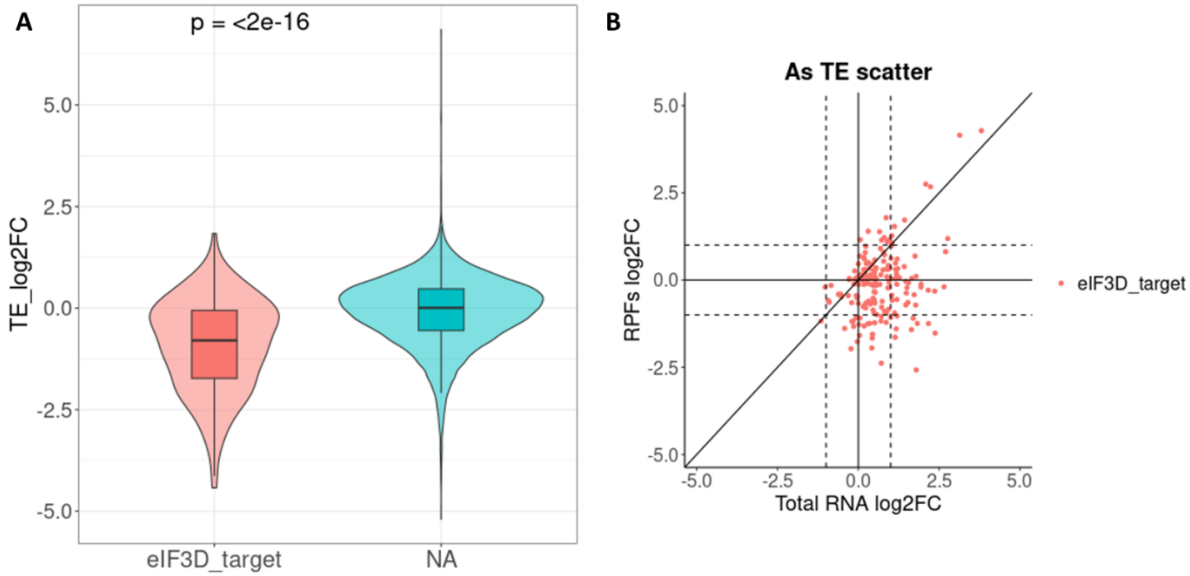


Figure 5.15 eIF3D targets are translationally downregulated in response to sodium arsenite

A) Violin plots showing distributions with boxplots representing minimum, maximum, interquartile range and median of TE log2 fold change in response to sodium arsenite. Groups represent eIF3D targets identified in (Lamper et al., 2020) or all other mRNAs (NA) P-values are from a Wilcoxon rank-sum test.

B) Scatter plot representing translation efficiency of eIF3D targets in response to sodium arsenite treatment

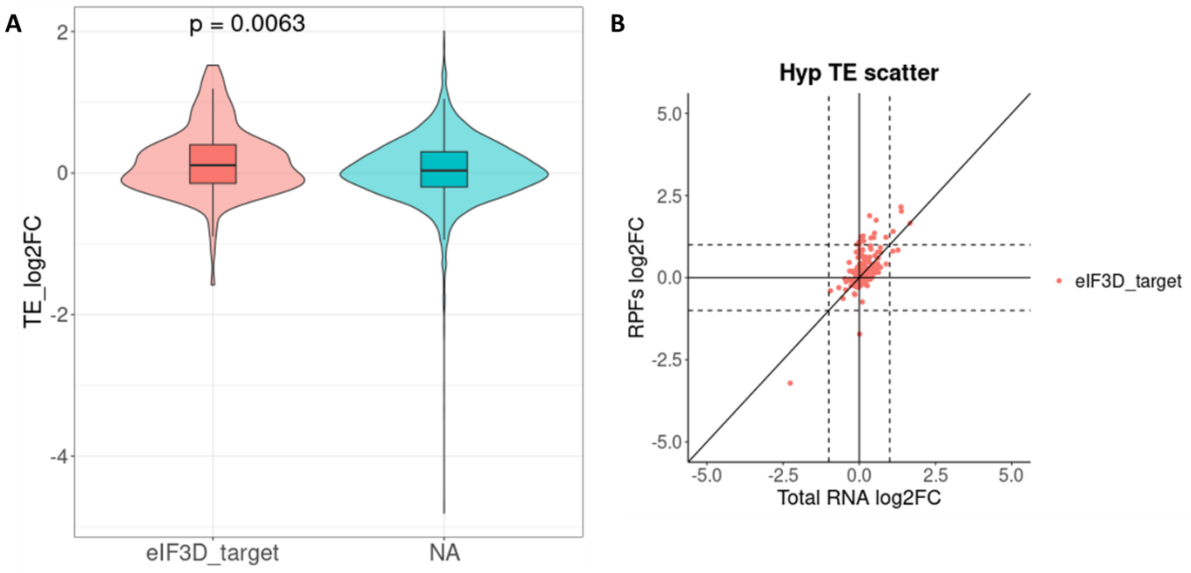


Figure 5.16 eIF3D targets are translationally upregulated in response to hypoxia

A) Violin plots showing distributions with boxplots representing minimum, maximum, interquartile range and median of TE log2 fold change in response to hypoxia. Groups represent eIF3D targets identified in (Lamper et al., 2020) or all other mRNAs (NA) P-values are from a Wilcoxon rank-sum test.

B) Scatter plot representing translation efficiency of eIF3D targets in response to hypoxia.

5.6 Chapter discussion.

Polysome sucrose gradient plots revealed strong translation inhibition in response to sodium arsenite treatment, as previously described (McEwen et al., 2005) (figure 5.1 B). Bulk protein synthesis rates in response to hypoxia were comparable to the control (figure 5.1 B, 5.2). Although this result is surprising in the context of stress granule formation, stress granules have been shown to form without major translation inhibition independently of eIF2α phosphorylation (Moutaoufik et al., 2014). In order to understand stress specific changes in translation a ribosome profiling approach was used (figures 5.3). Ribosome profiling libraries of high quality, with fragment length of ~30 nt (figure 5.5 B-D), strong positional enrichment within the CDS of mRNAs (figures 5.5 B-D, 5.6 B) and tri-nucleotide periodicity were generated (figure 6 B). PCA analysis revealed distinct patterns of gene expression at the level of total mRNA and ribosome occupancy in response to sodium arsenite and hypoxia induced stress (figure 5.7).

Integration of ribosome profiling data with stress granule RNAseq showed that, on average, stress granule mRNAs in response to sodium arsenite are translationally repressed (figure 5.9), in agreement with previous studies (Khong et al., 2017, Namkoong et al., 2018, Padrón et al., 2019). In contrast, hypoxia stress granule mRNAs, on average, were translationally upregulated (figure 5.10). These findings support the hypothesis developed in chapter 3, that hypoxia induced stress granule represent a more translationally active sub-cellular compartment due to localisation of large ribosomal subunit proteins. In (Padrón et al., 2019) the authors showed that stress granule mRNAs in response to inhibition of eIF4A1 with hippuristanol were not translationally repressed. These data together suggest that translational repression of stress granule mRNAs may be a feature in response to eIF2α phosphorylation but is not conserved in response to eIF4F complex perturbation.

Looking at average TE of gene sets is a useful tool for describing general trends, however, individual mRNA resolution is lost. Both sodium arsenite and hypoxia induced stress granule mRNAs could be further subsetted into groups of mRNAs with positive (figure 5.11) and negative TE values (figure 5.12). These findings support a model of selective partitioning of stress granule mRNAs and coordination of pro survival gene expression programmes,

similar to what was described for G3BP1 associated mRNAs in (Somasekharan et al., 2020). Categorising all stress granule mRNAs as translationally repressed is an over simplified, reductionist dogma that needs to be challenged. Stress granule enriched mRNAs with positive TE values converged on pathways and individual mRNAs synonymous with cancer progression (figures 5.11-12). An example of potential therapeutic significance was the localisation of VEGFA to stress granules coupled with its increased TE in response to both sodium arsenite and hypoxia (figure5.13). Anti angiogenic chemotherapeutics are a commonly used treatment strategy in human cancers. A challenge faced is development of resistance and activation of alternative pro-angiogenic cascades (Itatani et al., 2018). Stress granules represent a potential indirect druggable axis that may also be able to combat tumour angiogenesis. The sequestration and subsequent downregulation of the G1/S transition regulator E2F3 may also explain the link to stress granules and cell cycle arrest.

As translation is one of the most energetically exhaustive cellular processes (Buttgereit and Brand, 1995), it is reasonable to assume that multiple modalities of regulation will simultaneously coordinate stress responsive protein synthesis. eIF3D has previously been shown as an alternative mRNA cap binding protein (Lee et al., 2016) and regulator of stress responsive translation initiation in conditions of nutrient depletion (Lamper et al., 2020). In response to hypoxia there was a reduction in CK2 kinase activity (figure 5.14) which is known to stimulate eIF3D mRNA cap binding. eIF3D targets identified in (Lamper et al., 2020) showed increased TE specifically in response to hypoxia when compared with all other mRNAs (figures 5.15-5.16). These findings implicate eIF3D selective translation in response to hypoxia in addition to what has already been described in response to glucose deprivation.

In conclusion, the two primary aims of this chapter have been fulfilled. High quality ribosome profiling libraries were generated and this data was successfully integrated with the stress granule RNAseq from the previous chapter. The work presented in this thesis shows that stress granule morphology, protein and mRNA content is stress dependent. In addition, stress granule morphology is regulated by protein content and stress granule mRNAs show unique sequence features in response to sodium arsenite and hypoxia stress. Hypoxia stress granule mRNAs are translationally upregulated, in contrast to sodium arsenite stress granule mRNAs which are translationally repressed. Stress granules

coordinate pro-survival translation responses and represent a therapeutic opportunity in combatting human cancers.

6 Chapter 6: Final discussion

The aim of this thesis was to gain a better understanding of how stress granules regulate translation in response to physiological stresses relevant to cancer. Stress granules have been implicated in many aspects of cancer, including disease progression and response to treatment (section 1.3). Stress granules potentially represent an under-utilised druggable axis that may improve patient outcomes. To date, knowledge of the functions of stress granules remains limited. In order to build a model of stress granule mediated translational control, multiple high throughput methodologies were used to simultaneously characterise the stress granule proteome and mRNA content. Ribosome profiling was used to understand the stress responsive gene expression programmes. The key experimental findings and future perspectives will be outlined in this chapter.

