Examining the Role of Ataxin-2 Protein and its Domains in Cellular Function, Translational Control and Neurodegeneration

PhD Thesis

2023

Arnas Petrauskas

Supervisor: Prof Mani Ramaswami

Trinity College Institute of Neuroscience,
Smurfit Institute of Genetics

University of Dublin Trinity College

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I acknowledge that the thesis contains work that has been carried out jointly and declare that unpublished and/or published work of others is duly acknowledged and credited in the text wherever included.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).

Arnas Petrauskas
September 2022

Summary

Neurodegenerative diseases represent an increasing burden on an aging society with no cures and limited treatment options. One sub-group of neurodegenerative disease, collectively known as ataxias, affects cerebellar neurons and typically result in movement control impairments. Ataxin-2 (Atxn2) is the causative gene for spinocerebellar ataxia type II (SCA2), and is implicated in amyotrophic lateral sclerosis (ALS) and Parkinsonism. Specifically: an increase in the length of a CAG triplet-repeat tract within the Atxn2 gene has been shown as the causal mutation for SCA2. Subsequently, intermediate length CAG expansions were also implicated in ALS. In addition to causing neurodegeneration directly, Atxn2 may be an important contributor to progression of multiple forms of other degenerative disease. Thus, recent mouse and Drosophila studies show that reducing levels of Atxn2 protein can slow down the progression of ALS and SCA2 disease models.

Atxn2 (or Atx2 in insects) can function in translational activation, translational repression, mRNA stability and in the assembly of messenger ribonucleoprotein granules (mRNPs), a process mediated by intrinsically disordered regions (IDRs). Work from our lab has shown that the LSm (Like-Sm) domain of Atx2, which can help stimulate mRNA translation, antagonizes mRNP-granule assembly. In my thesis work, I and colleagues built on previous research through a series of experiments on *Drosophila* and human Ataxin-2 proteins to identify molecular mechanisms through which the protein normally functions and how alterations in these activities drive neurodegeneration. In particular, I replicate and increase the applicability of our models for the role of the IDR in granule assembly and degeneration (particularly tying observations from fly Atx2 to mammalian Atxn2), as well as highlight a “poly-A tail – polyA-binding protein (PABP) – Ataxin-2 Pam2 domain” interaction driven localisation mechanism for the protein that is important for translational control of mRNAs and necessary degeneration in *Drosophila* models.

Our lab’s work as well as of several others have shown Ataxin-2’s varied roles in translation control as well as the involvement of multiple mechanisms; through both conserved and structured domains of the protein that likely act at the level of individual mRNA/mRNPs as well as via IDRs that bring mRNPs into higher-order assemblies such as stress granules and neuronal granules. Work I contributed to in *Drosophila* has shown that IDRs of the protein, including an endogenous poly-glutamine domain, are

specifically required for both promoting neurodegeneration in fly models of ALS, and for regulating neuronal mRNA translation necessary for consolidation of long-term memory. My experiments provided the key observation that deletions of these IDRs conferred resistance to degeneration in fly models of ALS. We proposed a working model in which IDRs allow Atx2 to form granules that facilitate spatial and temporal translational control required for long-term memory formation, but also inadvertently, in susceptible individuals, support key steps in the progression of neurodegeneration. I further expand on this work by identifying an analogous IDR in human Atxn2 and setting the groundwork of experiments (such comparatively testing the IDR in a cellular granule formation assay, and testing the functional interchangeability of human and fly Atx2 *in vivo*) crucial to the understanding of the protein as a therapeutic target for humans. This is timely given the increased interest in and even clinical trials currently in operation aimed at Atxn2 and its functions regarding granule formation.

Acknowledging that single domains of a protein rarely act in isolation of other domains, or co-factors and binding partners, my work further examines the interplay between the three structured domains (LSm, LSm-AD, and Pam2) of the protein and the functions of Ataxin-2 that are mediated by the IDR. I add to our understanding of the protein and its interactors by examining radically truncated Atx2 constructs in cells and transgenic flies to separate and independently assay the effects of structured and unstructured domains – finding that the Pam2 domain, through its interaction with PABP and thus poly-A mRNA is essential to localising and directing all other activities of the protein. Targets of RNA-Binding Proteins Identified by Editing (TRIBE) experiments as well as co-localization and pull-downs showed that this domain is the major determinant of the mRNA and protein content of Ataxin-2 mRNP granules and thus essential to the translational control of target mRNAs. Meanwhile, transgenic *Drosophila* degeneration assays show the structured PAM2 and unstructured IDR interactions both, but through distinct modes of action, are required for Atx2-induced cytotoxicity.

Altogether, the work I carried out as part of this thesis improves our understanding of the forces at work controlling the sometimes antagonistic functions of Ataxin-2 and its domains. We propose a model whereby Ataxin-2 complexes are remodelled based on context during translational control and how structured and non-structured interactions affect the specificity and efficiency of mRNP condensation. In turn, these forces, that can be moved from equilibrium by disease causing mutations, add up to modulate pathways that may lead to neurodegeneration. This work may prove crucial for suggesting more targeted therapeutic approaches towards Atxn2.

Acknowledgements

I would like to thank everyone who contributed to, supported, and made this work possible. It has been an exciting and at times unpredictable 5 years filled with great joy and great disruption. But from my first internship and throughout the rest of my time here, I have found the Ramaswami lab a truly welcoming and inspiring place to work in and to grow in.

Thank you to my co-workers, collaborators, and indeed friends: Dr. Jörn Hülsmeier, Dr. Jens Hillebrand, Dr. Márcia Aranha, Dr. Amanjot Singh, Dr. Baskar Bakthavachalu, Dr. Ankita Chodankar, and Dr. Swati Trisal. A special thank you to my fellow PhD students Camilla Roselli, Dr. Daniel Fortunati, Guillaume Thuery, and Georgia Brown, together with whom I shared the good times and the bad. This PhD was only possible because of the great help, feedback and bench-side chats that always reignited my curiosity and raised my spirits. Thank you to Dr. Adrian Dervan, Emily Walker, Dr. Isabel Twick, and all the inspiring scientists and students I have met and had the pleasure to work with, learn from, or teach along the way. Thank you to my friends Remsha, Ieva, Christie, Mihai, Craig, Luke and my family – Arūnas, Švitra, Simona and Nerijus – who have been nothing but supportive during this time. Finally, thank you to my supervisor and mentor, Prof. Mani Ramaswami, who's casual wisdom and immense knowledge drove me forward and kept me on the path. I am truly grateful to everyone.

Most importantly, I would like to specifically thank my love, my partner and my closest friend – Dr. Marlena Mucha, whose support and care inspired me to strive and persevere through it all. You are always there when I need you, and you are the source of my happiness. The peace you bring to my life is the only reason I can find the space to grow as a person and pursue the mysteries of science. I could not have made it without you.

These years have been the beginning of my consciousness, and the beginning of what I know will be a life-long love of learning.

Table of Contents

Abbreviations

Ataxin-2 – The Ataxin-2 protein in general, referring to either *Drosophila* or human protein and the shared properties of the homologs and paralogs
Atx2 – *Drosophila* Ataxin-2 protein specifically
Atxn2 – Human Ataxin-2 protein specifically
Atxn2L – Human Ataxin-2 Like protein specifically
SG/SGs – Stress granule/stress granules
S2/S2R+ – Schneider 2 R+ *Drosophila* cultured cell line
U2OS – U2 Osteosarcoma human cultured cancer cell line
HEK293T – Human Embryonic Kidney 293T cultured immortalised cell line
SCA2 – Spino-Cerebellar Ataxia type 2:
ALS – Amyotrophic Lateral Sclerosis
IDR – Intrinsically Disordered Region, either generically or referring to all such regions within a protein
CIDR – The C-terminal Intrinsically Disordered Region of Ataxin-2, consisting of the sequence downstream from the PAM2 domain and up to the stop codon
MIDR – The Middle Intrinsically Disordered Region of Ataxin-2, consisting of the glutamine tract(s) and prion-like sequences, downstream of the LSm-AD but before the PAM2 domains
Poly-Q – Poly-Glutamine repeat
mRNP – messenger ribonucleoprotein
iPSC – induced pluripotent stem cells
LLPS – liquid-liquid phase separation
WB – Western Blot
IP – Immune Pull-Down of proteins
KD – Knock Down of target mRNA and protein, through RNAi
KO – Knock Out of a target gene, through site-directed techniques such as P-element integration or CRISPR/Cas9

Introduction

It is common to start any research work dealing with a topic even somewhat tangential to a condition in humans by stating the personal and financial toll the condition exerts on its sufferers, and often highlight the lack of cures or treatment options for it. However, it is with deep empathy that I would like to make the argument that neurodegenerative diseases truly deserve to be highlighted in such a way due to their varied but similarly insidious pathology. Degenerative conditions, especially later onset ailments are increasing in prevalence as average populations in developed countries continue to get older. Conditions such as Alzheimer's disease slowly take away essential aspects of the fundamental human experience of life: memory and self-awareness. While degeneration of motor neurons such as in Amyotrophic Lateral Sclerosis (ALS) or key motor co-ordination neurons in Spinocerebellar Ataxia type 2 (SCA2) – the conditions tied to the work of this thesis – have the opposite effect of locking-in a healthy mind in a slowly failing body. There are no effective cures for any neurodegenerative condition and even the most advanced therapeutics currently available extend life and/or function by only a few months (such as in the case of Lecanemab, FDA approved for Alzheimer's disease in 2023). My work as part of this PhD thesis focuses on basic research into the subcellular mechanisms of my protein of interest and utilises degeneration models in flies as a means to answer specific, narrow questions. But all the while it is important to humbly remember that there are countless people affected by great suffering who have enduring hope that this field of research will eventually produce a means to give them, or people like them, their lives back. This study of the Ataxin-2 protein and its function contributes but the smallest drop to this vast, dynamic and important field.

The *Ataxin-2* gene (*Atxn2*) was first identified in the context of SCA2 patients (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996). Before any other aspects of the gene and protein were elucidated, this work identified critical mutations in the poly-glutamine (poly-Q) tract unique to the disease-associated allele of *Atxn2* as causative for associated neurodegeneration. Expansions of the normally 23Q tract to ≥33Q were found exclusively in patients of SCA2 and implied at the time to be causative. Subsequently, intermediate expansions were identified to be associated with, and modifiers of severity in ALS (Elden *et al.*, 2010). Increased interest in *Atxn2* led to a focus of research towards the gene and its protein products, firstly showing that it is a highly conserved RNA-binding protein with at least one broadly-expressed homolog

present in all eukaryotes, from humans to *S. cerevisiae* (Satterfield and Pallanck, 2006; Jiménez-López and Guzmán, 2014). This suggested that the protein plays essential biological functions and led to the hypothesis that these normal functions may be altered by mutations to drive neurodegeneration in affected patients. My work as part of this thesis and the work of my colleagues and lab specifically harkens back to this key idea of normal protein function being sabotaged in the disease scenario that eventually leads down to cell degeneration. All our work attempts to add to the understanding of the function of Atxn2 from the scale of molecules to neurodegeneration models in fruit flies. By breaking down the individual forces acting upon the protein through its various domains we strive to attain a model of the equilibrium between them, and how it may be disrupted in SCA2 or ALS patients to drive these modes of degeneration.