6.1 Isolation and proteomic analysis of stress granule cores

The aims of this chapter were:

1. Develop methodologies for stress granule induction.

2. Develop a methodology for isolation of stress granule cores to enable proteomic characterisation of stress granules.

Stress granules with distinct morphologies form in response to sodium arsenite and hypoxia stress (figure 3.1-2). Larger poorly defined stress granules formed in response to hypoxia which had no docking P-bodies. Sodium arsenite stress granule formation was regulated by phosphorylation of eIF2α whereas hypoxia induced stress granule formation was regulated by reduced mTOR activity and hypophosphorylation of 4E-BP1, preventing eIF4F complex assembly (figure 3.3-3.4). Both sodium arsenite and hypoxia induced stress granules contain eIF3B and are therefore considered cytoprotective as pro-apoptotic stress granules lack eIF3B (Aulas et al., 2018) (figure 3.5). Methods developed in (Jain et al., 2016) were utilised to isolate stress granule cores (figure 3.7-8). Stress granule proteomics yielded a number of interesting findings. Firstly, the identification of the DEAD box helicases eIF4A1 and DDX6

were identified as being differentially enriched in hypoxia and sodium arsenite induced stress granules respectively (figure 3.15-3.17). DEAD box helicases, including eIF4A1 and DDX6 are known regulators of liquid-liquid phase separation. eIF4A1 is believed to limit RNA-RNA mediated interactions and thus liquid-liquid phase separation (Tauber et al., 2020), whereas DDX6 promotes phase separation (Hondele et al., 2019, Ayache et al., 2015). DDX6 is known to be a member of the mRNA decapping complex, and two other complex members (EDC4 and DCP1A) were also identified as enriched in sodium arsenite induced stress granules. Treatment with the eIF4A1 inhibitor, hippuristanol, resulted in reduced stress granule size and induced P-body docking on stress granules (figure 3.19). This led to the development of a hypothesis that eIF4A1 and DDX6 may regulated the stress granule P-body docking interaction, with eIF4A1 limiting and DDX6 (potentially as part of the decapping complex) promoting P-body docking. eIF4A1 may also be altering granule morphology by preventing the formation of stress granule cores and supporting the creation of a more liquid-like compartment. Multiple complimentary techniques showed that eIF4A1 and G3BP1 interact with each other *in cellulo* and *in vitro* potentially through their opposingly charged IDRs (figure 3.24-26). This may also represent a very physiologically relevant interaction, as both eIF4A1 and G3BP1 are overexpressed in lung adenocarcinoma (figure 3.21). There is a positive correlation between their expression (figure 3.22) and high expression of both represents the worst clinical outcomes in lung adenocarcinoma patients (figure 3.23).

Future work stemming from this chapter will focus on *in vitro* biochemical work better characterising the G3BP1 eIF4A1 interaction. Truncated mutants will allow for the identification of the interacting residues. *In vitro* interaction work will then be validated by expressing the truncated mutants in cells. Another critical next step is using pure *in vitro* systems to understand how the eIF4A1 G3BP1 interactions regulate liquid-liquid phase separation. Using fluorescently tagged RNAs, imaging methods will be used as described in (Tauber et al., 2020) to achieve this. Although no evidence presented supports a direct interaction between DDX6 and G3BP1, DDX6 contains an IDR that may facilitate a direct interaction with G3BP1. Therefore, similar approaches can be used to test whether DDX6 and G3BP1 interact, and how this interaction alters liquid-liquid phase separation dynamics.

Stress granule proteomics also highlighted the enrichment of members of the large ribosomal subunit as enriched in hypoxia (figure 3.15-18). Recent evidence shows that translation can occur within a stress granule (Mateju et al., 2020). This observation, in conjunction with the localisation of eIF4A1 which may be limiting RNA-RNA mediated interactions led to the hypothesis that hypoxia induced stress granules may represent a more dynamic translationally active compartment than what has previously described in response to sodium arsenite.

Limitations of the work presented in this chapter include a lack orthogonal techniques used to validate mass spectrometry results and lack of quantitation of microscopy images. Using immunofluorescence approaches or expression of fluorescently tagged constructs could be used to validate localisation of enriched proteins to stress granules. This would allow the quantitation of the proportion of the protein that localises to stress granules in response to sodium arsenite or hypoxia. In tandem, metrics such as stress granule size, number of cells with stress granules etc could be analysed to improve the robustness and informativeness of the observations highlighted in this chapter.

6.2 Isolation and RNAseq analysis of stress granule cores

The aims of this chapter were:

1. Adapt the stress granule core isolation methodology to enable RNAseq analysis of stress granule mRNAs.

2. Characterise stress granule mRNAs and their sequence features in sodium arsenite and hypoxia induced stress granules.

The stress granule isolation methodology was successfully used to generate libraries that were suitable for RNAseq (figure 4.1-3). This enabled the identification of stress granule mRNAs in response to both sodium arsenite and hypoxia (figure 4.5-7). It is important to

note that a high level of overlap was observed between the mRNAs identified in the control conditions and in both stress conditions. Indeed, the mRNAs identified in all conditions, including the controls, had a more statistically significant overlap with the previously published stress granule transcriptome (figure 4.6-4.9). These results show the absolute necessity for adequately controlling RNAseq based experiments to account for technique biases. In order to potentially reduce the levels of background observed, a CRISPR approach was used to endogenously tag G3BP1 with GFP and APEX figure (figure 4.16-19).

Analysis of the stress granule RNAseq data revealed that mRNAs with distinct sequence features localise to stress granules in response to hypoxia and sodium arsenite. In response to sodium arsenite, stress granule mRNAs had a GC rich sequence bias (figure 4.12, figure 4.15). In response to hypoxia stress granule mRNAs had an AU rich sequence bias (figure 4.14-15). Alternate sequence composition of stress granule mRNAs suggests that different RNA binding proteins may be regulating localisation to stress granules. Using publicly available eCLIP datasets may enable identification of specific mRNA motifs associated with RNA binding proteins. If specific RNA binding protein motifs were identified as differentially enriched an immunofluorescence approach could be used to see if the RNA binding proteins were differentially enriched in stress granules. Furthermore, the Bushell lab have generated eCLIP datasets for both eIF4A1 and DDX6. These data could be combined with stress granule RNA sequencing to understand whether there is an enrichment in motifs associated with either eIF4A1 or DDX6 in hypoxia and sodium arsenite stress granule mRNAs respectively. Both hypoxia and sodium arsenite induced stress granule mRNAs were longer than stress granule depleted mRNAs (figure 4.11-12). mRNA length has been identified as a determining feature of stress granule localisation in this thesis and multiple independent studies (Namkoong et al., 2018, Khong et al., 2017, Padrón et al., 2019). This is believed to contribute to stress granule assembly, as longer mRNAs can provide more nucleotides for protein-RNA or RNA-RNA mediated interactions (Khong et al., 2022). The alternate sequence features identified in addition to the identification of stress specific localisation of mRNAs suggests that localisation of an mRNA to stress granules is a stress specific response and is tailored depending on the stress applied. The stress granule mRNA datasets provided the framework for understanding how stress granule localisation influences translation rates.

As with the previous chapter, limitations of the work presented include a lack of orthogonal technique validation of omic results. smFISH could be used to validate candidate mRNAs localising to stress granules. In addition, a machine learning approach could be used to identify further sequence features of stress granule mRNAs in a more unbiased approach (Gillen et al., 2021). This could potentially enable the design of a synthetic prototypical stress granule mRNA for sodium arsenite and hypoxia, which in turn could be validated by an imaging approach. An additional limitation is that the analysis in this chapter is purely focussed on identifying differentially enriched stress granule mRNAs compared to total mRNA. The analysis needs to be expanded to understand what differences there at the level of total mRNA in each of the various conditions. Understanding the transcriptomic landscape may further highlight other biomolecular processes such as transcription (i.e. are newly transcribed mRNAs localising to stress granules) or RNA degradation (i.e. are stress granule mRNAs protected from degradation). The analysis should also be expanded to include non-coding RNAs to understand whether certain non-coding RNAs are differentially enriched in stress granules in response to either sodium arsenite or hypoxia and whether they play a role in stress granule biology.

6.3 Ribosome profiling in sodium arsenite and hypoxia induced stress

The aims of this chapter were:

1. Generate high quality ribosome profiling libraries in conditions of sodium arsenite and hypoxia induced stress.

2. Integrate ribosome profiling data with stress granule RNAseq data to understand how stress granule localisation regulates translation.

High quality ribosome profiling libraries were generated with fragments of the correct length, with CDS enrichment and showing high levels of trinucleotide periodicity (figure 5.5-6). The hypothesis developed at the end of chapter 3 and discussed in section 6.1 of this chapter, that hypoxia induced stress granules represent a more translationally active compartment is supported by the ribosome profiling data. In response to sodium arsenite stress granule mRNAs were translationally repressed (figure 5.9). This is in concordance with other studies that have shown in response to eIF2α phosphorylation stress granule localised mRNAs are translationally repressed (Namkoong et al., 2018, Khong et al., 2017,

Padrón et al., 2019). In response to hypoxia stress granule localised mRNAs were translationally upregulated, and on average showed increased TE values compared to stress granule depleted or non-significantly enriched or depleted mRNAs (figure 5.10). It has also previously been shown in response to eIF4A1 inhibition by hippuristanol that stress granule mRNAs were not significantly downregulated translationally compared to stress granule depleted or non-significantly enriched or depleted mRNAs. This leads to the question as to whether translational repression of stress granule mRNAs is eIF2α phosphorylation dependent and not conserved for other stresses. Further work inducing stress granule formation independently of eIF2α phosphorylation is required to better understand this.

Increased TE values and localisation of the large ribosomal subunit proteins support that hypoxia induced stress granules are a translationally active compartment however the evidence is not unequivocal. A number of complementary imaging approaches will be used to address this. Using cryo-electron tomography will potentially answer a number of questions simultaneously. This will allow for characterisation of stress granule substructure on a previously unprecedented scale. This will show whether eIF4A1 is supporting the creation of a more liquid compartment by preventing stress granule core formation in hypoxia. Recent advances in this technique have enabled structural characterisation of translating ribosomes *in cellulo* (Hoffmann et al., 2022). Obtaining structural evidence as to whether elongating ribosomes are present in greater quantities in response to hypoxia is a key piece of evidence needed to support this hypothesis. It is also a possibility that stress granule enriched mRNAs are engaged in translation outside of stress granules. smFISH will be used to validate the stress granule sequencing data. It will also enable more accurate quantitation of what proportion of the total pool of an mRNA is localised to stress granules. Similar methods to those described in (Mateju et al., 2020) can then be used on validated targets to image translation in real time and better understand the relationship between stress granule localisation and translation rates.

6.4 Understanding the disease relevance of stress granules

The great endeavour of cancer research is to try and improve patient outcomes. Whilst there is mounting evidence supporting stress granules as having a role in cancer pathophysiology, work within the field needs to move beyond observational studies to functional studies to understand if stress granules can be exploited therapeutically. VEGFA and E2F3, which localised to stress granules in response to both sodium arsenite and hypoxia represent two interesting candidate mRNAs for follow up future work. Despite stress granule mRNAs having apparent alternate fates of translational regulation (i.e., downregulated in response to sodium arsenite and upregulated in response to hypoxia) these two mRNAs belonged to a subset where their translational regulation was conserved between the two conditions (figure 5.11-12). In response to both stresses, VEGFA localised to stress granules and had positive TE values (figure 5.13). VEGFA has been implicated in the progression of cancer by promoting angiogenesis. VEGFA expression is known to increase in response to hypoxia and the integrated stress response (section 1.4). In contrast, E2F3 localised to stress granules in both conditions but had negative TE values. E2F3 is a transcription actor that promotes cell cycle progression from G1 to S phase. Multiple studies have implicated stress granules as regulators of cell cycle, with accumulation of cells in G1 phase of cell cycle in response to sodium arsenite (Somasekharan et al., 2020) and UVC (Moutaoufik et al., 2014) stress. The model proposed is that stress granules arrest cell cycle to allow the cell to overcome the stress applied before progressing through cell cycle, thus protecting the cell in conditions not favourable for cell division (Moutaoufik et al., 2014). Irrespective of the stress applied, perhaps the function of stress granules is to promote the translation of stress responsive mRNAs such as VEGFA whilst simultaneously downregulating mRNAs that may negatively impact cell survival. This concept of selective partitioning has already been shown for G3BP1 bound mRNAs in (Somasekharan et al., 2020) and is a fascinating concept. A more in-depth analysis of the mRNAs identified in this thesis may provide insight as to whether they contain sequence regulatory elements that allows for this differential regulation upon stress granule localisation.

In order to truly understand the contribution of stress granules to cancer a transition to in vivo work is required. Multiplex immunohistochemistry is now being used to simultaneously assess stress granule formation and the upstream signalling events regulating formation in

human lung adenocarcinoma tissue samples. This will be combined with markers of proliferation and cell death to try to understand how stress granules are impacting disease progression. Understanding what signalling events lead to stress granule formation in a diseased tissue and assessing granule morphology will elucidate which state of stress granule formation is disease relevant. However, in the complex environment of a tissue, it is likely that a combination of both will be observed. As discussed in (Riggs et al., 2020) the field is lacking suitable in vivo disease models to truly understand the contribution of stress granules. Development of appropriate models would allow in depth *in vivo* characterisation of stress granules and serve as a means of testing compounds that alter stress granule dynamics as potential therapeutics. Furthermore, numerous compounds targeting translation initiation are currently of therapeutic interest in cancer. These compounds are often looked at over simplistically as just dysregulating bulk translation rates. In addition to this function, they also have the potential of disrupting cellular organisation by altering stress granule dynamics. The contribution of altering stress granule dynamics is yet to be fully understood as a therapeutic strategy.

Appendix

Table 3 Primary Antibodies

Antibody	Supplier	Product Number	Use
peIF2α Ser 51	Cell Signalling Technology	119A11	Western blot
eIF2α	Cell Signalling Technology	9722	Western blot
p4E-BP1	Cell Signalling Technology	2855	Western blot
4E-BP1	Cell Signalling Technology	9644	Western blot
pRPS6 Ser 240/244	Cell Signalling Technology	2215	Western blot
RPS6	Cell Signalling Technology	2217s	Western blot
EDC3	Abcam	ab168851	Immunofluoresence
eIF4A1	Abcam	ab31217	Western blot, Immunoprecipitation
eIF4G	Cell Signalling Technology	2498S	Western blot
eIF4E	Invitrogen	MA1-089	Western blot
eIF3B	Santa Cruz Biotechnology	sc-137214	Immunofluoresence
GFP	Santa Cruz Biotechnology	sc-9996	Western blot
PABP	Abcam	ab21060	Western blot
GAPDH	Cell Signalling Technology	5174S	Western blot
RPL11	Cell Signalling Technology	18163S	Western blot
G3BP1	Proteintech	13057-2-AP	Western blot
turbo GFP	Evrogen	AB513	Western blot
Vinculin	Abcam	ab129002	Western blot
TIA1	Abcam	ab40693	Immunofluoresence
pAKT Ser 129	Abcam	ab133458	Western blot
AKT	Cell Signalling Technology	9272	Western blot
6X His	Abcam	ab137839	Immunoprecipitation
Rabbit IGG	Cell Signalling Technology	2729S	Immunoprecipitation

Table 4 Secondary Antibodies

Antibody	Supplier	Product Number	Use
Goat Anti-Mouse IgG Polyclonal Antibody (IRDye® 680RD)	Li-Cor	926-68070	Western blot
Goat Anti-Rabbit IgG Antibody (IRDye 680RD)	Li-Cor	926-68071	Western blot
Donkey Anti-Mouse IgG Polyclonal Antibody (IRDye® 800CW)	Li-Cor	926-32212	Western blot
Donkey Anti-Rabbit IgG Polyclonal Antibody (IRDye® 800CW)	Li-Cor	926-32213	Western blot
Goat Anti-Rabbit IgG H&L (Alexa Fluor® 568)	Abcam	ab175471	Immunofuoresence

References

ACKERMAN, D. & SIMON, M. C. 2014. Hypoxia, lipids, and cancer: surviving the harsh tumor microenvironment. *Trends in Cell Biology,* 24**,** 472-478.

ADJIBADE, P., ST-SAUVEUR, V. G., QUEVILLON-HUBERDEAU, M., FOURNIER, M. J., SAVARD, A., COUDERT, L., KHANDJIAN, E. W. & MAZROUI, R. 2015. Sorafenib, a multikinase inhibitor, induces formation of stress granules in hepatocarcinoma cells. *Oncotarget,* 6.

ALKALAEVA, E. Z., PISAREV, A. V., FROLOVA, L. Y., KISSELEV, L. L. & PESTOVA, T. V. 2006. In Vitro Reconstitution of Eukaryotic Translation Reveals Cooperativity between Release Factors eRF1 and eRF3. *Cell,* 125**,** 1125-1136.

ANAND, A. A. & WALTER, P. 2020. Structural insights into ISRIB, a memory-enhancing inhibitor of the integrated stress response. *Febs j,* 287**,** 239-245.

ANDERS, M., CHELYSHEVA, I., GOEBEL, I., TRENKNER, T., ZHOU, J., MAO, Y., VERZINI, S., QIAN, S. B. & IGNATOVA, Z. 2018. Dynamic m(6)A methylation facilitates mRNA triaging to stress granules. *Life Sci Alliance,* 1**,** e201800113.

ANDERSEN, G. R., PEDERSEN, L., VALENTE, L., CHATTERJEE, I., KINZY, T. G., KJELDGAARD, M. & NYBORG, J. 2000. Structural Basis for Nucleotide Exchange and Competition with tRNA in the Yeast Elongation Factor Complex eEF1A:eEF1Bα. *Molecular Cell,* 6**,** 1261-1266.

ANDERSON, P. & KEDERSHA, N. 2002. Stressful initiations. *J Cell Sci,* 115**,** 3227-34.

ANDERSON, P. & KEDERSHA, N. 2006. RNA granules. *J Cell Biol,* 172**,** 803-8.

ANDERSON, P. & KEDERSHA, N. 2009. RNA granules: post-transcriptional and epigenetic modulators of gene expression. *Nat Rev Mol Cell Biol,* 10**,** 430-6.

ANDREI, M. A., INGELFINGER, D., HEINTZMANN, R., ACHSEL, T., RIVERA-POMAR, R. & LÜHRMANN, R. 2005. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *Rna,* 11**,** 717-27.

ANGER, A. M., ARMACHE, J. P., BERNINGHAUSEN, O., HABECK, M., SUBKLEWE, M., WILSON, D. N. & BECKMANN, R. 2013. Structures of the human and Drosophila 80S ribosome. *Nature,* 497**,** 80-5.

ARCONDÉGUY, T., LACAZETTE, E., MILLEVOI, S., PRATS, H. & TOURIOL, C. 2013. VEGF-A mRNA processing, stability and translation: a paradigm for intricate regulation of gene expression at the post-transcriptional level. *Nucleic Acids Res,* 41**,** 7997-8010.

ARIMOTO, K., FUKUDA, H., IMAJOH-OHMI, S., SAITO, H. & TAKEKAWA, M. 2008. Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nat Cell Biol,* 10**,** 1324-32.

ASH, P. E. A., VANDERWEYDE, T. E., YOUMANS, K. L., APICCO, D. J. & WOLOZIN, B. 2014. Pathological stress granules in Alzheimer's disease. *Brain Research,* 1584**,** 52-58.

AULAS, A., LYONS, S. M., FAY, M. M., ANDERSON, P. & IVANOV, P. 2018. Nitric oxide triggers the assembly of "type II" stress granules linked to decreased cell viability. *Cell Death Dis,* 9**,** 1129.

AYACHE, J., BÉNARD, M., ERNOULT-LANGE, M., MINSHALL, N., STANDART, N., KRESS, M. & WEIL, D. 2015. P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. *Mol Biol Cell,* 26**,** 2579-95.

AYLETT, C. H. S., BOEHRINGER, D., ERZBERGER, J. P., SCHAEFER, T. & BAN, N. 2015. Structure of a Yeast 40S–eIF1–eIF1A–eIF3–eIF3j initiation complex. *Nature Structural & Molecular Biology,* 22**,** 269-271.

BABU, M. M. 2016. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem Soc Trans,* 44**,** 1185-1200.

BAILEY, T. L. 2021. STREME: accurate and versatile sequence motif discovery. *Bioinformatics,* 37**,** 2834-2840.

BAN, N., NISSEN, P., HANSEN, J., MOORE, P. B. & STEITZ, T. A. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science,* 289**,** 905-920.

BANANI, S. F., LEE, H. O., HYMAN, A. A. & ROSEN, M. K. 2017. Biomolecular condensates: organizers of cellular biochemistry. *Nature Reviews Molecular Cell Biology,* 18**,** 285-298.

BASSELL, G. J. & WARREN, S. T. 2008. Fragile X Syndrome: Loss of Local mRNA Regulation Alters Synaptic Development and Function. *Neuron,* 60**,** 201-214.

BERTOLI, C., SKOTHEIM, J. M. & DE BRUIN, R. A. 2013. Control of cell cycle transcription during G1 and S phases. *Nat Rev Mol Cell Biol,* 14**,** 518-28.

BOHLEN, J., ROIUK, M. & TELEMAN, A. A. 2021. Phosphorylation of ribosomal protein S6 differentially affects mRNA translation based on ORF length. *Nucleic Acids Research,* 49**,** 13062-13074.

BRANGWYNNE, C. P., ECKMANN, C. R., COURSON, D. S., RYBARSKA, A., HOEGE, C., GHARAKHANI, J., JÜLICHER, F. & HYMAN, A. A. 2009. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science,* 324**,** 1729-32.

BRANGWYNNE, C. P., MITCHISON, T. J. & HYMAN, A. A. 2011. Active liquid-like behavior of nucleoli determines their size and shape in Xenopus laevis oocytes. *Proc Natl Acad Sci U S A,* 108**,** 4334-9.

BRODY, E. & ABELSON, J. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science,* 228**,** 963-7.

BROWN, R. H. & BENNETT, J. J. Observations On The Organs And Mode Of Fecundation In Orchideae And Asclepiadeae. 2012.

BUCHAN, J. R., KOLAITIS, R. M., TAYLOR, J. P. & PARKER, R. 2013. Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell,* 153**,** 1461-74.

BUNDRED, J. R., HENDRIX, E. & COLEMAN, M. L. 2018. The emerging roles of ribosomal histidyl hydroxylases in cell biology, physiology and disease. *Cellular and Molecular Life Sciences,* 75**,** 4093-4105.

BURSET, M., SELEDTSOV, I. A. & SOLOVYEV, V. V. 2000. Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res,* 28**,** 4364-75.

BUSHELL, M., STONELEY, M., KONG, Y. W., HAMILTON, T. L., SPRIGGS, K. A., DOBBYN, H. C., QIN, X., SARNOW, P. & WILLIS, A. E. 2006. Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis. *Molecular Cell,* 23**,** 401-412.

BUTLER, J. E. & KADONAGA, J. T. 2002. The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev,* 16**,** 2583-92.

BUTTGEREIT, F. & BRAND, M. D. 1995. A hierarchy of ATP-consuming processes in mammalian cells. *Biochem J,* 312 (Pt 1)**,** 163-7.

CARMELIET, P. 2005. VEGF as a key mediator of angiogenesis in cancer. *Oncology,* 69 Suppl 3**,** 4-10.

CERAMI, E., GAO, J., DOGRUSOZ, U., GROSS, B. E., SUMER, S. O., AKSOY, B. A., JACOBSEN, A., BYRNE, C. J., HEUER, M. L., LARSSON, E., ANTIPIN, Y., REVA, B., GOLDBERG, A. P., SANDER, C. & SCHULTZ, N. 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov,* 2**,** 401-4.