Functions of Ataxin-2

Currently it is known that higher level functions of Ataxin-2 across species include cell growth, cell development and neuronal plasticity (Nonis *et al.*, 2008; Sudhakaran *et al.*, 2014; Van den Heuvel *et al.*, 2014). Looking at a molecular level and typical of an RNA-binding protein, Ataxin-2 functions in translational control as both activator and repressor: it is required for translational repression for several mRNAs such as *miR12*, as well as translational activation of others such as *Period* (Nonhoff *et al.*, 2007; McCann *et al.*, 2011; Lim and Allada, 2013; Lee *et al.*, 2017). These are vital biological functions: in organisms such as *Drosophila,* where only one paralog is present, the gene is essential for survival. In humans and mice, two paralogs are present: *Atxn2* and *Atxn2-Like.* Three highly conserved structured domains are present between homologs and paralogs: the LSm, LSm-AD and PAM2 domains (Jiménez-López and Guzmán, 2014). The LSm and LSm-AD domains allow direct binding of the protein to RNA motifs and are involved in protein-protein interactions that regulate RNA translational control performed by Ataxin-2 (Neuwald and Koonin, 1998; He and Parker, 2000; Kufel *et al.*, 2003; Yokoshi *et al.*, 2014). The PAM2 domain on the other hand is essential for interaction with poly(A) binding protein PABP, for enhancement of bound mRNA stability and thus implicates Ataxin-2 in translational activation (Kozlov et al., 2001; Albrecht and Lengauer, 2004; Lim and Allada, 2013; Yokoshi et al., 2014). Further proline-rich regions and motifs commonly seen within the Atx2 protein sequences may play a part in receptor endocytosis and cellular growth signalling pathways that the protein is involved in (Nonis *et al.*, 2008; Auburger *et al.*, 2017). With partner proteins such as TDP-43, InsP(3)R1, Atx-2 binding protein, Staufen1 and FMRP; Ataxin-2 was shown as important for RNA splicing, translational repression,

neuronal mRNA retrograde transport and long-term habituation (Shibata, Huynh and Pulst, 2000; Liu *et al.*, 2009; Sudhakaran *et al.*, 2014; Becker *et al.*, 2017; Bakthavachalu *et al.*, 2018; Paul *et al.*, 2018).

Ataxin-2 functions in a degeneration context

While the roles of Ataxin-2 within the normal function of cells are increasingly understood, its role in degeneration was simultaneously examined. In particular, cellular degeneration, such as the degeneration of cerebellar Purkinje neurons in SCA2 or motor neurons in ALS converges on a number of pathways. Anything that affects mitochondria, for example, will have downstream effects in cell metabolism (limiting ATP), release of reactive oxygen species, DNA damage and cytoplasmic exposure induced apoptosis, or any combination of the mentioned and unmentioned modes of failure that lead to cell death (Ramanan and Saykin, 2013; Gan *et al.*, 2018). In this example, Atxn2 was shown to be involved in mitochondrial function through association with Pink1 in mice (Sen *et al.*, 2016; Auburger *et al.*, 2017). While my work will focus on one model of how Ataxin-2 activity may lead to a pathway of cellular degeneration in disease contexts, it is worth keeping in mind the system of interactions and roles a protein may carry out within cells. Affecting one system or protein may often lead to a systemic collapse of the rest and have homeostasis-wide effects on a cell. Thus mitochondrial usage of Ataxin-2, for example, will always be impacted by mutations that may not primary target this mechanism, but contribute to the overall threshold of factors that drive a cell to die. Similarly, embryonic development and cell growth (Kiehl, Shibata and Pulst, 2000; Drost *et al.*, 2013; Vianna *et al.*, 2016), endocytosis (via endophilin 1/2) (Ralser *et al.*, 2005; Nonis *et al.*, 2008), centrosome and microtubule dynamics (Stubenvoll *et al.*, 2016) as well as Golgi apparatus association and organisation (Huynh *et al.*, 2003) are all cell functions in which an Ataxin-2 homolog has been implicated in at least one organism. These functions in turn correspond to the key ways in which cell degeneration may be driven: impaired differentiation and cell growth, metabolic defects, subcellular transport and organisation errors, as well as protein synthesis and secretion/signalling defects respectively (Ramanan and Saykin, 2013; Gan *et al.*, 2018).

However, the mechanistic link between Ataxin-2 and degeneration becomes particularly clear when considering the role of the protein in translational control. RNA-binding proteins such as Ataxin-2 are highly enriched in screens of degeneration modifying targets and converge on the pathways of translational control, protein quality

control, stress responses, autophagy, and mRNA transport (Hentze *et al.*, 2018; Gebauer *et al.*, 2020). Ataxin-2 fits this profile and impacts several of these mechanisms. Firstly and fundamentally for my work, Ataxin-2 is a stress-granule component (Nonhoff *et al.*, 2007; Buchan, Muhlrad and Parker, 2008; Swisher and Parker, 2010), however in normal conditions, the protein is found in a generally diffuse state in the cytosol. Particularly, the protein is associated with translational machinery in P-bodies and at the rough endoplasmic reticulum (Nonhoff et al., 2007; van de Loo et al., 2009). This alludes to the protein's function in the translation initiation complex together with LSM12 and TYF in *Drosophila* (Lim and Allada, 2013; Lee *et al.*, 2017) or as part of the translation loop or polysome complex (Fittschen *et al.*, 2015; Lastres-Becker *et al.*, 2016) for activating translation of some target mRNAs; while the protein is also a part of a translational repression complex with Me31B and Not1 (Sudhakaran *et al.*, 2014; Lee *et al.*, 2017) or as a key component of the generally mRNA silencing membrane-less organelles known stress granules (SGs) (Kedersha *et al.*, 2005; Bakthavachalu *et al.*, 2018). In subsequent sections, I will examine these functions of Ataxin-2 and pathways of degeneration in particular, and will attempt to expand on this knowledge through experiments carried out as part of this thesis work.

Stress granules, disordered protein regions, and aggregation in degeneration

Together with other co-factors such as G3BP/Rasputin and PABP(C1), Ataxin-2 serves to sequester some target mRNAs and otherwise severely slow down the translation of most mRNAs in the cell for the duration of this stress response (Kedersha *et al.*, 2005; Nonhoff *et al.*, 2007; Lastres-Becker *et al.*, 2016; Khong *et al.*, 2017). Stress granules (SGs) are a type of cytoplasmic granule that forms in eukaryotic cells when they are subjected to cellular stress signals induced by external factors such as heat, amino-acid depletion, viral infection, oxidative stress or addition of chemicals such as sodium arsenite or sorbitol that induce the aforementioned effects (Kedersha *et al.*, 1999; Protter and Parker, 2016; Khong and Parker, 2020). Key to classifying SGs is the presence of essential protein components within these granules (such as G3BP1/2 and PABP) and the involvement of the stress signalling pathway consisting of eIF2α kinase activation that causes individual messenger ribonucleoprotein granules (mRNPs) which are present in unstressed cell cytoplasm to arrest translation and coalesce into multi-mRNP assemblies (Protter and Parker, 2016). Notably, this feature of SGs and the ability to induce them in cell culture assays allows them to serve as a proxy for probing the functions of their mRNP precursors too. To tie these observations to the role of Ataxin-2 in cells and in degeneration, our lab together with Roy Parker and Paul Taylor,

articulated a model where mRNP granules provide micro-environments of concentrated proteins and RNA, where toxic irreversible aggregates could form and in turn lead to altered cell functions that contribute to neurodegeneration and disease progression (Ramaswami, Taylor and Parker, 2013).

This model predicts that drugs or mutations that either slow down RNP granule formation or promote their disassembly would delay neurodegeneration. Precedent for this model has long existed since protein aggregates were first identified in neurodegeneration (Strassnig and Ganguli, 2005). Furthermore, albeit controversially, the amyloid hypothesis of neurodegeneration (originally specific to Amyloid beta in Alzheimer's disease but generalisable to other conditions involving protein aggregates) would suggest that it is the protein aggregates explicitly that cause cellular degeneration and the condition associated with it (Glenner and Wong, 1984). However, an evidence gap remained in tying mRNP or other granule formation to toxic aggregates until critical, *in vitro* experiments, showed the transition between granules – or membrane-less droplets of concentrated proteins held together as a liquid – and prion/amyloid-like protein fibrils for the ALS implicated protein FUS (Patel *et al.*, 2015). Even protein aggregation specifically within SGs has by now been shown in the context of SOD1, another ALS associated protein (Mateju *et al.*, 2017). What both these proteins, along with some other modulators of neurodegeneration such as hnRNPA1 or Tau, have in common are intrinsically disordered regions (IDRs) in their protein structure and the ability to phase separate in the cytoplasm or *in vitro* (Lin *et al.*, 2015; Patel *et al.*, 2015; Mateju *et al.*, 2017; Wegmann *et al.*, 2018).

IDRs allow for a high number of weaker, more promiscuous interactions to take place between these amino acid tracts and structured proteins, cofactors, or RNA/DNA strands. But most importantly, they allow for IDR-IDR interactions that are thermodynamically favourable and drive de-mixing of proteins from their surrounding substrate, leading to liquid-liquid phase separation (LLPS) under the right conditions (Babinchak and Surewicz, 2020). This is a phenomenon that can provide a mechanistic insight into how molecular condensates are generated and controlled within cells and how this control may be upset in the disease state. Folded domains have fewer, specific interactions that may be sterically hindered or may block the binding surface. A structured-only peptide must carry out a high number of them in order to form an assembly, such as an amyloid fibril, Tau neurofibrillary tangle or a prion protein inclusion. Disordered domains have weaker interactions that are more promiscuous and multivalent. A disordered domain may rapidly shift between a large number of

temporarily stable secondary and tertiary structures that may find a local free-energy minimum interacting with one partner while having a very low activation energy to facilitate a shift to another temporary structure with different binding partners. Critically, these weaker interactions may be sufficient to maintain a loose assembly of various partners and by their very nature have a higher valence than stably folded domains. This behaviour is to an extent analogous to the Hydrogen bonds and Van Der Waals interactions in between water molecules, and further analogous to the Gibbs free energy minimising hydrophobic-hydrophilic molecule separation seen between water and oil – giving rise to the term liquid-liquid phase separation. The number of promiscuous interactions in a peptide would more effectively contribute to and drive the formation of a loose assembly of molecules, in conjunction with other multivalent molecules such as RNA. As the assembly is not as tightly bound together as one formed by structured domains alone, it would be more dynamic and similar to a condensate (Protter *et al.*, 2018; Alberti, Gladfelter and Mittag, 2019). IDRs and LLPS will be the key property and mechanism of action examined in a number of experiments for this thesis and are essential to our models of Ataxin-2 function and pathology.