CHANDRASHEKAR, D. S., BASHEL, B., BALASUBRAMANYA, S. A. H., CREIGHTON, C. J., PONCE-RODRIGUEZ, I., CHAKRAVARTHI, B. &

VARAMBALLY, S. 2017. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia,* 19**,** 649-658.
CHEN, P.-S., CHIU, W.-T., HSU, P.-L., LIN, S.-C., PENG, I. C., WANG, C.-Y. & TSAI, S.-J. 2020. Pathophysiological implications of hypoxia in human diseases. *Journal of Biomedical Science,* 27**,** 63.
CHEN, Y. & BELMONT, A. S. 2019. Genome organization around nuclear speckles. *Curr Opin Genet Dev,* 55**,** 91-99.
CHENG, H., DUFU, K., LEE, C. S., HSU, J. L., DIAS, A. & REED, R. 2006. Human mRNA export machinery recruited to the 5' end of mRNA. *Cell,* 127**,** 1389-400.
CHOTHANI, S., ADAMI, E., OUYANG, J. F., VISWANATHAN, S., HUBNER, N., COOK, S. A., SCHAFER, S. & RACKHAM, O. J. L. 2019. deltaTE: Detection of Translationally Regulated Genes by Integrative Analysis of Ribo-seq and RNA-seq Data. *Current Protocols in Molecular Biology,* 129**,** e108.
CHOTHIA, C. 1974. Hydrophobic bonding and accessible surface area in proteins. *Nature,* 248**,** 338-9.
CHOTHIA, C. 1975. Structural invariants in protein folding. *Nature,* 254**,** 304-8.
CIGAN, A. M., FENG, L. & DONAHUE, T. F. 1988. tRNAimet Functions in Directing the Scanning Ribosome to the Start Site of Translation. *Science,* 242**,** 93-97.
COHEN, G. B., REN, R. & BALTIMORE, D. 1995. Modular binding domains in signal transduction proteins. *Cell,* 80**,** 237-48.
COLLER, J. M., TUCKER, M., SHETH, U., VALENCIA-SANCHEZ, M. A. & PARKER, R. 2001. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *Rna,* 7**,** 1717-27.
COUREL, M., CLÉMENT, Y., BOSSEVAIN, C., FORETEK, D., VIDAL CRUCHEZ, O., YI, Z., BÉNARD, M., BENASSY, M. N., KRESS, M., VINDRY, C., ERNOULT-LANGE, M., ANTONIEWSKI, C., MORILLON, A., BREST, P., HUBSTENBERGER, A., ROEST CROLLIUS, H., STANDART, N. & WEIL, D. 2019. GC content shapes mRNA storage and decay in human cells. *Elife,* 8.
COX, J., NEUHAUSER, N., MICHALSKI, A., SCHELTEMA, R. A., OLSEN, J. V. & MANN, M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res,* 10**,** 1794-805.
CRAMER, P., ARMACHE, K. J., BAUMLI, S., BENKERT, S., BRUECKNER, F., BUCHEN, C., DAMSMA, G. E., DENGL, S., GEIGER, S. R., JASIAK, A. J., JAWHARI, A., JENNEBACH, S., KAMENSKI, T., KETTENBERGER, H., KUHN, C. D., LEHMANN, E., LEIKE, K., SYDOW, J. F. & VANNINI, A. 2008. Structure of eukaryotic RNA polymerases. *Annu Rev Biophys,* 37**,** 337-52.
CRICK, F. 1970. Central Dogma of Molecular Biology. *Nature,* 227**,** 561-563.
DANG, C. V. 2012. MYC on the path to cancer. *Cell,* 149**,** 22-35.
DECKER, C. J., BURKE, J. M., MULVANEY, P. K. & PARKER, R. 2022. RNA is required for the integrity of multiple nuclear and cytoplasmic membrane-less RNP granules. *Embo j,* 41**,** e110137.
DERENZINI, M., MONTANARO, L. & TRERÉ, D. 2009. What the nucleolus says to a tumour pathologist. *Histopathology,* 54**,** 753-62.
DEVER, T. E., FENG, L., WEK, R. C., CIGAN, A. M., DONAHUE, T. F. & HINNEBUSCH, A. G. 1992. Phosphorylation of initiation factor 2 alpha by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. *Cell,* 68**,** 585-96.
DOLICKA, D., FOTI, M. & SOBOLEWSKI, C. 2021. The Emerging Role of Stress Granules in Hepatocellular Carcinoma. *Int J Mol Sci,* 22.
DOMOTO, T., PYKO, I. V., FURUTA, T., MIYASHITA, K., UEHARA, M., SHIMASAKI, T., NAKADA, M. & MINAMOTO, T. 2016. Glycogen synthase

kinase-3β is a pivotal mediator of cancer invasion and resistance to therapy. *Cancer Sci,* 107**,** 1363-1372.
DURÁN, R. V. & HALL, M. N. 2012. Regulation of TOR by small GTPases. *EMBO Rep,* 13**,** 121-8.
ELBAUM-GARFINKLE, S., KIM, Y., SZCZEPANIAK, K., CHEN, C. C., ECKMANN, C. R., MYONG, S. & BRANGWYNNE, C. P. 2015. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A,* 112**,** 7189-94.
EMARA, M. M., FUJIMURA, K., SCIARANGHELLA, D., IVANOVA, V., IVANOV, P. & ANDERSON, P. 2012. Hydrogen peroxide induces stress granule formation independent of eIF2α phosphorylation. *Biochem Biophys Res Commun,* 423**,** 763-9.
FARLEY-BARNES, K. I., OGAWA, L. M. & BASERGA, S. J. 2019. Ribosomopathies: Old Concepts, New Controversies. *Trends Genet,* 35**,** 754-767.
FEI, J., JADALIHA, M., HARMON, T. S., LI, I. T. S., HUA, B., HAO, Q., HOLEHOUSE, A. S., REYER, M., SUN, Q., FREIER, S. M., PAPPU, R. V., PRASANTH, K. V. & HA, T. 2017. Quantitative analysis of multilayer organization of proteins and RNA in nuclear speckles at super resolution. *J Cell Sci,* 130**,** 4180-4192.
FISHER, R. A. 1992. *Statistical methods for research workers*, Springer.
FRINGER, J. M., ACKER, M. G., FEKETE, C. A., LORSCH, J. R. & DEVER, T. E. 2007. Coupled Release of Eukaryotic Translation Initiation Factors 5B and 1A from 80S Ribosomes following Subunit Joining. *Molecular and Cellular Biology,* 27**,** 2384-2397.
FU, L., NIU, B., ZHU, Z., WU, S. & LI, W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics,* 28**,** 3150-2.
FUJIMURA, K., SASAKI, A. T. & ANDERSON, P. 2012. Selenite targets eIF4E-binding protein-1 to inhibit translation initiation and induce the assembly of non-canonical stress granules. *Nucleic Acids Res,* 40**,** 8099-110.
GAO, J., AKSOY, B. A., DOGRUSOZ, U., DRESDNER, G., GROSS, B., SUMER, S. O., SUN, Y., JACOBSEN, A., SINHA, R., LARSSON, E., CERAMI, E., SANDER, C. & SCHULTZ, N. 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal,* 6**,** pl1.
GATTO, L. & LILLEY, K. S. 2012. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. *Bioinformatics,* 28**,** 288-9.
GHADDAR, N., WANG, S., WOODVINE, B., KRISHNAMOORTHY, J., VAN HOEF, V., DARINI, C., KAZIMIERCZAK, U., AH-SON, N., POPPER, H., JOHNSON, M., OFFICER, L., TEODÓSIO, A., BROGGINI, M., MANN, K. K., HATZOGLOU, M., TOPISIROVIC, I., LARSSON, O., LE QUESNE, J. & KOROMILAS, A. E. 2021. The integrated stress response is tumorigenic and constitutes a therapeutic liability in KRAS-driven lung cancer. *Nature Communications,* 12**,** 4651.
GHOSH, R., LIPSON, K. L., SARGENT, K. E., MERCURIO, A. M., HUNT, J. S., RON, D. & URANO, F. 2010. Transcriptional regulation of VEGF-A by the unfolded protein response pathway. *PLoS One,* 5**,** e9575.
GILKS, N., KEDERSHA, N., AYODELE, M., SHEN, L., STOECKLIN, G., DEMBER, L. M. & ANDERSON, P. 2004. Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell,* 15**,** 5383-98.
GILLEN, S. L., GIACOMELLI, C., HODGE, K., ZANIVAN, S., BUSHELL, M. & WILCZYNSKA, A. 2021. Differential regulation of mRNA fate by the human Ccr4-Not complex is driven by coding sequence composition and mRNA localization. *Genome Biol,* 22**,** 284.

GINGRAS, A.-C., RAUGHT, B. & SONENBERG, N. 1999. eIF4 Initiation Factors: Effectors of mRNA Recruitment to Ribosomes and Regulators of Translation. *Annual Review of Biochemistry,* 68**,** 913-963.

GINGRAS, A. C., KENNEDY, S. G., O'LEARY, M. A., SONENBERG, N. & HAY, N. 1998. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev,* 12**,** 502-13.

GOLDMAN, M. J., CRAFT, B., HASTIE, M., REPEČKA, K., MCDADE, F., KAMATH, A., BANERJEE, A., LUO, Y., ROGERS, D., BROOKS, A. N., ZHU, J. & HAUSSLER, D. 2020. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol,* 38**,** 675-678.

GONZÁLEZ-MARISCAL, L., MIRANDA, J., GALLEGO-GUTIÉRREZ, H., CANO-CORTINA, M. & AMAYA, E. 2020. Relationship between apical junction proteins, gene expression and cancer. *Biochim Biophys Acta Biomembr,* 1862**,** 183278.

GOTTSCHALD, O. R., MALEC, V., KRASTEVA, G., HASAN, D., KAMLAH, F., HEROLD, S., ROSE, F., SEEGER, W. & HÄNZE, J. 2010. TIAR and TIA-1 mRNA-binding proteins co-aggregate under conditions of rapid oxygen decline and extreme hypoxia and suppress the HIF-1α pathway. *J Mol Cell Biol,* 2**,** 345-56.

GRABOCKA, E. & BAR-SAGI, D. 2016. Mutant KRAS Enhances Tumor Cell Fitness by Upregulating Stress Granules. *Cell,* 167**,** 1803-1813.e12.

GRABOWSKI, P. J., SEILER, S. R. & SHARP, P. A. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell,* 42**,** 345-353.

GRAY, N. K. & HENTZE, M. W. 1994. Regulation of protein synthesis by mRNA structure. *Mol Biol Rep,* 19**,** 195-200.

GUILLÉN-BOIXET, J., KOPACH, A., HOLEHOUSE, A. S., WITTMANN, S., JAHNEL, M., SCHLÜßLER, R., KIM, K., TRUSSINA, I., WANG, J., MATEJU, D., POSER, I., MAHARANA, S., RUER-GRUß, M., RICHTER, D., ZHANG, X., CHANG, Y. T., GUCK, J., HONIGMANN, A., MAHAMID, J., HYMAN, A. A., PAPPU, R. V., ALBERTI, S. & FRANZMANN, T. M. 2020. RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell,* 181**,** 346-361.e17.

GUPTA, S., TAKEBE, N. & LORUSSO, P. 2010. Targeting the Hedgehog pathway in cancer. *Ther Adv Med Oncol,* 2**,** 237-50.

GWINN, D. M., SHACKELFORD, D. B., EGAN, D. F., MIHAYLOVA, M. M., MERY, A., VASQUEZ, D. S., TURK, B. E. & SHAW, R. J. 2008. AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol Cell,* 30**,** 214-26.

HAGHIGHAT, A. & SONENBERG, N. 1997. eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J Biol Chem,* 272**,** 21677-80.

HANAHAN, D. 2022. Hallmarks of Cancer: New Dimensions. *Cancer Discov,* 12**,** 31-46.

HANDWERGER, K. E. & GALL, J. G. 2006. Subnuclear organelles: new insights into form and function. *Trends Cell Biol,* 16**,** 19-26.

HAYES, J. D., DINKOVA-KOSTOVA, A. T. & TEW, K. D. 2020. Oxidative Stress in Cancer. *Cancer Cell,* 38**,** 167-197.

HE, Y., FANG, J., TAATJES, D. J. & NOGALES, E. 2013. Structural visualization of key steps in human transcription initiation. *Nature,* 495**,** 481-6.

HE, Y., SUN, M. M., ZHANG, G. G., YANG, J., CHEN, K. S., XU, W. W. & LI, B. 2021. Targeting PI3K/Akt signal transduction for cancer therapy. *Signal Transduction and Targeted Therapy,* 6**,** 425.

HELLEDAY, T., PETERMANN, E., LUNDIN, C., HODGSON, B. & SHARMA, R. A. 2008. DNA repair pathways as targets for cancer therapy. *Nature Reviews Cancer,* 8**,** 193-204.

HIROSE, T., NINOMIYA, K., NAKAGAWA, S. & YAMAZAKI, T. 2023. A guide to membraneless organelles and their various roles in gene regulation. *Nature Reviews Molecular Cell Biology,* 24**,** 288-304.

HOESEL, B. & SCHMID, J. A. 2013. The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer,* 12**,** 86.

HOFFMANN, P. C., KREYSING, J. P., KHUSAINOV, I., TUIJTEL, M. W., WELSCH, S. & BECK, M. 2022. Structures of the eukaryotic ribosome and its translational states in situ. *Nature Communications,* 13**,** 7435.

HOLMES, K., ROBERTS, O. L., THOMAS, A. M. & CROSS, M. J. 2007. Vascular endothelial growth factor receptor-2: structure, function, intracellular signalling and therapeutic inhibition. *Cell Signal,* 19**,** 2003-12.

HOLSTEGE, F. C., FIEDLER, U. & TIMMERS, H. T. 1997. Three transitions in the RNA polymerase II transcription complex during initiation. *Embo j,* 16**,** 7468-80.

HON, W. C., WILSON, M. I., HARLOS, K., CLARIDGE, T. D., SCHOFIELD, C. J., PUGH, C. W., MAXWELL, P. H., RATCLIFFE, P. J., STUART, D. I. & JONES, E. Y. 2002. Structural basis for the recognition of hydroxyproline in HIF-1 alpha by pVHL. *Nature,* 417**,** 975-8.

HONDELE, M., SACHDEV, R., HEINRICH, S., WANG, J., VALLOTTON, P., FONTOURA, B. M. A. & WEIS, K. 2019. DEAD-box ATPases are global regulators of phase-separated organelles. *Nature,* 573**,** 144-148.

HUBER, W., VON HEYDEBRECK, A., SÜLTMANN, H., POUSTKA, A. & VINGRON, M. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics,* 18 Suppl 1**,** S96-104.

HUBSTENBERGER, A., COUREL, M., BÉNARD, M., SOUQUERE, S., ERNOULT-LANGE, M., CHOUAIB, R., YI, Z., MORLOT, J.-B., MUNIER, A., FRADET, M., DAUNESSE, M., BERTRAND, E., PIERRON, G., MOZZICONACCI, J., KRESS, M. & WEIL, D. 2017. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Molecular Cell,* 68**,** 144-157.e5.

INGELFINGER, D., ARNDT-JOVIN, D. J., LÜHRMANN, R. & ACHSEL, T. 2002. The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *Rna,* 8**,** 1489-501.

INGOLIA, N. T., GHAEMMAGHAMI, S., NEWMAN, J. R. S. & WEISSMAN, J. S. 2009. Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science,* 324**,** 218-223.

ITATANI, Y., KAWADA, K., YAMAMOTO, T. & SAKAI, Y. 2018. Resistance to Anti-Angiogenic Therapy in Cancer-Alterations to Anti-VEGF Pathway. *Int J Mol Sci,* 19.

JACKSON, R. & STANDART, N. 2015. The awesome power of ribosome profiling. *Rna,* 21**,** 652-4.

JAIN, S., WHEELER, J. R., WALTERS, R. W., AGRAWAL, A., BARSIC, A. & PARKER, R. 2016. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell,* 164**,** 487-98.

JOHN A. RYAN. *Cell Cloning By Serial Dilution in 96 Well Plates* [Online]. corning.com. Available: https://www.corning.com/catalog/cls/documents/protocols/Single_cell_cloning_protocol.pdf [Accessed].

JOHNSON, J. O., MANDRIOLI, J., BENATAR, M., ABRAMZON, Y., VAN DEERLIN, V. M., TROJANOWSKI, J. Q., GIBBS, J. R., BRUNETTI, M., GRONKA, S., WUU, J., DING, J., MCCLUSKEY, L., MARTINEZ-LAGE, M., FALCONE, D., HERNANDEZ, D. G., AREPALLI, S., CHONG, S., SCHYMICK, J. C., ROTHSTEIN, J., LANDI, F., WANG, Y.-D., CALVO, A., MORA, G., SABATELLI, M., MONSURRÒ, M. R., BATTISTINI, S., SALVI, F., SPATARO, R., SOLA, P., BORGHERO, G., GALASSI, G., SCHOLZ, S. W., TAYLOR, J. P.,

RESTAGNO, G., CHIÒ, A. & TRAYNOR, B. J. 2010. Exome Sequencing Reveals VCP Mutations as a Cause of Familial ALS. *Neuron,* 68**,** 857-864.

KAEHLER, C., ISENSEE, J., HUCHO, T., LEHRACH, H. & KROBITSCH, S. 2014. 5-Fluorouracil affects assembly of stress granules based on RNA incorporation. *Nucleic Acids Research,* 42**,** 6436-6447.

KAPP, L. D. & LORSCH, J. R. 2004. The molecular mechanics of eukaryotic translation. *Annu Rev Biochem,* 73**,** 657-704.

KEDERSHA, N., CHEN, S., GILKS, N., LI, W., MILLER, I. J., STAHL, J. & ANDERSON, P. 2002. Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Mol Biol Cell,* 13**,** 195-210.

KEDERSHA, N., CHO, M. R., LI, W., YACONO, P. W., CHEN, S., GILKS, N., GOLAN, D. E. & ANDERSON, P. 2000. Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J Cell Biol,* 151**,** 1257-68.

KEDERSHA, N., PANAS, M. D., ACHORN, C. A., LYONS, S., TISDALE, S., HICKMAN, T., THOMAS, M., LIEBERMAN, J., MCINERNEY, G. M., IVANOV, P. & ANDERSON, P. 2016. G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J Cell Biol,* 212**,** 845-60.

KEDERSHA, N., STOECKLIN, G., AYODELE, M., YACONO, P., LYKKE-ANDERSEN, J., FRITZLER, M. J., SCHEUNER, D., KAUFMAN, R. J., GOLAN, D. E. & ANDERSON, P. 2005. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol,* 169**,** 871-84.

KHAPERSKYY, D. A., HATCHETTE, T. F. & MCCORMICK, C. 2012. Influenza A virus inhibits cytoplasmic stress granule formation. *The FASEB Journal,* 26**,** 1629-1639.

KHONG, A., JAIN, S., MATHENY, T., WHEELER, J. R. & PARKER, R. 2018. Isolation of mammalian stress granule cores for RNA-Seq analysis. *Methods,* 137**,** 49-54.

KHONG, A., MATHENY, T., HUYNH, T. N., BABL, V. & PARKER, R. 2022. Limited effects of m6A modification on mRNA partitioning into stress granules. *Nature Communications,* 13**,** 3735.

KHONG, A., MATHENY, T., JAIN, S., MITCHELL, S. F., WHEELER, J. R. & PARKER, R. 2017. The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Mol Cell,* 68**,** 808-820.e5.

KIM, D. H., SARBASSOV, D. D., ALI, S. M., KING, J. E., LATEK, R. R., ERDJUMENT-BROMAGE, H., TEMPST, P. & SABATINI, D. M. 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell,* 110**,** 163-75.

KIM, D. H., SARBASSOV, D. D., ALI, S. M., LATEK, R. R., GUNTUR, K. V., ERDJUMENT-BROMAGE, H., TEMPST, P. & SABATINI, D. M. 2003. GbetaL, a positive regulator of the rapamycin-sensitive pathway required for the nutrient-sensitive interaction between raptor and mTOR. *Mol Cell,* 11**,** 895-904.

KIMBALL, S. R., HORETSKY, R. L., RON, D., JEFFERSON, L. S. & HARDING, H. P. 2003. Mammalian stress granules represent sites of accumulation of stalled translation initiation complexes. *Am J Physiol Cell Physiol,* 284**,** C273-84.

KIMONIS, V. E., MEHTA, S. G., FULCHIERO, E. C., THOMASOVA, D., PASQUALI, M., BOYCOTT, K., NEILAN, E. G., KARTASHOV, A., FORMAN, M. S., TUCKER, S., KIMONIS, K., MUMM, S., WHYTE, M. P., SMITH, C. D. & WATTS, G. D. J. 2008. Clinical studies in familial VCP myopathy associated with Paget disease of bone and frontotemporal dementia. *American Journal of Medical Genetics Part A,* 146A**,** 745-757.

KOBAYASHI, T., FUNAKOSHI, Y., HOSHINO, S.-I. & KATADA, T. 2004. The GTP-binding Release Factor eRF3 as a Key Mediator Coupling Translation Termination

to mRNA Decay *. *Journal of Biological Chemistry,* 279**,** 45693-45700.

KORNEEVA, N. L., LAMPHEAR, B. J., HENNIGAN, F. L. & RHOADS, R. E. 2000. Mutually cooperative binding of eukaryotic translation initiation factor (eIF) 3 and eIF4A to human eIF4G-1. *J Biol Chem,* 275**,** 41369-76.

KOZAK, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res,* 15**,** 8125-48.

KULESHOV, M. V., JONES, M. R., ROUILLARD, A. D., FERNANDEZ, N. F., DUAN, Q., WANG, Z., KOPLEV, S., JENKINS, S. L., JAGODNIK, K. M., LACHMANN, A., MCDERMOTT, M. G., MONTEIRO, C. D., GUNDERSEN, G. W. & MA'AYAN, A. 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res,* 44**,** W90-7.

KURIHARA, M., KATO, K., SANBO, C., SHIGENOBU, S., OHKAWA, Y., FUCHIGAMI, T. & MIYANARI, Y. 2020. Genomic Profiling by ALaP-Seq Reveals Transcriptional Regulation by PML Bodies through DNMT3A Exclusion. *Mol Cell,* 78**,** 493-505.e8.

LAFONTAINE, D. L. J., RIBACK, J. A., BASCETIN, R. & BRANGWYNNE, C. P. 2021. The nucleolus as a multiphase liquid condensate. *Nature Reviews Molecular Cell Biology,* 22**,** 165-182.

LAMPER, A. M., FLEMING, R. H., LADD, K. M. & LEE, A. S. Y. 2020. A phosphorylation-regulated eIF3d translation switch mediates cellular adaptation to metabolic stress. *Science,* 370**,** 853-856.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-9.

LE HIR, H., GATFIELD, D., IZAURRALDE, E. & MOORE, M. J. 2001. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *Embo j,* 20**,** 4987-97.

LE HIR, H., IZAURRALDE, E., MAQUAT, L. E. & MOORE, M. J. 2000. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *Embo j,* 19**,** 6860-9.

LE, Q. T., CHEN, E., SALIM, A., CAO, H., KONG, C. S., WHYTE, R., DONINGTON, J., CANNON, W., WAKELEE, H., TIBSHIRANI, R., MITCHELL, J. D., RICHARDSON, D., O'BYRNE, K. J., KOONG, A. C. & GIACCIA, A. J. 2006. An evaluation of tumor oxygenation and gene expression in patients with early stage non-small cell lung cancers. *Clin Cancer Res,* 12**,** 1507-14.