Ataxin-2 shares similarities to the aforementioned proteins: it is an RNA-binding protein with large IDRs within its domain structure, and thus susceptible to the same forces driving its LLPS (Bakthavachalu *et al.*, 2018). In Atxn2 disease pathology, we know that cytoplasmic aggregates are present in the neurons of SCA2 patients (Seidel *et al.*, 2017) and we have evidence form cell overexpression studies of polyQ expanded Atxn2 (as in SCA2), showing cell death via apoptosis in transfected cells caused specifically by this mutation (Huynh *et al.*, 2003; Ng, Pulst and Huynh, 2007). Furthermore, Atxn2 modulation of ALS severity, where TDP-43 containing neuronal inclusions were observed, confirm another scenario where Atxn2 directly affects protein aggregation (Hart *et al.*, 2012). Therefore the involvement of Ataxin-2 protein both in RNP granule formation and neurodegeneration – through an expanded polyQ tract which is implicated in protein aggregation specifically as well as its other disordered regions that may drive LLPS and granule formation – flagged it for further detailed study to test and elaborate on this model of pathology (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996; Bakthavachalu *et al.*, 2018). Meanwhile, key exploratory studies knocking-down or disrupting Ataxin-2 activities in mice and fly models of degeneration have identified it as an important therapeutic target for a variety of neurodegenerative conditions, from SCA2 and ALS to Huntington's disease (Kim *et al.*, 2014; Becker *et al.*, 2017; Scoles *et al.*, 2017; Huelsmeier *et al.*, 2021).

The role of phase separation and biomolecular condensates in normal cell and organism function

Evolution is astonishing in the efficiency with which physical laws, or biochemical principles affecting the building blocks of life are utilised for organismal and cellular functions. Biomolecular condensates are one such example and have been understood in recent years as being ubiquitous in biology. Superenhancer puncta on actively transcribed chromatin regions act like crowding agents for the co-factors and transcriptional machinery required for efficient catalysis of this reaction. The nucleolus acts in an analogous way as a means to localise the materials for rRNA production and assembly of the subunits of ribosomes. Similarly, paraspeckles and Cajal bodies also within the nucleus are loose assemblies of various co-factors for specific control of RNAs and gene transcription (Banani *et al.*, 2017; Verdile, De Paola and Paronetto, 2019). All these membraneless organelles rely on the previously discussed principles of biomolecule de-mixing in order to generate these distinct functional granules, and are clearly indispensable to the efficient function of the essential reactions they catalyse or facilitate. While many of the biological processes that occur in and around membraneless organelles may still take place without this phenomenon, the localisation of the components in time and space within the cell must increase the efficiency of reactions between them in order to be as conserved and ubiquitous as it is reported (Korkmazhan, Tompa and Dunn, 2021; Tartakoff *et al.*, 2022). Through the course of evolution of the eukaryotic cell, tight biological control over how, when and where these granules form had to arise, as well as precise mechanisms of content control.

With this in mind, we will be focusing more on the somatic cell cytosolic granules that are directly relevant to the questions studied in this thesis. P-bodies are one such group of granules found in the cytosol and act as the localised sites where degradation and processing for a specific subset of mRNAs in the cytoplasm takes place, but not exclusively (Kedersha *et al.*, 2005; Franks and Lykke-Andersen, 2008). Interestingly, P-bodies contain known Ataxin-2 interactors such as DDX6/Me31B whereas Ataxin-2 itself is more clearly associated with stress granules. These well studied granules are induced by the internal or external factors such as viral infection, oxidative stress or heat (endoplasmic reticulum stress), heme deprivation, and amino acid deprivation and contain stalled polysomes on a large subset of translationally repressed mRNAs not essential to the stress signalling pathway, cell survival or recovery (some chaperone mRNAs, for example, are excluded). It is speculated that such a shut-down of general

translation in the cytosol evolved to increase the ability of a stressed cell to survive until the stressing stimulus passes (Protter and Parker, 2016). Uncontrolled translation of proteins in this case could lead to increased errors and misfolding, or potentially subtract resources and machinery from the highly transcriptionally and translationally activated stress response genes. The stress stimuli each have a corresponding eIF2α kinase in mammals that gets activated upon sensing stress: PERK for oxidative, heat and endoplasmic reticulum stress, PKR for viral infection, HRI for heme deprivation, and GCN2 for amino acid deprivation. In turn, phosphorylated eIF2α, in addition to activating the transcription of stress response genes through ATF4, initiates Cap dependent translation arrest of target mRNAs in the cytosol. Ataxin-2 is found on these cytosolic mRNAs and at the time of cell-wide translational arrest, acts together with stalled ribosomes, partner RNA-binding and IDR containing proteins such as G3BP/Rasputin, and the intrinsic self-assembly of mRNA molecules to form SGs. The weak multivalent interactions between IDR-IDR, IDR-structured domains, RNA-RNA, and IDR-RNA, as well as stronger structured interactions between proteins and between proteins and RNA pull together and hold these granules until the cell no longer senses stress and they are disassembled (Kedersha, Ivanov and Anderson, 2013; Protter and Parker, 2016; Alberti *et al.*, 2017; Banani *et al.*, 2017; Khong and Parker, 2020). While such is the beneficial role of stress granules in cells, we know that this function of granule forming proteins such as Ataxin-2 is susceptible to sabotage in the disease context to drive the assembly of pathogenic protein aggregates (Ramaswami, Taylor and Parker, 2013; Becker *et al.*, 2017; Scoles *et al.*, 2017; Bakthavachalu *et al.*, 2018).

Finally, it is worth considering the more subtle utilisation of phase separation activity for the creation of mRNA transport granules, mRNPs and creation of microenvironments for specific cell biological tasks. Ataxin-2, as alluded to earlier, is one of the RNA binding proteins found on both translating mRNAs (such as *Period*) as well as on repressed RNAs in the cytosol (such as *CamKII*) (Lim and Allada, 2013; Sudhakaran *et al.*, 2014). As part of these complexes, it is likely the IDRs of Ataxin-2 contribute to some form of condensed microenvironment that could be defined as a messenger Ribonucleoprotein granule (mRNP), which in turn gets specifically repurposed in some cases as a transport mRNP for localised translation (Hirokawa, 2006; Abouward and Schiavo, 2020). Both transport and other types of mRNPs are not fully characterised however it is clear from the trends in the architecture of RNA binding proteins found in them that IDRs and therefore phase separation ability are indispensable for their varied functions and activities that take place in them. Similarly, phase separated

microenvironments have been described as essential within the central passage of the nuclear pore complex in order to catalyse transport of molecules through them. Nucleoporins (in particular Nup62) that make up the inner ring of the pore contain IDR tails that are specifically organised to fill the opening – in a fine example of tight biological control of an intrinsically promiscuous biological activity. Furthermore, in human neurodegenerative disease and animal models (for example ALS caused by FUS), defects in nucleocytoplasmic transport are observed and may betray the underlying pathogenicity when control over the nuclear pore microenvironment is lost (Kim and Taylor, 2017).

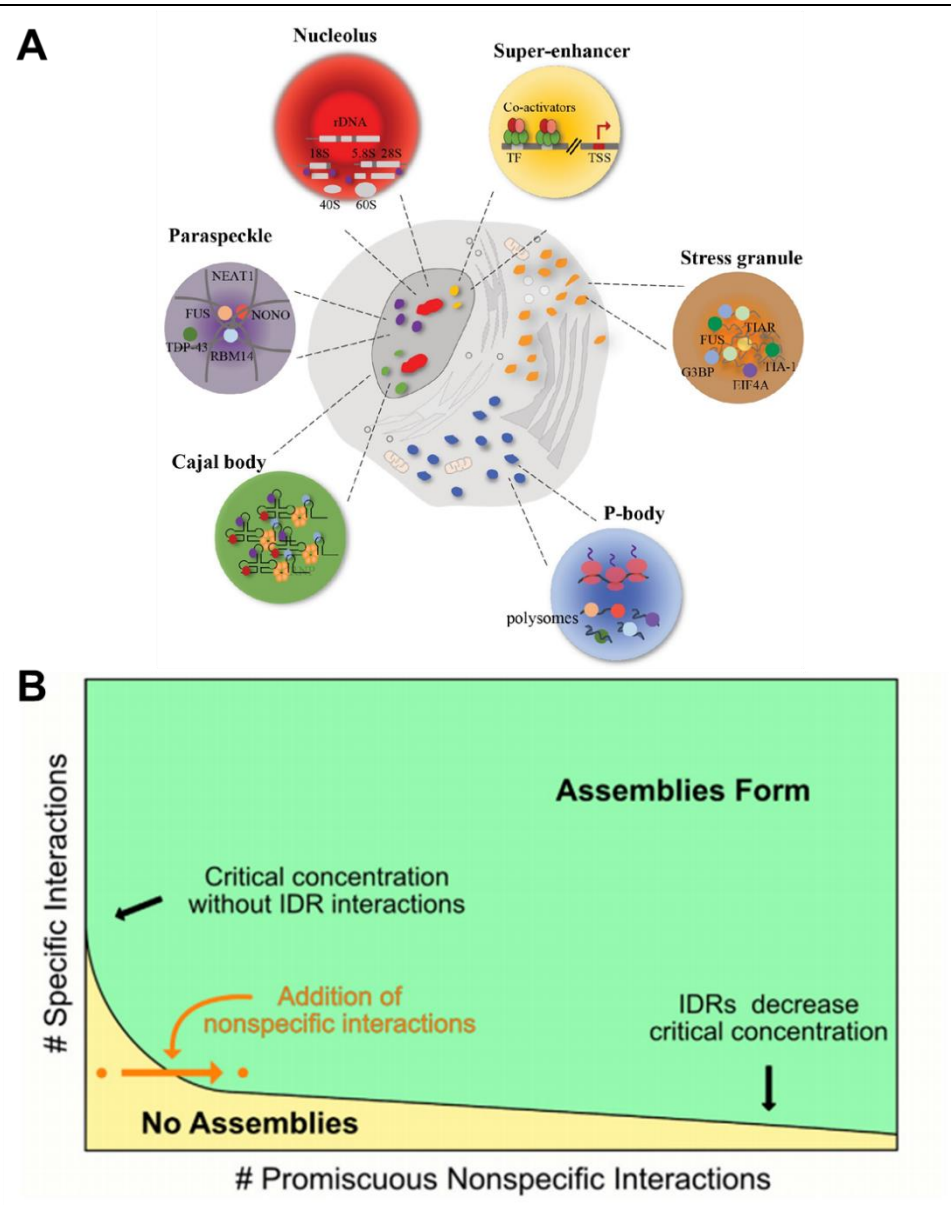

Figure 1: The variety of membraneless organelles/granules within the eukaryotic cell, and a summary phase diagram of the interactions between biomolecules required for

granule assembly. (A) Eukaryotic cells contain a large number of compositionally distinct and specifically localised membraneless organelles and granules that are tightly controlled (such as stress granules that arise only in response to stress signalling pathway activation). These compartments catalyse or facilitate various biological functions within the cell, showing the wide utilisation of the phenomenon of liquid-liquid de-mixing in cells by evolution. Figure adapted from (Verdile, De Paola and Paronetto, 2019). (B) A hypothetical phase transition diagram showing the main principle behind the intrinsic drive of multi-valent, promiscuously interacting disordered regions of a protein to "assemble" or drive phase separation. Structured interactions are stronger, but less valent. Promiscuous interactions are weaker, but more readily interact with more targets. A theoretical critical number of interactions between biomolecules must take place for assemblies to coalesce from a diffuse mixture in solvent. For a protein with some structured interactions, an addition of a disordered region could provide a sufficient total number of potential interactions that liquid-liquid de-mixing could lead to the separation of the protein from the solvent. Figure adapted from (Protter *et al.*, 2018).