LEE, A. S. Y., KRANZUSCH, P. J., DOUDNA, J. A. & CATE, J. H. D. 2016. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature,* 536**,** 96-99.

LEE, J. E., CATHEY, P. I., WU, H., PARKER, R. & VOELTZ, G. K. 2020. Endoplasmic reticulum contact sites regulate the dynamics of membraneless organelles. *Science,* 367.

LEE, J. W., KO, J., JU, C. & ELTZSCHIG, H. K. 2019. Hypoxia signaling in human diseases and therapeutic targets. *Experimental & Molecular Medicine,* 51**,** 1-13.

LEE, M., SADOWSKA, A., BEKERE, I., HO, D., GULLY, BENJAMIN S., LU, Y., IYER, K. S., TREWHELLA, J., FOX, A. H. & BOND, C. S. 2015. The structure of human SFPQ reveals a coiled-coil mediated polymer essential for functional aggregation in gene regulation. *Nucleic Acids Research,* 43**,** 3826-3840.

LEVINE, A. J. 2022. Targeting the P53 Protein for Cancer Therapies: The Translational Impact of P53 Research. *Cancer Research,* 82**,** 362-364.

LI, B. & DEWEY, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics,* 12**,** 323.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., DURBIN, R. & SUBGROUP, G. P. D. P. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-2079.

LIBERZON, A., BIRGER, C., THORVALDSDÓTTIR, H., GHANDI, M., MESIROV, J. P. & TAMAYO, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* 1**,** 417-425.

LIU-YESUCEVITZ, L., BILGUTAY, A., ZHANG, Y. J., VANDERWEYDE, T., CITRO, A., MEHTA, T., ZAARUR, N., MCKEE, A., BOWSER, R., SHERMAN, M., PETRUCELLI, L. & WOLOZIN, B. 2010. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One,* 5**,** e13250.

LÖBRICH, M. & JEGGO, P. A. 2007. The impact of a negligent G2/M checkpoint on genomic instability and cancer induction. *Nature Reviews Cancer,* 7**,** 861-869.

LORSCH, J. R. & HERSCHLAG, D. 1998. The DEAD Box Protein eIF4A. 1. A Minimal Kinetic and Thermodynamic Framework Reveals Coupled Binding of RNA and Nucleotide. *Biochemistry,* 37**,** 2180-2193.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology,* 15**,** 550.

LU, L., HAN, A. P. & CHEN, J. J. 2001. Translation initiation control by heme-regulated eukaryotic initiation factor 2alpha kinase in erythroid cells under cytoplasmic stresses. *Mol Cell Biol,* 21**,** 7971-80.

LU, P. D., HARDING, H. P. & RON, D. 2004. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol,* 167**,** 27-33.

MAAG, D., FEKETE, C. A., GRYCZYNSKI, Z. & LORSCH, J. R. 2005. A Conformational Change in the Eukaryotic Translation Preinitiation Complex and Release of eIF1 Signal Recognition of the Start Codon. *Molecular Cell,* 17**,** 265-275.

MADDEN, E., LOGUE, S. E., HEALY, S. J., MANIE, S. & SAMALI, A. 2019. The role of the unfolded protein response in cancer progression: From oncogenesis to chemoresistance. *Biol Cell,* 111**,** 1-17.

MAO, Y. S., SUNWOO, H., ZHANG, B. & SPECTOR, D. L. 2011a. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nature Cell Biology,* 13**,** 95-101.

MAO, Y. S., ZHANG, B. & SPECTOR, D. L. 2011b. Biogenesis and function of nuclear bodies. *Trends Genet,* 27**,** 295-306.

MARCOTRIGIANO, J., GINGRAS, A. C., SONENBERG, N. & BURLEY, S. K. 1999. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell,* 3**,** 707-16.

MARINHO, H. S., REAL, C., CYRNE, L., SOARES, H. & ANTUNES, F. 2014. Hydrogen peroxide sensing, signaling and regulation of transcription factors. *Redox Biol,* 2**,** 535-62.

MARKMILLER, S., SOLTANIEH, S., SERVER, K. L., MAK, R., JIN, W., FANG, M. Y., LUO, E. C., KRACH, F., YANG, D., SEN, A., FULZELE, A., WOZNIAK, J. M., GONZALEZ, D. J., KANKEL, M. W., GAO, F. B., BENNETT, E. J., LÉCUYER, E. & YEO, G. W. 2018. Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell,* 172**,** 590-604.e13.

MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011,* 17**,** 3.

MASSON, N., WILLAM, C., MAXWELL, P. H., PUGH, C. W. & RATCLIFFE, P. J. 2001. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. *Embo j,* 20**,** 5197-206.

MATEJU, D. & CHAO, J. A. 2022. Stress granules: regulators or by-products? *The FEBS Journal*, 289**,** 363-373.

MATEJU, D., EICHENBERGER, B., VOIGT, F., EGLINGER, J., ROTH, G. & CHAO, J. A. 2020. Single-Molecule Imaging Reveals Translation of mRNAs Localized to Stress Granules. *Cell*, 183**,** 1801-1812.e13.

MATSUKI, H., TAKAHASHI, M., HIGUCHI, M., MAKOKHA, G. N., OIE, M. & FUJII, M. 2013. Both G3BP1 and G3BP2 contribute to stress granule formation. *Genes Cells*, 18**,** 135-46.

MAZROUI, R., SUKARIEH, R., BORDELEAU, M. E., KAUFMAN, R. J., NORTHCOTE, P., TANAKA, J., GALLOUZI, I. & PELLETIER, J. 2006. Inhibition of ribosome recruitment induces stress granule formation independently of eukaryotic initiation factor 2alpha phosphorylation. *Mol Biol Cell*, 17**,** 4212-9.

MCEWEN, E., KEDERSHA, N., SONG, B., SCHEUNER, D., GILKS, N., HAN, A., CHEN, J. J., ANDERSON, P. & KAUFMAN, R. J. 2005. Heme-regulated inhibitor kinase-mediated phosphorylation of eukaryotic translation initiation factor 2 inhibits translation, induces stress granule formation, and mediates survival upon arsenite exposure. *J Biol Chem*, 280**,** 16925-33.

MCGLINCY, N. J. & INGOLIA, N. T. 2017. Transcriptome-wide measurement of translation by ribosome profiling. *Methods*, 126**,** 112-129.

MCINERNEY, G. M., KEDERSHA, N. L., KAUFMAN, R. J., ANDERSON, P. & LILJESTRÖM, P. 2005. Importance of eIF2alpha phosphorylation and stress granule assembly in alphavirus translation regulation. *Mol Biol Cell*, 16**,** 3753-63.

MELANSON, G., TIMPANO, S. & UNIACKE, J. 2017. The eIF4E2-Directed Hypoxic Cap-Dependent Translation Machinery Reveals Novel Therapeutic Potential for Cancer Treatment. *Oxid Med Cell Longev*, 2017**,** 6098107.

MEURS, E., CHONG, K., GALABRU, J., THOMAS, N. S., KERR, I. M., WILLIAMS, B. R. & HOVANESSIAN, A. G. 1990. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. *Cell*, 62**,** 379-90.

MIHAYLOVA, M. M. & SHAW, R. J. 2011. The AMPK signalling pathway coordinates cell growth, autophagy and metabolism. *Nat Cell Biol*, 13**,** 1016-23.

MOLLIEX, A., TEMIROV, J., LEE, J., COUGHLIN, M., KANAGARAJ, A. P., KIM, H. J., MITTAG, T. & TAYLOR, J. P. 2015. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, 163**,** 123-33.

MOORE, M. J. & SHARP, P. A. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*, 365**,** 364-8.

MORLEY, S. J., CURTIS, P. S. & PAIN, V. M. 1997. eIF4G: translation's mystery factor begins to yield its secrets. *Rna*, 3**,** 1085-104.

MOTEKI, S. & PRICE, D. 2002. Functional coupling of capping and transcription of mRNA. *Mol Cell*, 10**,** 599-609.

MOUTAOUFIK, M. T., EL FATIMY, R., NASSOUR, H., GAREAU, C., LANG, J., TANGUAY, R. M., MAZROUI, R. & KHANDJIAN, E. W. 2014. UVC-induced stress granules in mammalian cells. *PLoS One*, 9**,** e112742.

MUDGE, J. M., RUIZ-ORERA, J., PRENSNER, J. R., BRUNET, M. A., CALVET, F., JUNGREIS, I., GONZALEZ, J. M., MAGRANE, M., MARTINEZ, T. F., SCHULZ, J. F., YANG, Y. T., ALBÀ, M. M., ASPDEN, J. L., BARANOV, P. V., BAZZINI, A. A., BRUFORD, E., MARTIN, M. J., CALVIELLO, L., CARVUNIS, A.-R., CHEN, J., COUSO, J. P., DEUTSCH, E. W., FLICEK, P., FRANKISH, A., GERSTEIN, M., HUBNER, N., INGOLIA, N. T., KELLIS, M., MENSCHAERT, G., MORITZ, R. L., OHLER, U., ROUCOU, X., SAGHATELIAN, A., WEISSMAN, J. S. & VAN HEESCH, S. 2022. Standardized annotation of translated open reading frames. *Nature Biotechnology*, 40**,** 994-999.

NADEZHDINA, E. S., LOMAKIN, A. J., SHPILMAN, A. A., CHUDINOVA, E. M. & IVANOV, P. A. 2010. Microtubules govern stress granule mobility and dynamics. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* 1803**,** 361-371.

NAGANUMA, T., NAKAGAWA, S., TANIGAWA, A., SASAKI, Y. F., GOSHIMA, N. & HIROSE, T. 2012. Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *Embo j,* 31**,** 4020-34.

NAMKOONG, S., HO, A., WOO, Y. M., KWAK, H. & LEE, J. H. 2018. Systematic Characterization of Stress-Induced RNA Granulation. *Mol Cell,* 70**,** 175-187.e8.

NISSEN, P., HANSEN, J., BAN, N., MOORE, P. B. & STEITZ, T. A. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science,* 289**,** 920-30.

NOTT, T. J., PETSALAKI, E., FARBER, P., JERVIS, D., FUSSNER, E., PLOCHOWIETZ, A., CRAGGS, T. D., BAZETT-JONES, D. P., PAWSON, T., FORMAN-KAY, J. D. & BALDWIN, A. J. 2015. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell,* 57**,** 936-947.

NOWELL, C. S. & RADTKE, F. 2017. Notch as a tumour suppressor. *Nature Reviews Cancer,* 17**,** 145-159.

OZTURK, S. B. & KINZY, T. G. 2008. Guanine nucleotide exchange factor independence of the G-protein eEF1A through novel mutant forms and biochemical properties. *J Biol Chem,* 283**,** 23244-53.

PADRÓN, A., IWASAKI, S. & INGOLIA, N. T. 2019. Proximity RNA Labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules. *Mol Cell,* 75**,** 875-887.e5.

PAKOS-ZEBRUCKA, K., KORYGA, I., MNICH, K., LJUJIC, M., SAMALI, A. & GORMAN, A. M. 2016. The integrated stress response. *EMBO Rep,* 17**,** 1374-1395.

PANAS, M. D., IVANOV, P. & ANDERSON, P. 2016. Mechanistic insights into mammalian stress granule dynamics. *J Cell Biol,* 215**,** 313-323.

PANAS, M. D., VARJAK, M., LULLA, A., ER ENG, K., MERITS, A., KARLSSON HEDESTAM, G. B. & MCINERNEY, G. M. 2012. Sequestration of G3BP coupled with efficient translation inhibits stress granules in Semliki Forest virus infection. *Molecular Biology of the Cell,* 23**,** 4701-4712.