Drosophila Ataxin-2 and the role of its IDRs

Our lab focuses on *Drosophila melanogaster* Ataxin-2 (Atx2). In fruit flies there is one genetic ortholog to mammalian *Atxn2* and *Atxn2L,* located on chromosome III (Satterfield, Jackson and Pallanck, 2002). Homozygous deletion of the *Atx2* gene in Drosophila has been reported as being lethal during embryo development (no larvae are reported hatching from eggs) from the moment the first mutants were generated (Satterfield, Jackson and Pallanck, 2002). The structured LSm, LSm-AD, and PAM2 domains – as well as the CAG repeat tract (polyQ) – are highly conserved between human and fly Ataxin-2 (Jiménez-López and Guzmán, 2014). Work I contributed to as part of this thesis determined that the protein contains IDRs that include the disease associated polyQ tract within a prion-like domain referred to as the middle-IDR (mIDR) as well as a highly disordered C-terminal region (cIDR) that are not sequence-conserved between species but instead conserved in their relative placement and disordered structure (Bakthavachalu *et al.*, 2018). Knowing the importance of IDRs or prion-like domains to the biophysical process of liquid-liquid phase separation and degeneration we showed that the IDRs of fly Atx2 are necessary for the ability of the protein to efficiently form mRNP granules but, remarkably, not required for

development, survival and fertility of the animals. While homozygous flies carrying only IDR deleted versions of Atx2 appeared normal in most ways assayed, they showed specific defects in long-term (but not short term) olfactory habituation and, strikingly, reduced neurodegeneration in two fly eye models of ALS (GMR-Gal4 driven UAS C9Orf72 GR50 and GA50 dipeptide repeat overexpression, and UAS human FUS overexpression) (Bakthavachalu *et al.*, 2018). Further examination of how generalisable the modulating effect of Atx2 CIDR deletion is in other fly models of human neurodegeneration were carried out by our lab. We were able to determine that Huntingtin exon1 polyQ expanded peptide overexpression pathology is also made more severe by the presence of the CIDR and alleviated in homozygous deleted lines (Huelsmeier *et al.*, 2021). The generic granule forming ability of the promiscuous, multi-valent CIDR is observed as a node in the aggregation and pathology of a number of different, unrelated diseases modelled in *Drosophila.* Hence, we proposed a model where the ability of Ataxin-2 to form granules is linked with the localized RNA translation needed for long-term memory formation and possibly, for survival under conditions of stress. This at least partially explains the conservation of IDR elements despite their detrimental role in the enhancement of toxicity in individuals predisposed to neurodegeneration (Ramaswami, Taylor and Parker, 2013).

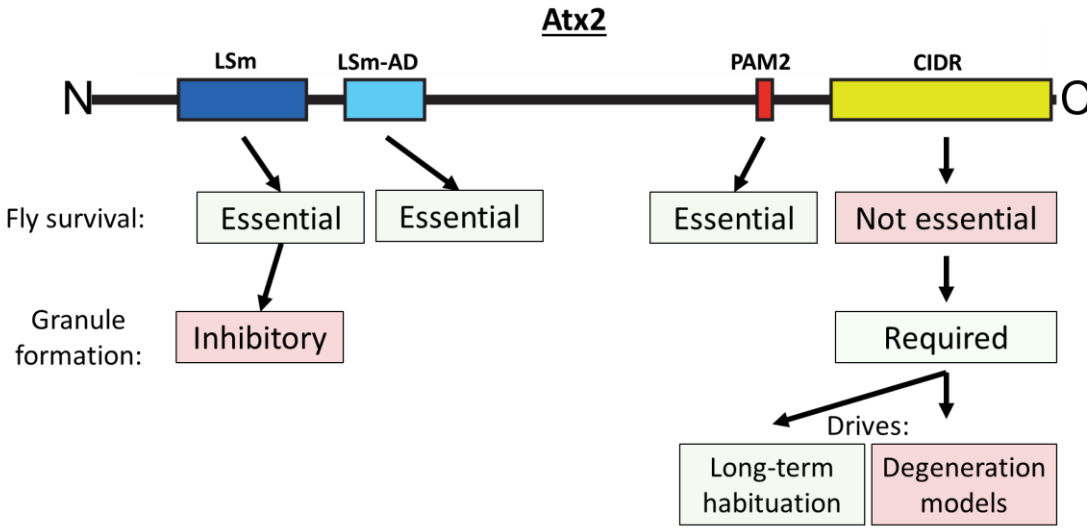

Figure 2: Schematic of the domains of Atx2 and a summary of their functions in *Drosophila* assays. The three structured domains of Atx2 are shown in the schematic (LSm, LSm-AD, and PAM2) as well as the CIDR. Data from Bakthavachalu *et al.* 2018 shows that the three structured domains are required for (homozygote) fly viability, while the homozygote endogenous Atx2 CIDR deleted flies remain viable and fertile. PAM2 deleted Atx2 homozygous flies are incompletely lethal as a low

level of escaper flies were observed, with <5% of the expected progeny numbers surviving. Furthermore, the same paper shows that the CIDR of Atx2 is required for granule formation in S2R+ cell assays. Data from Singh *et al.* 2022 subsequently show that the LSm domain, on the other hand, represses the granule forming ability of Atx2 in S2R+ cell assays. Finally, Bakthavachalu *et al.* 2018 and Huelsmeier *et al.* 2021 detail that the CIDR of Atx2 facilitates long-term olfactory habituation in *Drosophila* Y-maze assays while also being required to drive the severity of two ALS and a Huntington's disease models in flies. Deleting the CIDR alleviated the phenotypes of these degeneration models.

Specific Aims

The central questions of my research fall into two categories corresponding to the chapters of this thesis. Firstly, to confirm and investigate the effect of the IDRs of Atx2 on its molecular and organismal functions, while trying to tie mechanisms observed in *Drosophila* Atx2 with human Atxn2 protein. Defining and examining a functionally homologous C-terminal IDR in human Atxn2 would translate our model of how this region promotes granule formation in cells – to the human cell scenario. Ataxin-2, through the action of its IDRs may have the potential to seed pathogenic protein aggregation and therefore an elucidation of this specific activity may potentially inspire therapeutic approaches, some of which are being currently developed based on whole-protein KD data. I will study this broad issue *in vivo* in flies and in various insect and human cellular models, while further detailing how Atx2 IDRs modulate the proteins' functions and promote neurodegeneration.

Secondly, to examine and detail the interactions of the structured domains of Ataxin-2 (human and fly) and probe how they may be affecting or modifying the functions driven by the IDR. As the IDR has no definite conformational shape and many promiscuous interactions, identifying structured domains of Ataxin-2 that may be modulating the IDRs' effects on degeneration – for example – may reveal mechanisms of how the protein functions in disparate complexes and provide a more concise target interaction for therapeutics screens. I will utilise radically truncated Atx2 and Atxn2 constructs as well as single domain deleted mutants in cultured cell granule forming assays, immunoprecipitation experiments, and, with collaborators, target RNA-sequencing and fly neurodegeneration assays to define structured domains that modify either the normal or pathologic function of Ataxin-2.

Taken together, I hope to propose a model of how Ataxin-2 interactions are remodelled during translational control and how structured and non-structured interactions contribute differently to the specificity and efficiency of RNP granule condensation as well as to neurodegeneration. The protein may be recruited to different contexts and be part of different or remodelled complexes depending on the equilibrium of internal forces coming from its domains and regions as well as external signals and interactors. This model not only aims to explain how normal functions such as mRNP formation get sabotaged by disease mutations to drive neurodegeneration, but also seeks to provide inspiration for future therapeutic approaches.

Materials and Methods

Reagents table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent (*D. melanogaster*)	*UAS-Atx2-WT-ADARcd*	Singh *et al.* 2021	N/A	
Genetic reagent (*D. melanogaster*)	*UAS-Atx2-ΔLsm-ADARcd*	Singh *et al.* 2021	N/A	
Genetic reagent (*D. melanogaster*)	(*UAS-Atx2-ΔPAM2-ADARcd*)	Singh *et al.* 2021	N/A	
Genetic reagent (*D. melanogaster*)	*UAS-Atx2-ΔcIDR-ADARcd*	Singh *et al.* 2021	N/A	
Genetic reagent (*D. melanogaster*)	*UAS-Atx2-only Lsm/Lsm-AD - ADARcd*	Singh *et al.* 2021	N/A	
Genetic reagent (*D. melanogaster*)	*Mef2-Gal4; Tub-Gal80^{ts}*	Bloomington *Drosophila* Stock center		
Genetic reagent (*D. melanogaster*)	*Elav-Gal4; Tub-Gal80^{ts}*	Bloomington *Drosophila* Stock center		
Cell line (*D. melanogaster*)	S2R+ cells	DGRC	RRID:CVCL_Z831	
Cell line (Human)	HEK293T	Gift: Adrian Bracken lab	N/A	
Cell line (Human)	U2OS	Gift: Martina Schroeder lab	N/A	
Recombinant DNA reagent	pUASt-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly WT Atx2-SNAP
Recombinant DNA reagent	pUASt-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human WT Atxn2-SNAP

Recombinant DNA reagent	pUASt-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly mini Atx2-SNAP
Recombinant DNA reagent	pUASt-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human mini Atx2-SNAP
Recombinant DNA reagent	pUASt-ΔLSm-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly ΔLSm mini Atx2-SNAP
Recombinant DNA reagent	pUASt-ΔLSm-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human ΔLSm mini Atxn2-SNAP
Recombinant DNA reagent	pUASt-ΔLSmAD-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly ΔLSm-AD mini Atx2-SNAP
Recombinant DNA reagent	pUASt-ΔLSmAD-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human ΔLSm-AD mini Atxn2-SNAP
Recombinant DNA reagent	pUASt-ΔPAM2-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly ΔPAM2 mini Atx2-SNAP
Recombinant DNA reagent	pUASt-ΔPAM2-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human ΔPAM2 mini Atxn2-SNAP
Recombinant DNA reagent	pUASt-L859A-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly PAM2* L859A mini Atx2-SNAP
Recombinant DNA reagent	pUASt-L914A-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human PAM2* L914A mini Atxn2-SNAP

Recombinant DNA reagent	pUASt-F866A-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly PAM2* F866A mini Atx2-SNAP
Recombinant DNA reagent	pUASt-F921A-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human PAM2* F921A mini Atxn2-SNAP
Recombinant DNA reagent	pUASt-L859A-F866A-mini-Atx2-SNAP_fly (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express fly PAM2* L859A and F866A mini Atx2-SNAP
Recombinant DNA reagent	pUASt-L914A-F921A-Q928A-mini-Atxn2-SNAP_hum (Plasmid)	Petrauskas and Fortunati *et al.* BioRxiv, 2022	N/A	Construct to express human PAM2* L914A, F921A and Q928A mini Atxn2-SNAP
Recombinant DNA reagent	pAcman-Atx2-GFP (Fosmid)	Sudhakaran et al. 2013	N/A	Genomic construct to express fly WT Atxn2-GFP
Recombinant DNA reagent	pAcman-ΔLSm-Atx2-GFP (Fosmid)	Sudhakaran et al. 2013	N/A	Genomic construct to express fly ΔLSm Atxn2-GFP
Recombinant DNA reagent	pAcman-ΔPAM2-Atx2-GFP (Fosmid)	Sudhakaran et al. 2013	N/A	Genomic construct to express fly ΔPAM2 Atxn2-GFP
Recombinant DNA reagent	pActin-Gal4	DGRC		Actin promoter Gal4 for insect UAS expression
Recombinant DNA reagent	pCMV-Gal4	Addgene	#24345	CMV promoter Gal4 for mammalian UAS expression
Antibody	Anti-Atx2 (chicken polyclonal)	Bakthavachalu et al., 2018	N/A	IF (1:1000) WB (1:1000)

Antibody	Anti-Caprin (rabbit polyclonal)	Papoulas et al., 2010	N/A	IF (1:1000)
Antibody	Anti-dFMR (mouse monoclonal)	DSHB	# 5A11	IF (1:1000) deposited to the DSHB by Siomi, H.
Antibody	Anti-GFP (chicken polyclonal)	Abcam	Cat# mAb 13970	IF (1:1000)
Antibody	Anti-V5 (rabbit polyclonal)	Santa Cruz Biotechnology	Cat# sc83849-R	IF (1:1000) WB (1:1000)
Antibody	PABP (rabbit polyclonal serum)	Lee et al. 2017	N/A	IF (1:500) WB (1:500)
Antibody	Me31B (rabbit polyclonal serum)	Lee et al. 2018	N/A	IF (1:500)
Antibody	Rox8 (rat polyclonal)	Buddika et al. 2020	N/A	IF (1:1000)
Antibody	SNAP (rabbit polyclonal)	NEB	Cat# P9310S	WB (1:1000)
Antibody	PABPC1 (rabbit polyclonal)	Abcam		WB (1:1000)
Antibody	LSM12 (rabbit anti human polyclonal)	Abcam		WB (1:1000)
Antibody	G3BP (mouse monoclonal)	BD Bioscience	Cat# 611126	IF (1:1000)
Antibody	Histone H3 (mouse polyclonal)	Abcam		WB (1:1000)
Antibody	Baf155 (mouse polyclonal)	Abcam		WB (1:1000)
Antibody	HRP Goat anti-rabbit	Invitrogen		WB (1:10,000)
Antibody	HRP Goat anti-mouse	Invitrogen		WB (1:10,000)

Antibody	Alexa Fluor 555 (polyclonal goat anti-chicken IgG)	Invitrogen	Cat# A21437	IF (1:1000)
Antibody	Alexa Fluor 488 (polyclonal goat anti-chicken IgG)	Invitrogen	Cat# A11039	IF (1:1000)
Antibody	Alexa Fluor 647 (polyclonal goat anti-chicken IgG)	Invitrogen	Cat# A21449	IF (1:1000)
Antibody	Alexa Fluor 555 (polyclonal goat anti-rabbit IgG)	Invitrogen	Cat# A21428	IF (1:1000)
Antibody	Alexa Fluor 488 (polyclonal goat anti-rabbit IgG)	Invitrogen	Cat# A11078	IF (1:1000)
Antibody	Alexa Fluor 647 (polyclonal goat anti-rabbit IgG)	Invitrogen	Cat# A21244	IF (1:1000)
Antibody	Alexa Fluor 555 (polyclonal goat anti-mouse IgG)	Invitrogen	Cat# A21422	IF (1:1000)
Antibody	Alexa Fluor 488 (polyclonal goat anti-mouse IgG)	Invitrogen	Cat# A21121	IF (1:1000)
Antibody	Alexa Fluor 647 (polyclonal goat anti-mouse IgG)	Invitrogen	Cat# A21235	IF (1:1000)

Chemical compound	MOWIOL mounting medium	Sigma (Merck)	Cat# 81381	
Chemical compound	SNAP-TmrStar	New England Biolabs		IF (1:1000)
Chemical compound	SNAP-Surface 488	New England Biolabs		IF (1:1000)
Software, algorithm	TRIBE	McMahon et al., 2016	https://github.com/rosbashlab/TRIBE	
Software, algorithm	STAR v2.5.3	Dobin et al., 2013	https://github.com/alexdobin/STAR	
Software, algorithm	HTSeq v0.11.2	Anders et al., 2015	https://github.com/htseq/htseq	
Software, algorithm	DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html	
Software, algorithm	ARE*Score*	Spasic et al., 2012	http://arescore.dkfz.de/arescore.pl	
Software, algorithm	Guitar	Cui et al., 2016	https://bioconductor.org/packages/release/bioc/html/Guitar.html	
Software, algorithm	Bedtools	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2	
Software, algorithm	twoBitToFa	-	https://genome.ucsc.edu/goldenPath/help/twoBit.html	
Software, algorithm	MEME suite	Bailey et al., 2009	http://meme-suite.org/tools/meme	
Software, algorithm	ImageJ	Schneider et al., 2012	https://imagej.nih.gov/ij/	
Software, algorithm	Ggplot2	Wilkinson, 2011	https://github.com/tidyverse/ggplot2	
Software, algorithm	Pheatmap		https://cran.r-project.org/web/packages/pheatmap/index.html	

Software, algorithm	SnapDragon		https://www.flyrnai.org/snapdragon

Molecular cloning

Cloning of constructs and plasmids was carried out through combination of restriction/ligation and Gibson cloning protocols. In brief, plasmid maps were analysed through the online gene viewer and tool Benchling (www.benchling.com). Plasmid maps were confirmed with common primer Sanger sequencing where necessary, but particularly around multiple cloning sites, sites of previous insertion or promoter regions. Restriction enzymes were used to digest plasmids and DNA products as per required application. PCR reactions were run with plasmids/DNA fragments if a sequence needed to be amplified, while using sequence added or mutagenesis primers allowed for manipulation of the product sequence. TBE-agarose gels were used to confirm sizes of DNA sequences, while NEB gel extraction kits and protocols were used to cut-out and purify specific amplified sequences for downstream applications. Ligations were carried out using either sticky-end/blunt end ligation (NEB QuickLigase), Gibson assembly or Gateway cloning protocols depending on plasmid backbones and individual construct constraints. Gene fragments were synthesised and codon optimised by GenScript Ltd or BioCat GmbH. Cloning oligos and primers are generated by Eurofins Genomics Ltd. Empty plasmid backbones or specific usable or source constructs were purchased from the Addgene repository, Drosophila Genome Resource Centre, or cDNA repositories such as Genscript Ltd. Constructs are sequence verified by PCR/restriction digestion and Sanger sequencing (Source Biosciences Ltd.). We would like to acknowledge and thank Georg Auburger lab in particular for providing a large number of Ataxin-2 plasmids at the start of my PhD work.

Cell culturing and transfection

Drosophila S2R+ cells were obtained from DGRC, Indiana University and were grown in Gibco Schneider's S2 media (ThermoFisher) with 10% FBS, 1x penicillin and streptomycin, at 25°C, according to the protocols in in Ceriani, 2007; and Fundaci, 2014. Transfections were performed using either FugeneHD (Active Motif) or more efficiently with TransIT-X2 (Mirus) reagents at 1:2 ratio µl reagent to µg plasmid DNA for 24-72H depending on subsequent applications and plasmids used.

HEK293T (human embryonic kidney) cells were kindly gifted by the Adrian Bracken lab, TCD, and grown in Gibco Dulbecco's Modified Eagle Media (ThermoFisher) with

10% FBS, 2mM l-glutamine addition, penicillin and streptomycin, at 37°C and 5% CO2. U2OS (osteo sarcoma) cells were kindly gifted by the Martina Schroeder lab, Maynooth University, and grown at the same conditions as HEK293T. Culturing protocols were adapted from the Farnham lab protocol and as detailed by Kedersha & Anderson, 2007. Mammalian cell transfections were carried out with 1mM PEI (Polysciences) solution at 1:2 ratio µl reagent to µg plasmid DNA for 24-72H depending on downstream use.

For confocal imaging applications cells were grown in 24-well plates on glass cover-slips for 24H before transfection for up to 48H. For Western blotting and IP, cells were grown in 75cm2 flasks until >80% confluent before transfection for up to 72H before harvesting.

Inducing oxidative stress in cell culture

Oxidative stress was induced in Drosophila S2R+ cells with addition of sodium arsenite solution to a final concentration of 50µM in media for 3H. In mammalian cells, oxidative stress was induced in the same way except for only 1H. Following treatment with sodium arsenite solution cells were immediately fixed with 4% para-formaldehyde solution in PBS, or pelleted and flash-frozen – depending on the application. For stress recovery assays, after treatment with sodium arsenite solution cells were washed with fresh media three times after contaminated media was aspirated and disposed of separately as cytotoxic waste, before aforementioned fixation or pelleting.