PANDYA-JONES, A. 2011. Pre-mRNA splicing during transcription in the mammalian system. *Wiley Interdiscip Rev RNA,* 2**,** 700-17.

PAULIN, F. E. M., CAMPBELL, L. E., O'BRIEN, K., LOUGHLIN, J. & PROUD, C. G. 2001. Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Current Biology,* 11**,** 55-59.

PAVITT, G. D., YANG, W. & HINNEBUSCH, A. G. 1997. Homologous segments in three subunits of the guanine nucleotide exchange factor eIF2B mediate translational regulation by phosphorylation of eIF2. *Mol Cell Biol,* 17**,** 1298-313.

PAWSON, T. & NASH, P. 2003. Assembly of cell regulatory systems through protein interaction domains. *Science,* 300**,** 445-52.

PEDERSON, T. 2011. The nucleolus. *Cold Spring Harb Perspect Biol,* 3.

PESTKA, S. 1971. Inhibitors of ribosome functions. *Annu Rev Microbiol,* 25**,** 487-562.

PESTOVA, T. V., LOMAKIN, I. B., LEE, J. H., CHOI, S. K., DEVER, T. E. & HELLEN, C. U. 2000. The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature,* 403**,** 332-5.

PISAREV, A. V., SKABKIN, M. A., PISAREVA, V. P., SKABKINA, O. V., RAKOTONDRAFARA, A. M., HENTZE, M. W., HELLEN, C. U. & PESTOVA, T. V. 2010. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell,* 37**,** 196-210.

POKHOLOK, D. K., HANNETT, N. M. & YOUNG, R. A. 2002. Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. *Mol Cell,* 9**,** 799-809.

PORIA, D. K. & RAY, P. S. 2017. Polysome Analysis. *Bio Protoc,* 7.

PORT, J., MUTHALAGU, N., RAJA, M., CETECI, F., MONTEVERDE, T., KRUSPIG, B., HEDLEY, A., KALNA, G., LILLA, S., NEILSON, L., BRUCOLI, M., GYURASZOVA, K., TAIT-MULDER, J., MEZNA, M., SVAMBARYTE, S., BRYSON, A., SUMPTON, D., MCVIE, A., NIXON, C., DRYSDALE, M., ESUMI, H., MURRAY, G. I., SANSOM, O. J., ZANIVAN, S. R. & MURPHY, D. J. 2018. Colorectal Tumors Require NUAK1 for Protection from Oxidative Stress. *Cancer Discovery,* 8**,** 632-647.

PROTTER, D. S. W. & PARKER, R. 2016. Principles and Properties of Stress Granules. *Trends Cell Biol,* 26**,** 668-679.

RABL, J., LEIBUNDGUT, M., ATAIDE, S. F., HAAG, A. & BAN, N. 2011. Crystal Structure of the Eukaryotic 40S Ribosomal Subunit in Complex with Initiation Factor 1. *Science,* 331**,** 730-736.

REINEKE, L. C., TSAI, W. C., JAIN, A., KAELBER, J. T., JUNG, S. Y. & LLOYD, R. E. 2017. Casein Kinase 2 Is Linked to Stress Granule Dynamics through Phosphorylation of the Stress Granule Nucleating Protein G3BP1. *Mol Cell Biol,* 37.

RHEE, H. J., KIM, E. J. & LEE, J. K. 2007. Physiological polyamines: simple primordial stress molecules. *J Cell Mol Med,* 11**,** 685-703.

RIBATTI, D., TAMMA, R. & ANNESE, T. 2020. Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Transl Oncol,* 13**,** 100773.

RICHARD, P. & MANLEY, J. L. 2009. Transcription termination by nuclear RNA polymerases. *Genes Dev,* 23**,** 1247-69.

RIES, R. J., ZACCARA, S., KLEIN, P., OLARERIN-GEORGE, A., NAMKOONG, S., PICKERING, B. F., PATIL, D. P., KWAK, H., LEE, J. H. & JAFFREY, S. R. 2019. m6A enhances the phase separation potential of mRNA. *Nature,* 571**,** 424-428.

RIGGS, C. L., KEDERSHA, N., IVANOV, P. & ANDERSON, P. 2020. Mammalian stress granules and P bodies at a glance. *Journal of Cell Science,* 133.

RITCHIE, M. E., PHIPSON, B., WU, D., HU, Y., LAW, C. W., SHI, W. & SMYTH, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* 43**,** e47.

RODNINA, M. V., BERINGER, M. & WINTERMEYER, W. 2007. How ribosomes make peptide bonds. *Trends in biochemical sciences,* 32**,** 20-26.

RODNINA, M. V., SAVELSBERGH, A., KATUNIN, V. I. & WINTERMEYER, W. 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature,* 385**,** 37-41.

SAITOH, M., PULLEN, N., BRENNAN, P., CANTRELL, D., DENNIS, P. B. & THOMAS, G. 2002. Regulation of an activated S6 kinase 1 variant reveals a novel mammalian target of rapamycin phosphorylation site. *J Biol Chem,* 277**,** 20104-12.

SAITOH, N., SPAHR, C. S., PATTERSON, S. D., BUBULYA, P., NEUWALD, A. F. & SPECTOR, D. L. 2004. Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell,* 15**,** 3876-90.

SALAS-MARCO, J. & BEDWELL, D. M. 2004. GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. *Mol Cell Biol,* 24**,** 7769-78.

SANDERS, D. W., KEDERSHA, N., LEE, D. S. W., STROM, A. R., DRAKE, V., RIBACK, J. A., BRACHA, D., EEFTENS, J. M., IWANICKI, A., WANG, A., WEI, M. T., WHITNEY, G., LYONS, S. M., ANDERSON, P., JACOBS, W. M.,

IVANOV, P. & BRANGWYNNE, C. P. 2020. Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell,* 181**,** 306-324.e28.

SASIKUMAR, A. N., PEREZ, W. B. & KINZY, T. G. 2012. The many roles of the eukaryotic elongation factor 1 complex. *WIREs RNA,* 3**,** 543-555.

SCHÄFER, G., CRAMER, T., SUSKE, G., KEMMNER, W., WIEDENMANN, B. & HÖCKER, M. 2003. Oxidative Stress Regulates Vascular Endothelial Growth Factor-A Gene Transcription through Sp1- and Sp3-dependent Activation of Two Proximal GC-rich Promoter Elements*. *Journal of Biological Chemistry,* 278**,** 8190-8198.

SCHINDELIN, J., ARGANDA-CARRERAS, I., FRISE, E., KAYNIG, V., LONGAIR, M., PIETZSCH, T., PREIBISCH, S., RUEDEN, C., SAALFELD, S., SCHMID, B., TINEVEZ, J.-Y., WHITE, D. J., HARTENSTEIN, V., ELICEIRI, K., TOMANCAK, P. & CARDONA, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods,* 9**,** 676-682.

SCHMIDT, T., DABROWSKA, A., WALDRON, J. A., HODGE, K., KOULOURAS, G., GABRIELSEN, M., MUNRO, J., TACK, D. C., HARRIS, G., MCGHEE, E., SCOTT, D., CARLIN, L. M., HUANG, D., LE QUESNE, J., ZANIVAN, S., WILCZYNSKA, A. & BUSHELL, M. 2023. eIF4A1-dependent mRNAs employ purine-rich 5'UTR sequences to activate localised eIF4A1-unwinding through eIF4A1-multimerisation to facilitate translation. *Nucleic Acids Res,* 51**,** 1859-1879.

SCHWARTZ, S. & MOTORIN, Y. 2017. Next-generation sequencing technologies for detection of modified nucleotides in RNAs. *RNA Biol,* 14**,** 1124-1137.

SEGUIN, S. J., MORELLI, F. F., VINET, J., AMORE, D., DE BIASI, S., POLETTI, A., RUBINSZTEIN, D. C. & CARRA, S. 2014. Inhibition of autophagy, lysosome and VCP function impairs stress granule assembly. *Cell Death Differ,* 21**,** 1838-51.

SEMENZA, G. L. 2003. Targeting HIF-1 for cancer therapy. *Nature Reviews Cancer,* 3**,** 721-732.

SERGUSHICHEV, A. A. 2016. An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv***,** 060012.

SFAKIANOS, A. P., MELLOR, L. E., PANG, Y. F., KRITSILIGKOU, P., NEEDS, H., ABOU-HAMDAN, H., DÉSAUBRY, L., POULIN, G. B., ASHE, M. P. & WHITMARSH, A. J. 2018. The mTOR-S6 kinase pathway promotes stress granule assembly. *Cell Death Differ,* 25**,** 1766-1780.

SHAHBAZI, J., LOCK, R. & LIU, T. 2013. Tumor Protein 53-Induced Nuclear Protein 1 Enhances p53 Function and Represses Tumorigenesis. *Front Genet,* 4**,** 80.

SHATKIN, A. J. & MANLEY, J. L. 2000. The ends of the affair: capping and polyadenylation. *Nat Struct Biol,* 7**,** 838-42.

SHETH, U. & PARKER, R. 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science,* 300**,** 805-8.

SHIN, Y. & BRANGWYNNE, C. P. 2017. Liquid phase condensation in cell physiology and disease. *Science,* 357**,** eaaf4382.

SIDRAUSKI, C., MCGEACHY, A. M., INGOLIA, N. T. & WALTER, P. 2015. The small molecule ISRIB reverses the effects of eIF2α phosphorylation on translation and stress granule assembly. *eLife,* 4**,** e05033.

SMITH, K. P., MOEN, P. T., WYDNER, K. L., COLEMAN, J. R. & LAWRENCE, J. B. 1999. Processing of endogenous pre-mRNAs in association with SC-35 domains is gene specific. *J Cell Biol,* 144**,** 617-29.

SMITH, T., HEGER, A. & SUDBERY, I. 2017. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res,* 27**,** 491-499.

SOKABE, M. & FRASER, C. S. 2017. A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Proc Natl Acad Sci U S A,* 114**,** 6304-6309.

SOMASEKHARAN, S. P., EL-NAGGAR, A., LEPRIVIER, G., CHENG, H., HAJEE, S., GRUNEWALD, T. G., ZHANG, F., NG, T., DELATTRE, O., EVDOKIMOVA, V., WANG, Y., GLEAVE, M. & SORENSEN, P. H. 2015. YB-1 regulates stress granule formation and tumor progression by translationally activating G3BP1. *J Cell Biol.*

SOMASEKHARAN, S. P., ZHANG, F., SAXENA, N., HUANG, J. N., KUO, I. C., LOW, C., BELL, R., ADOMAT, H., STOYNOV, N., FOSTER, L., GLEAVE, M. & SORENSEN, P. H. 2020. G3BP1-linked mRNA partitioning supports selective protein synthesis in response to oxidative stress. *Nucleic Acids Res,* 48**,** 6855-6873.

SONENBERG, N., RUPPRECHT, K. M., HECHT, S. M. & SHATKIN, A. J. 1979. Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled m7GDP. *Proc Natl Acad Sci U S A,* 76**,** 4345-9.

SONESON, C., LOVE, M. I. & ROBINSON, M. D. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res,* 4**,** 1521.

SONG, H., MUGNIER, P., DAS, A. K., WEBB, H. M., EVANS, D. R., TUITE, M. F., HEMMINGS, B. A. & BARFORD, D. 2000. The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell,* 100**,** 311-21.

SOUQUERE, S., MOLLET, S. P., KRESS, M., DAUTRY, F. O., PIERRON, G. R. & WEIL, D. 2009. Unravelling the ultrastructure of stress granules and associated P-bodies in human cells. *Journal of Cell Science,* 122**,** 3619-3626.