Live imaging of cells

S2 cells were live-imaged on a custom made slide adaptor. Briefly, a standard microscope slide was affixed with two or more layers of thin, double-sided adhesive tape with circular holes slightly smaller than the coverslips used for growing adherent cells. This way, the glass slide would have a small reservoir surrounded with a ridge of adherent tape into which a S2++ media can be pipetted. When the media reservoir is filled to the top, a coverslip with cells from a multi-well plate can be mounted on to the adherent tape upside-down, so that cells adhering to the top of the coverslip while growing will be suspended upside-down and covered by the media on the reservoir on the slide preparation. The coverslip would be pressed to the adherent tape, cleaned on the exposed surface with a wipe and, if oil immersion lenses are to be used for imaging, sealed on the sides with small quantities of nail varnish or other resin to prevent mixing of media and oil. All procedures except imaging were carried out in a sterile laminar flow cabinet using sterilised constituent materials. The live cell preparation is a single use only and, depending on rigour while setting up, sufficient for

keeping cells in a healthy condition for up to 12H post preparation, after which it is disposed of as biologically contaminated, sharps waste. Where live-imaging of cells relies on cell-membrane permeable dyes, such as SNAP/HALO ligands (TMR-Star, Oregon Green) or Hoechst stain, these dyes must be added to the growing cells in the multi-well plate before they are mounted for live imaging (for approximately 1H at recommended concentrations, followed by a single media wash), however dyes can be added to the media in the preparation as well depending on the application (e.g. Dapi to detect cell death as it is not membrane permeable). Live cell imaging was carried out in the same way as for fixed samples, while FRAP experiments were carried out and analysed by the FRAP/bleaching toolkit and interface within the Zeiss ZEN Black software suite.

Western blotting and protein immunoprecipitation

Total protein extracts were prepared from S2 and HEK cells by collecting pellets of aspirated cells after 3days growth in 75cm3 flasks in 12ml appropriate media at optimal growth temperature as described in the cell culturing section (approximately 1x10e7 mammalian cells and 5x10e7 S2 cells). Suspended media were spun in 15ml sterile falcon flasks for 5mins at 1.5K RCF, and the supernatant decanted. The pellets were washed with 10ml ice-cold sterile PBS, vortexed briefly, centrifuged again and the supernatant was decanted. This was repeated three times in total. The remaining pellets were flash-frozen in liquid nitrogen before storage at -20C.

Cells were lysed in high salt buffer with 0.1% NP-100 replacement or 0.1% Triton X-100 by sonication; or in diluted and complemented Chromotech lysis buffer on ice by aspiration (p200 pipette with tip, 3x times for 1min with 10min wait in between aspirations).

Up to 10 µg total protein in 1X sample loading buffer (SDS for BioRad system, LDS for ThermoFisher iBlot2) was loaded per well for detecting Atx2-SNAP constructs, partner proteins and loading controls on 8-12% SDS-PAGE gels and transferred to nitrocellulose membranes. The blots were probed in 5% skim milk in PBS using rabbit anti-SNAP (1:1000), rabbit anti-PABP (1:1000), rabbit anti-LSM12 (1:1000) antibodies, and mouse anti-histone H3 (1:5000) and mouse anti-BAF155 (1:2000) loading control antibodies. Corresponding HRP-conjugated secondary antibodies were used at 1:10,000 dilution and developed using Pierce ECL western blotting substrate (ThermoFisher) as per the manufacturer's instructions. Imaging was carried out using a LiCor or FujiFilm camera and darkbox systems.

For Atx2-SNAP construct immunoprecipitation, transfected cell lysates were normalised to the same volume and concentration using the Bradford test (BioRad protein assay solution), 10% of the volume was saved and diluted as an input control, and Chromotek anti-SNAP-tag conjugated agarose beads and IP kits were used according to the manufacturer's specifications. Elution of proteins was done either with citric acid (followed by neutralisation) or by boiling beads in sample loading buffer. Pulled-down proteins together with corresponding sample input controls were blotted as described above.

Immunohistochemistry and imaging of cultured cells

Transfected cells on coverslips were fixed with 4% para-formaldehyde in PBS solution for 15mins, followed by three 5min washes in PBS. Permeabilization was performed on all cells with 0.5% TritonX100 in PBS solution for 3mins, before three more 5min washes in PBS. Cells were blocked with 3% BSA in PBS for 1H at room temperature before staining with primary antibodies at appropriate dilutions in 3% BSA overnight at 4°C. Corresponding fluorescent secondary antibodies in 3% BSA were used to stain the sample cells for 1H at room temperature after primaries were washed off. Where SNAP-tagged proteins were being visualized, SNAP-ligand TMR-Star (NEB) or SNAP-surface-Alexa488 (NEB) were added at the secondary antibody staining stage. Following staining and washing, cells were mounted upside-down on microscopy slides in MOWIOL, allowed to cure for >12H at 4°C, and imaged on a Zeiss LSM880 Airyscan/AiryscanFast confocal microscope with a 20x air objective. Live-cell imaging is performed on glass slides modified with adhesive spacers to accommodate coverslips with adherent cells as well as culturing media with cell-permeable fluorophore probe dilutions. Specialist dyes employed in cell experiments include Proteostat (protein aggregate stain, Enzo Life Sciences Ltd.); CellNavigator RNA, endoplasmic reticulum and lysosome stains (AAT Bioquest Ltd.); and Thioflavin-T (amyloid stain, Sigma-Aldrich Ltd.). Bleaching is performed using either 488nm or 555nm lasers on the confocal system.

Bioimage analysis

Where relevant, Airyscan images were processed with Zen Black software (Zeiss) with recommended settings. Confocal microscopy images were analysed using macros within ImageJ/FIJI and Excel. Quantification of co-localisation was performed by comparing stress granule marker staining intensity profile across a random selection of Atx2 granules within transfected cells, with the intensity profile of the Atx2 staining. Any signal 10% or higher than background (adjusted for fluorophore bleed through) was

deemed evidence of co-localisation within that particular granule. This was reversed when looking for exclusion of mini-Atx2-SNAP constructs from stress induced granules – the Caprin or G3BP1 staining was used as independent identifier of stress granules and Atx2 profiles were compared to them. 48-120 granules were quantified in each co-staining , and 28-70 granules were quantified for each construct transfection (Figure 4).

Rearing, crossing, and genome edited line generation of Drosophila stocks

Flies were grown at room temperature in a separate fly room in flugged vials with polenta and golden syrup based solid food. Progeny generations were transferred into new vials to propagate and/or maintain fly lines at most once every month. To generate a genetic cross between two fly lines, only males or only virgin females of the required genotype and phenotype were selected from the population of that line of flies under CO2 anaesthesia and placed into a separate vial with their corresponding mates of the other line being crossed. After 3 days, the parental generation is transferred to another vial or euthanised in order to not contaminate the progeny. Flies not generated as part of this thesis, generated by the Ramaswami lab or previously acquired by the Ramaswami lab were bought from the Bloomington, Vienna or Kyoto repositories or sent as gifts from collaborating labs and are detailed in the reagents table.

Drosophila control strain (CantonS, CS) flies were engineered using CRISPR to replace the endogenous Atx2 locus. Atx2 locus replaced flies published in Bakthavachalu *et al.*, 2018 were generated by members of the Ramaswami and VijayRaghavan labs and utilised the fly facility services at CCAMP/NCBS Bangalore. These consisted of either middle IDR or C-terminal IDR deleted or w.t. control Atx2 alleles, sequences containing their introns and tagged in the first intron with 3xP3 dsRed reporter, and were generated by independent CRISPR events using the same sgRNA sequences. UAS-Atx2-ADARcd alleles were generated by Baskar Bakthavachalu and Amanjot Singh at NCBS Bangalore for Singh *et al.*, 2021 and this thesis work using PhiC31 site-directed integration protocols (targeted to III) and screened by marker presence and PCR.

The rescue of fly lethality by various Atx2 constructs caused by endogenous Atx2 deletion was tested by substituting the endogenous Atx2 ORF with a reporter ORF within the same context. The CRISPR engineered Atx2 locus replacement cassette (Atx2 null allele) used as part of this thesis work was engineered by WellGenetics Ltd, Taiwan utilising the same sgRNA sequences as previously described. *Atx2* gene open reading frame from the first start codon to the stop codon including all introns was

replaced with two mirrored attP targeted integration sites with a 3xP3 dsRed reporter sequence in between them for screening. *Nanos-Cas9* fly eggs were injected into the posterior pole with gRNA and homology-repair template plasmids, raised to adulthood, crossed with Tm3 balancer flies and screened visually for red fluorescence in eyes. This homozygous lethal integration cassette fly was sequence confirmed and used for site-directed replacement with Ataxin-2 alleles being screened – SNAP-tagged w.t. cDNA, human Atxn2 cDNA, and fly Mini-Atx2. The protocol was modified for the MIMIC system (Venken *et al.*, 2011) and the site-directed integration was subsequently employed to exchange the receiver/reporter cassette for an Atx2 construct of choice. Exchange cassette and PhiC31 integrase plasmids were co-injected into posterior poles of the CRISPR modified-fly eggs. The integration can occur in either the correct or the antisense directionality therefore progeny where the dsRed reporter expression is abolished were screened via PCR and sequencing to confirm the correct directionality integration of the construct. Reverse directionality integrated flies were utilised as controls. All generated flies were balanced on III with Tm6 Tb B.

Fly brain dissections and sample preparation

Fly brains were dissected for the purposes of fluorescent immunostaining and confocal imaging, as well as for fly brain protein western blotting. Flies were anaesthetised with CO_2 and euthanised by removing the head. Heads were then suspended in PBS or PBS+Tween 0.5% and the individual brains were removed and cleaned of sclera. For western blot applications, at least 20 brains were then collected and lysed in sample buffer through mashing with a pestle and boiling of the sample. For imaging purposes, dissected brains were fixed in 4% PFA in PBS for 30mins and washed in PBS+Tween 0.5% three times with rotation before primary antibody staining, subsequent washing and secondary antibody staining. Unbound secondary antibodies were washed off three times before the samples were permeated in VectaShield imaging media overnight and mounted on slides with coverslips.

TRIBE protocol and data analysis

TRIBE data acquisition and analysis pipelines were performed by collaborators in NCBS Bangalore, India as described in Singh *et al.*, 2021. TRIBE protocol was designed by and adapted for use from the Rosbash lab, as detailed in McMahon *et al.*, 2016. In brief, CRISPR edited UAS-Atx2-ADARcd allele carrying flies, were crossed to generate a neuronal targeted, temperature induced construct expressing animals: elav-Gal4 (pan-neuronal localised), actin promoter Gal80ts (suppresses Gal4 activation in temperatures below 24C), and UAS controlled Atx2 allele paired with the RAN editing

catalytic domain of *Drosophila* ADAR. Adult flies were grown for 5-10 days at 28C post ecclosure before head dissection and RNA extraction, followed by RNA sequencing and data analysis.