SPRIGGS, K. A., STONELEY, M., BUSHELL, M. & WILLIS, A. E. 2008. Re-programming of translation following cell stress allows IRES-mediated translation to predominate. *Biol Cell,* 100**,** 27-38.

SRIVASTAVA, S. P., KUMAR, K. U. & KAUFMAN, R. J. 1998. Phosphorylation of eukaryotic translation initiation factor 2 mediates apoptosis in response to activation of the double-stranded RNA-dependent protein kinase. *J Biol Chem,* 273**,** 2416-23.

STUMPFF, J., GHULE, P. N., SHIMAMURA, A., STEIN, J. L. & GREENBLATT, M. 2014. Spindle microtubule dysfunction and cancer predisposition. *J Cell Physiol,* 229**,** 1881-3.

SVITKIN, Y. V., PAUSE, A., HAGHIGHAT, A., PYRONNET, S., WITHERELL, G., BELSHAM, G. J. & SONENBERG, N. 2001. The requirement for eukaryotic initiation factor 4A (elF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *Rna,* 7**,** 382-94.

SYED, V. 2016. TGF-β Signaling in Cancer. *J Cell Biochem,* 117**,** 1279-87.

TAKAHARA, T. & MAEDA, T. 2012. Transient sequestration of TORC1 into stress granules during heat stress. *Mol Cell,* 47**,** 242-52.

TAUBER, D., TAUBER, G., KHONG, A., VAN TREECK, B., PELLETIER, J. & PARKER, R. 2020. Modulation of RNA Condensation by the DEAD-Box Protein eIF4A. *Cell,* 180**,** 411-426.e16.

THEDIECK, K., HOLZWARTH, B., PRENTZELL, M. T., BOEHLKE, C., KLÄSENER, K., RUF, S., SONNTAG, A. G., MAERZ, L., GRELLSCHEID, S. N., KREMMER, E., NITSCHKE, R., KUEHN, E. W., JONKER, J. W., GROEN, A. K., RETH, M., HALL, M. N. & BAUMEISTER, R. 2013. Inhibition of mTORC1 by astrin and stress granules prevents apoptosis in cancer cells. *Cell,* 154**,** 859-74.

THIRY, M. 1993. Differential location of nucleic acids within interchromatin granule clusters. *Eur J Cell Biol,* 62**,** 259-69.

TIMALSINA, S., ARIMOTO-MATSUZAKI, K., KITAMURA, M., XU, X., WENZHE, Q., ISHIGAMI-YUASA, M., KAGECHIKA, H. & HATA, Y. 2018. Chemical compounds that suppress hypoxia-induced stress granule formation enhance cancer drug sensitivity of human cervical cancer HeLa cells. *The Journal of Biochemistry,* 164**,** 381-391.

TIMPANO, S., MELANSON, G., EVAGELOU, S. L., GUILD, B. D., SPECKER, E. J. & UNIACKE, J. 2016. Analysis of Cap-binding Proteins in Human Cells Exposed to Physiological Oxygen Conditions. *J Vis Exp*.

TOURRIÈRE, H., CHEBLI, K., ZEKRI, L., COURSELAUD, B., BLANCHARD, J. M., BERTRAND, E. & TAZI, J. 2003. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *J Cell Biol,* 160**,** 823-31.

UNBEHAUN, A., BORUKHOV, S. I., HELLEN, C. U. & PESTOVA, T. V. 2004. Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. *Genes Dev,* 18**,** 3078-93.

UNIACKE, J., HOLTERMAN, C. E., LACHANCE, G., FRANOVIC, A., JACOB, M. D., FABIAN, M. R., PAYETTE, J., HOLCIK, M., PAUSE, A. & LEE, S. 2012. An oxygen-regulated switch in the protein synthesis machinery. *Nature,* 486**,** 126-129.

UPRETY, D. & ADJEI, A. A. 2020. KRAS: From undruggable to a druggable Cancer Target. *Cancer Treat Rev,* 89**,** 102070.

UVERSKY, V. N., GILLESPIE, J. R. & FINK, A. L. 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins,* 41**,** 415-27.

VALÁŠEK, L., NIELSEN, K. H., ZHANG, F., FEKETE, C. A. & HINNEBUSCH, A. G. 2004. Interactions of Eukaryotic Translation Initiation Factor 3 (eIF3) Subunit NIP1/c with eIF1 and eIF5 Promote Preinitiation Complex Assembly and Regulate Start Codon Selection. *Molecular and Cellular Biology,* 24**,** 9437-9455.

VALAVANIDIS, A., VLACHOGIANNI, T., FIOTAKIS, K. & LORIDAS, S. 2013. Pulmonary oxidative stress, inflammation and cancer: respirable particulate matter, fibrous dusts and ozone as major causes of lung carcinogenesis through reactive oxygen species mechanisms. *Int J Environ Res Public Health,* 10**,** 3886-907.

VAN OOIJEN, M. P., JONG, V. L., EIJKEMANS, M. J. C., HECK, A. J. R., ANDEWEG, A. C., BINAI, N. A. & VAN DEN HAM, H. J. 2018. Identification of differentially expressed peptides in high-throughput proteomics data. *Brief Bioinform,* 19**,** 971-981.

VAN TREECK, B., PROTTER, D. S. W., MATHENY, T., KHONG, A., LINK, C. D. & PARKER, R. 2018. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proceedings of the National Academy of Sciences,* 115**,** 2734-2739.

VAN ZUNDERT, G. C. P., RODRIGUES, J., TRELLET, M., SCHMITZ, C., KASTRITIS, P. L., KARACA, E., MELQUIOND, A. S. J., VAN DIJK, M., DE VRIES, S. J. & BONVIN, A. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol,* 428**,** 720-725.

VIPHAKONE, N., HAUTBERGUE, G. M., WALSH, M., CHANG, C.-T., HOLLAND, A., FOLCO, E. G., REED, R. & WILSON, S. A. 2012. TREX exposes the RNA-binding domain of Nxf1 to enable mRNA export. *Nature Communications,* 3**,** 1006.

WAGNER, S., BOHLEN, J., HERRMANNOVA, A., JELÍNEK, J., PREISS, T., VALÁŠEK, L. S. & TELEMAN, A. A. 2022. Selective footprinting of 40S and 80S ribosome subpopulations (Sel-TCP-seq) to study translation and its control. *Nat Protoc,* 17**,** 2139-2187.

WANG, T., TIAN, X., KIM, H. B., JANG, Y., HUANG, Z., NA, C. H. & WANG, J. 2022. Intracellular energy controls dynamics of stress-induced ribonucleoprotein granules. *Nature Communications,* 13**,** 5584.

WEINSTEIN, J. N., COLLISSON, E. A., MILLS, G. B., SHAW, K. R., OZENBERGER, B. A., ELLROTT, K., SHMULEVICH, I., SANDER, C. & STUART, J. M. 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet,* 45**,** 1113-20.

WETTSTEIN, F. O. & NOLL, H. 1965. Binding of transfer ribonucleic acid to ribosomes engaged in protein synthesis: Number and properties of ribosomal binding sites. *Journal of Molecular Biology,* 11**,** 35-53.

WHITE, J. P., CARDENAS, A. M., MARISSEN, W. E. & LLOYD, R. E. 2007. Inhibition of Cytoplasmic mRNA Stress Granule Formation by a Viral Proteinase. *Cell Host & Microbe,* 2**,** 295-305.

WHYTE, D., SKALKA, G., WALSH, P., WILCZYNSKA, A., PAUL, N. R., MITCHELL, C., NIXON, C., CLARKE, W., BUSHELL, M., MORTON, J. P., MURPHY, D. J. & MUTHALAGU, N. 2023. NUAK1 governs centrosome replication in pancreatic cancer via MYPT1/PP1β and GSK3β-dependent regulation of PLK4. *Mol Oncol*.

WILCZYNSKA, A., GILLEN, S. L., SCHMIDT, T., MEIJER, H. A., JUKES-JONES, R., LANGLAIS, C., KOPRA, K., LU, W. T., GODFREY, J. D., HAWLEY, B. R., HODGE, K., ZANIVAN, S., CAIN, K., LE QUESNE, J. & BUSHELL, M. 2019. eIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5'UTR. *Genome Biol,* 20**,** 262.

YAMAZAKI, T., NAKAGAWA, S. & HIROSE, T. 2019. Architectural RNAs for Membraneless Nuclear Body Formation. *Cold Spring Harb Symp Quant Biol,* 84**,** 227-237.

YAMAZAKI, T., SASAKI, N., NISHI, M., YAMAZAKI, D., IKEDA, A., OKUNO, Y., KOMAZAKI, S. & TAKESHIMA, H. 2007. Augmentation of drug-induced cell death by ER protein BRI3BP. *Biochem Biophys Res Commun,* 362**,** 971-5.

YAMAZAKI, T., SOUQUERE, S., CHUJO, T., KOBELKE, S., CHONG, Y. S., FOX, A. H., BOND, C. S., NAKAGAWA, S., PIERRON, G. & HIROSE, T. 2018. Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Molecular Cell,* 70**,** 1038-1053.e7.

YANG, P., MATHIEU, C., KOLAITIS, R. M., ZHANG, P., MESSING, J., YURTSEVER, U., YANG, Z., WU, J., LI, Y., PAN, Q., YU, J., MARTIN, E. W., MITTAG, T., KIM, H. J. & TAYLOR, J. P. 2020. G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell,* 181**,** 325-345.e28.

YOUN, J. Y., DUNHAM, W. H., HONG, S. J., KNIGHT, J. D. R., BASHKUROV, M., CHEN, G. I., BAGCI, H., RATHOD, B., MACLEOD, G., ENG, S. W. M., ANGERS, S., MORRIS, Q., FABIAN, M., CÔTÉ, J. F. & GINGRAS, A. C. 2018. High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol Cell,* 69**,** 517-532.e11.

YOUN, J. Y., DYAKOV, B. J. A., ZHANG, J., KNIGHT, J. D. R., VERNON, R. M., FORMAN-KAY, J. D. & GINGRAS, A. C. 2019. Properties of Stress Granule and P-Body Proteomes. *Mol Cell,* 76**,** 286-294.

ZEMAN, J., ITOH, Y., KUKAČKA, Z., ROSŮLEK, M., KAVAN, D., KOUBA, T., JANSEN, M. E., MOHAMMAD, M. P., NOVÁK, P. & VALÁŠEK, L. S. 2019. Binding of eIF3 in complex with eIF5 and eIF1 to the 40S ribosomal subunit is accompanied by dramatic structural changes. *Nucleic Acids Research,* 47**,** 8282-8300.

ZHANG, X., SMITS, A. H., VAN TILBURG, G. B., OVAA, H., HUBER, W. & VERMEULEN, M. 2018. Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat Protoc,* 13**,** 530-550.

ZHENG, H., ZHAN, Y., ZHANG, Y., LIU, S., LU, J., YANG, Y., WEN, Q. & FAN, S. 2019. Elevated expression of G3BP1 associates with YB1 and p-AKT and predicts poor prognosis in nonsmall cell lung cancer patients after surgical resection. *Cancer Med,* 8**,** 6894-6903.

ZHENG, Z.-Q., WANG, S.-Y., XU, Z.-S., FU, Y.-Z. & WANG, Y.-Y. 2021. SARS-CoV-2 nucleocapsid protein impairs stress granule formation to promote viral replication. *Cell Discovery,* 7**,** 38.

ZHU, A., IBRAHIM, J. G. & LOVE, M. I. 2018. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics,* 35**,** 2084-2092.