Drosophila ALS degeneration model testing and analysis

Atx2 mutant allele locus replaced flies described earlier – specifically w.t. Atx2 controls, and MIDR and CIDR deleted alleles – were crossed with UAS-FUS (human w.t. sequence as well as the ALS patient point mutation R518K) containing degeneration model flies – gift from Udai Pandey lab, University of Pittsburgh. Recombinants were generated, screened by marker presence and PCR, and balanced on chromosome III. Flies were subsequently crossed with the eye-specific Gal4 driver line "GMR-Gal4" to generate +/GMR-Gal4; Atx2 "allele" recombined with FUS "allele" homozygote III flies that were to be assayed for degeneration. These flies were raised to adulthood at 25°C, aged to 3 days and frozen at -20°C before eye phenotype imaging on a Leica light microscope. At least 41 flies of each genotype were observed and the level of degeneration was scored into none, mild or severe – with severe being represented as a "rough eye" phenotype consisting of reduced eye size and disrupted boundary, loss of ommatidia structure, discoloration, loss of follicles between ommatidia and presence of necrotic patches when inspected visually. Mild degeneration was defined as presence of more than half of the ommatidia on an individual fly as healthy. Composite focal-plane images were generated using FIJI/ImageJ software.

Construct design, bioinformatics tools and sequence feature analysis

DNA sequence, plasmid and construct initial analysis and design was carried out using the cloud-based digital tool "Benchling" (www.benchling.com). Full length Atx2 cDNA construct sequence was based on the longest isoform cDNA as provided in the UniProt database and modified to facilitate cloning/codon optimised as appropriate. Functional domain locations were determined utilising the InterPro database and UniProt annotations. SNAP tag sequence was acquired from the SnapGene parts database. For minimal construct design, flexible domain-linker sequences of medium length were chosen from a review by Chen, Zaro, & Shen, 2013. Flexible cloning spacers are present to allow future cloning and manipulation. PhiC31 attB flanking sites are added to fly integration constructs according to the MIMIC construct design principles (Venken *et al.*, 2011). PhiC31 attB and attD sequence source: Michelle Callos Addgene submissions (Groth *et al.*, 2000). ClustalΩ and MAFFT sequence alignment tools were utilised to compare DNA and protein conservation across species and orthologs, accessed through the integrated Benchling interface.

Disordered and prion-like domain computational prediction was conducted according to principles set out by Alberti, Gladfelter and Mittag (2019). In brief: the FoldIndex as well as the meta-tool D2P2 (www.d2p2.com) were used to determine IDR candidates while the PLAAC tool (http://plaac.wi.mit.edu) was employed for sequence prion-like property predictions. As a means to verify disorder prediction, known structure based predictive modelling tools such as iTasser and later AlphaFold were utilised to confirm the unstructured nature of identified fly Atx2 and human Atxn2 and Atxn2L CIDRs.

Drosophila lethality assays

Atx2 locus replaced flies were tested for homozygous viability and fertility by searching for non-balanced individuals lacking the dominant marker in the case of the former, and isolating homozygotes and checking their ability to produce viable progeny in the case of the latter. The rescue ability of Atx2 cDNA replacement was further analysed by crossing out the newly generated allele with previously characterised Atx2 nulls such as $Atx2^{1\text{-}8}$ and $Atx2^{X1}$ and checking for survival of the progeny of the correct phenotype.

Contact for reagent and resource sharing

Further information and requests for resources and reagents can be directed to and will be fulfilled by the PhD supervisor, Prof Mani Ramaswami (mani.ramaswami@tcd.ie).

1. Ataxin-2 IDRs modulate vital granule assembly and neurodegeneration in flies and are functionally homologous and conserved in humans

The study of neurodegeneration has focused, arguably from its modern inception, on the role of protein aggregation as a key driver of cell death and thus pathology. Alois Alzheimer's pioneering 1907 paper describing his namesake disease makes mention of fibrils in the brain early on, with the trend in research eventually leading to the amyloid hypothesis (Glenner and Wong, 1984; Strassnig and Ganguli, 2005). It is with this in mind that the higher order assemblies of proteins and other biomolecules, especially within cells, have been focused on both in research and in therapeutic approaches. While clinical trials are yet to strongly support any therapeutic that removes, halts or disassembles amyloids or amyloid-like protein aggregates in implicated disease, a number of proof of concept studies maintain optimism in this field despite the challenges (Kametani and Hasegawa, 2018; Karran and Strooper, 2022). However, this has, in recent times, led to a very welcome broadening of the research horizons with regards to neurodegeneration and the mechanisms that might be at play to drive it.

One such new avenue, and the focus of this chapter, is the process of phase separation, or biomolecular condensate formation within living cells. While traditionally proteins were conceived as either being in a pathologic, detergent-resistant aggregate form in disease or "diffuse" and in solution/bound to other proteins or membranes, this view has been reframed when looking at puncta, granules or localised high concentrations of proteins – not to mention the shifting of focus from amyloid fibrils to soluble oligomers in the amyloid hypothesis (Banani *et al.*, 2017; Kametani and Hasegawa, 2018; Verdile, De Paola and Paronetto, 2019). In light of phase-separation being established as a mechanism for protein de-mixing from solution *in vitro*, phenomena such as stress granules or nucleolus in cells have been viewed as biomolecular condensates formed under the same principles as *in vitro* droplets or hydrogels in those fundamental experiments. More generally, this third state of proteins, between diffusion and aggregate, seems to be increasingly important in the their normal function and in the assembly and activity of the complexes they are constituents of. The condensed form of proteins, their co-factors and/or DNA/RNA appears analogous to the effects of "crowding agents" in biochemistry – bringing together the necessary partners for a reaction to be catalysed efficiently (Protter *et al.*,

2018). However, the potential of protein aggregation being also driven by or at least facilitated within such biomolecular condensates is immediately evident from the principle and dramatically demonstrated *in vitro* in the case of aggregate disease associated proteins, such as FUS (Patel *et al.*, 2015).

Work from our lab and others have focused on these mechanisms in the case of the SCA2 causative, aggregation capable protein Ataxin-2. Our work has looked into firstly identifying IDRs within fly Atx2, confirming the ability of these protein regions to phase separate *in vitro* and induce and form granules in cells, while finally examining the effects of the IDRs *in vivo* in flies on their normal behaviour and in degeneration models. This work to which I contributed as part of my PhD was published in Bakthavachalu *et al.*, 2018, however in this chapter I will expand on the published data and display follow-up experiments that were aimed at identifying the analogous IDRs within human Atxn2 and Atxn2L proteins, and verifying their ability to facilitate the same biomolecular condensate formation in cells as observed in the *Drosophila* study. Furthermore, through experiments focusing on isolating the IDRs and their function from the effects of the rest of the protein domains in cells and in transgenic flies, a tantalising thread will emerge that will be followed up in the subsequent chapter: that Ataxin-2 IDR function is actively modulated by the other domains of the protein and thus the IDRs driving role in both normal biology and degeneration models is in turn affected by this interplay of domains.

Replicating fly Atx2 IDR mediated effects in cell assays and *in vivo*, and verifying the region's requirement for degeneration

Atx2 is an essential gene in flies, with homozygous null mutant flies showing lethality in the early stages of embryo development. In work led by Joern Huelsmeier and Baskar Bakthavachalu that I contributed to as part of my PhD, flies were transformed through site directed integration (landing site on chromosome II) to have an extra allele of Atx2, with each of the domains sequentially deleted as well as a w.t. control. This allowed for a fly to be made through crossing with Atx2 null alleles (chromosome III) that would only have a mutant allele of Atx2 missing a particular domain and no w.t. Atx2 from its native locus, which would be homozygously null. The findings showed that while w.t. Atx2 controls produced viable and healthy progeny, the LSm and LSmAD domains were each, individually necessary for survival. In the case of the deletion of the Pam2 domain, there was a large reduction in viability of progeny – but not complete. Finally,

both MIDR (poly-glutamine plus prion like domains) and CIDR deletions of unstructured domains did not affect fly survival – thus the functions of these domains were thus implied to be outside of essential developmental and survival functions (Bakthavachalu *et al.*, 2018). However, it is precisely the dispensableness of this region for viability in flies while it remains conserved (not on an amino acid level but conserved as disordered regions separating structured domains and even conserving low-complexity and prion-like propensity within the sequence) that made its purpose and function the interest of subsequent work. What are the functions of IDRs in Atx2?

Atx2 CIDR is required for long-term habituation

A question naturally follows: if the IDR is not essential for survival yet functionally conserved – what is its biological function? This was first examined in *Drosophila* on an organismal, behavioural level. Previous work from our lab and others implicate Atx2 in memory circuitry, mRNA transport to synaptic terminals and thus synapse potentiation (McCann *et al.*, 2011; Sudhakaran *et al.*, 2014). We decided to test the effect of the IDRs on these aspects of animal function as such promiscuous regions would be key to facilitating the formation of potential mRNA transport granules in neurons (Hirokawa, 2006; Abouward and Schiavo, 2020). The short-term olfactory habituation paradigm involves exposing flies to an aversive or appetitive odorant for a short (15min-30min) period of time and testing in a Y-maze, T-maze or quadrant arena apparatus with the same odorant. While naïve flies are expected to show attraction or repulsion to an odour chamber (depending on the odour valence) when compared with an air-only chamber – habituated flies should show less of a response. The flies with correct brain wiring and neuronal function would “remember” the odour when being tested and so find the choice of air vs odour more ambiguous within the apparatus. Atx2 IDR deleted flies displayed normal short-term habituation, same as CS controls (not shown) (Bakthavachalu *et al.*, 2018).

The long-term habituation paradigm, while tested in a same manner and apparatus as short-term, involves exposing flies to an odorant at a lower concentration but for much longer amount of time (1-3 days). Sometimes, a day or similar period is inserted between training and testing flies in order to distinguish long-term habituation that relies of translation and middle-to-long term potentiation that is independent of translation and not resistant to anaesthesia. Interestingly, our lab showed that the CIDR is in fact required for long-term olfactory habituation in flies (figure 3). CIDR deleted flies reacted % more to an odour they should have been habituated to as compared to habituated

wild type flies. This was the first time the IDR region of a protein alone was showed to have an organism-level, specific effect. Further questions were prompted by this finding: what aspect of potentiation is affected by the IDR, what is its cellular function as well as whether the IDR acts independent of the other, structured domains.

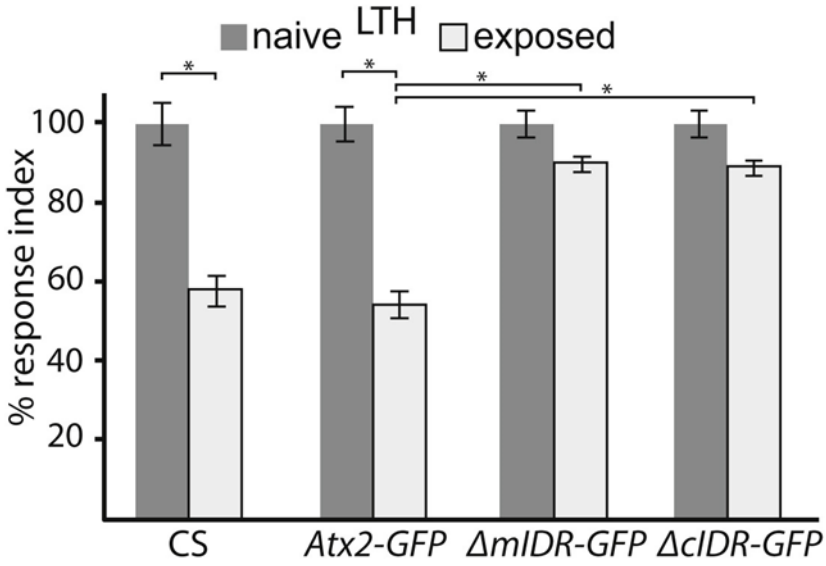

Figure 3: Atx2 intrinsically disordered regions are required for long-term olfactory habituation to ethyl butyrate in fruit flies. The Y-maze behavioural apparatus was utilised with naïve and odorant exposed (3 days) male and female flies. Naïve fly inherent avoidance of the ethyl butyrate arm of the Y-maze apparatus (as opposed to the "air" arm) was normalised as 100% of response index (y-axis). Control flies (CS) after exposure to ethyl butyrate showed significantly less avoidance (average 55%) of the odorant arm, displaying long-term habituation to the particular aversive odour. CRISPR-Cas9 edited homozygote flies were similarly tested and showed normal LTH in the case of the control, WT Atx2 + GFP tag allele replacement however a significant impairment to habituation in the case of flies edited to be missing either the MIDR (consisting of the poly-Q tract and prion-like regions) or the CIDR (consisting of the C-terminal region) parts of GFP-tagged Atx2 (average exposed fly response indexes of 90% in both cases). Work was carried out by collaborators in NCBS Bangalore Indulekha Sudhakaran and Devasena Thiagarajan. Figure is adapted from Bakthavachalu *et al.* 2018.

Atx2 CIDR activity is specific to neuronal function in LTH disruption

Imaging was carried out to pin down the location of Atx2 mutant and control protein in respective fly brains. Similarly, the correct synaptic innervation of olfactory lobe glomeruli was examined. In both of these assays, w.t. (+/+) and mutant, IDR deleted Atx2 (MIDR and CIDR) showed the same, normal result – which indicates any observed LTH issues in the mutant do not arise from the lack of Atx2 expression or from the death or mis-wiring of the neurons involved (figure 4). Notably at this point it was noticed that Me31B granules in neurons – indicative of mRNPs necessary for translational control and localisation of mRNAs (Sudhakaran *et al.*, 2014) – were notably missing in Atx2 delta CIDR homozygous fly brains. It was thus inferred that the CIDR function and the function responsible for LTH phenotypes may be related to the self and partner assembly of the Atx2 protein – i.e. granule forming ability as previously described with other mRNA binding proteins (Khong and Parker, 2020).

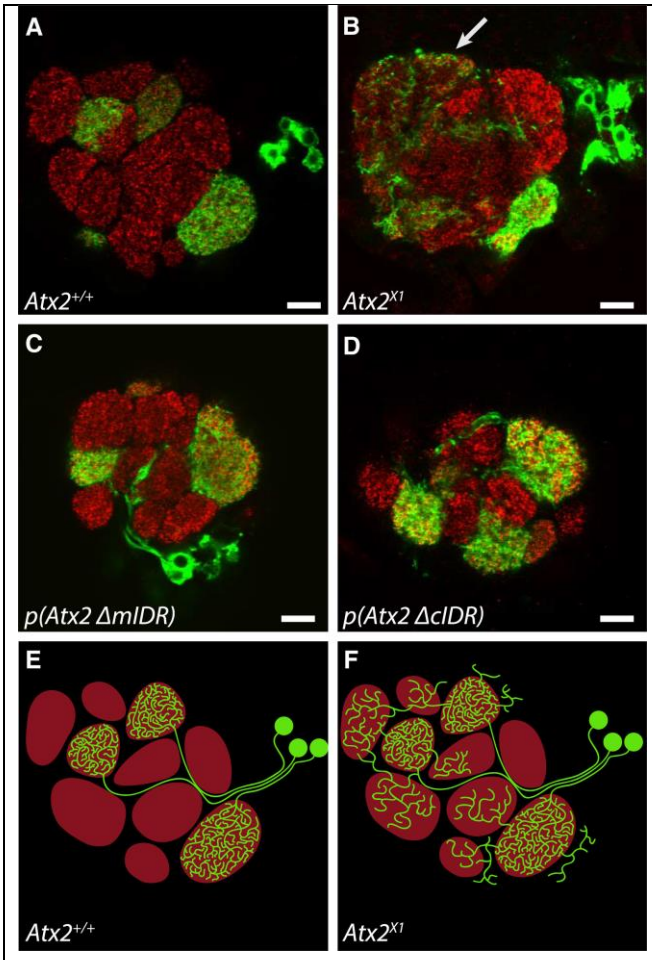

Figure 4: Atx2 IDR effects on olfactory long-term habituation are not the result of altered dendrite pathfinding in the glomeruli of the *Drosophila* antennal lobe. *Drosophila* adult antennal lobes were imaged using fluorescence microscopy showing multi-cell NP225-GAL4-expressing projection neuron clones (green) and the

synaptic marker Brp (red). (A) Atx2 is required for normal projection neuron dendrite pathfinding and targeting (n=13), whereas (B) in the Atx2 null allele atx2^{X1} containing projection neurons 100% of antennal lobes (n=10) showed defective dendrite localisation. *atx2*^{X1} mutant projection neuron dendritic pathfinding was rescued by an Atx2 transgene deleted for the mIDR (C; n = 12), and also rescued in 89% of the target neurons of flies deleted for the cIDR (D; n = 9). A schematic for normal dendritic pathfinding (E) and the defective scenario (F) is included. Scale bar = 10 µm. Work was carried out by Jens Hillebrand of Ramaswami lab and collaborators in NCBS Bangalore, led by Baskar Bakthavachalu and adapted directly from Bakthavachalu *et al.*, 2018.

The CIDR is the domain of Atx2 necessary for its self-assembly into granules via liquid-liquid phase separation

From the previous findings, a phenomenon emerged that became the focus of this thesis – the granule forming ability of the Atx2 protein. We tested candidate regions – the MIDR consisting of the poly-Q tract and prion-like regions of Atx2, and the CIDR encompassing the C-terminal of the protein – in an *in vitro* phase separation assay to determine intrinsic granule forming ability of these sequences arising from their lack of secondary and tertiary folded structures. Protein fragments, including a control peptide consisting of flanking sequences of the MIDR, were synthesised and purified before being set up in a crowding buffer (300 mM NaCl and 10% PEG3350) at a final concentration of 10 µM. Under these conditions, and imaged using differential interference contrast (DIC) microscopy, both the IDR peptides but not the control peptide exhibited liquid-liquid phase separation into droplets (figure 5). The result was our first indication of the intrinsic activity of Atx2 IDRs – an activity that may be conserved in the full length protein in the cytosol, as we proposed at the time. Without any additional binding partners or RNA, the IDRs were capable of harnessing the promiscuous and multivalent interactions between one another *in vitro*. Moving on, we can rely on this finding to inform our understanding of observations in more complex, *in vivo* in cell assays. If Atx2 IDRs do not form granules in cells under some conditions, the interpretation should consider not that the sequences are intrinsically unable to form them but that there may be factors to the system that are preventing or modulating this granule forming ability.

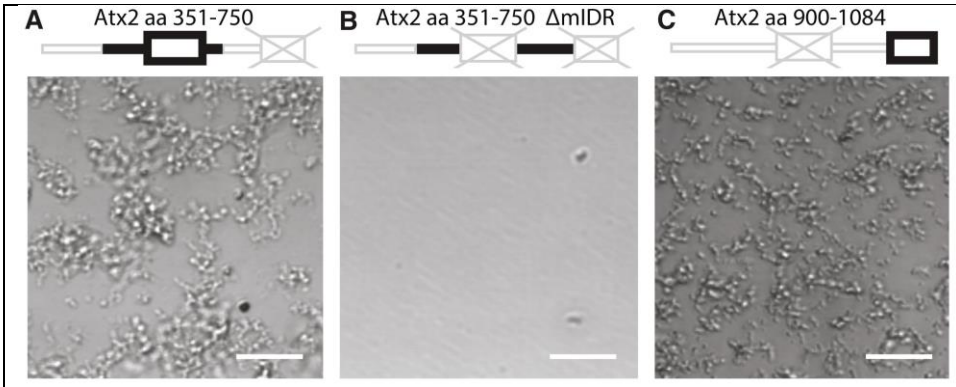

Figure 5: Atx2 IDRs, but not their flanking regions, are able to exhibit phase separation behaviour *in vitro* experiments*.* LLPS of purified MIDR and CIDR protein fragments was observed at 10 μM concentration (300 mM NaCl and 10% PEG3350) (A, C), imaged using DIC microscopy. Under the same conditions the purified protein fragments containing Atx2 amino acids 351-750 without the MIDR region (407-656) do not exhibit LLPS. Scale bar 20 μm. Work was carried out by collaborators in NCBS Bangalore, led by Baskar Bakthavachalu and adapted directly from Bakthavachalu *et al.*, 2018.

Drosophila Atx2 IDRs mediate degeneration in fly models of ALS

Complimentary to observations using the GR50 (C9ORF72 RAN-translated dipeptide overexpression in eyes) ALS model in flies, carried out by the Ramaswami lab (Bakthavachalu *et al.*, 2018), I conducted the same experiments with human FUS protein overexpression ALS model in flies. This involved generating recombinant flies between CRISPR-engineered alleles of Atx2 and the IIIrd chromosome inserted UAS-FUS degeneration model. Recombination success was confirmed through eye-colour/fluorescent reporter presence in eyes as well as single-fly PCR. Flies, reared at 25°C and expressing UAS-FUS driven by GMR-Gal4, were observed to predominantly show a “rough eye” phenotype in w.t. Atx2 background (figure 6). This is the result of cell death within the organ causing imperfect organisation and development of ommatidia. In middle-IDR deleted Atx2 background, the average level of degeneration is marginally lower in comparison. Finally, C-terminal IDR deleted Atx2 background led to a marked alleviation of the expected “rough eye” phenotypes. This evidence shows how the IDRs of Atx2 modulate phenotypes in at least two fly degeneration models, that granules or at least granule forming ability may be the underlying activity that bestows this observed modulation, and suggests a potential shared mechanism to the

pathology of these and other degeneration models, as per Ramaswami, Taylor and Parker, 2013.

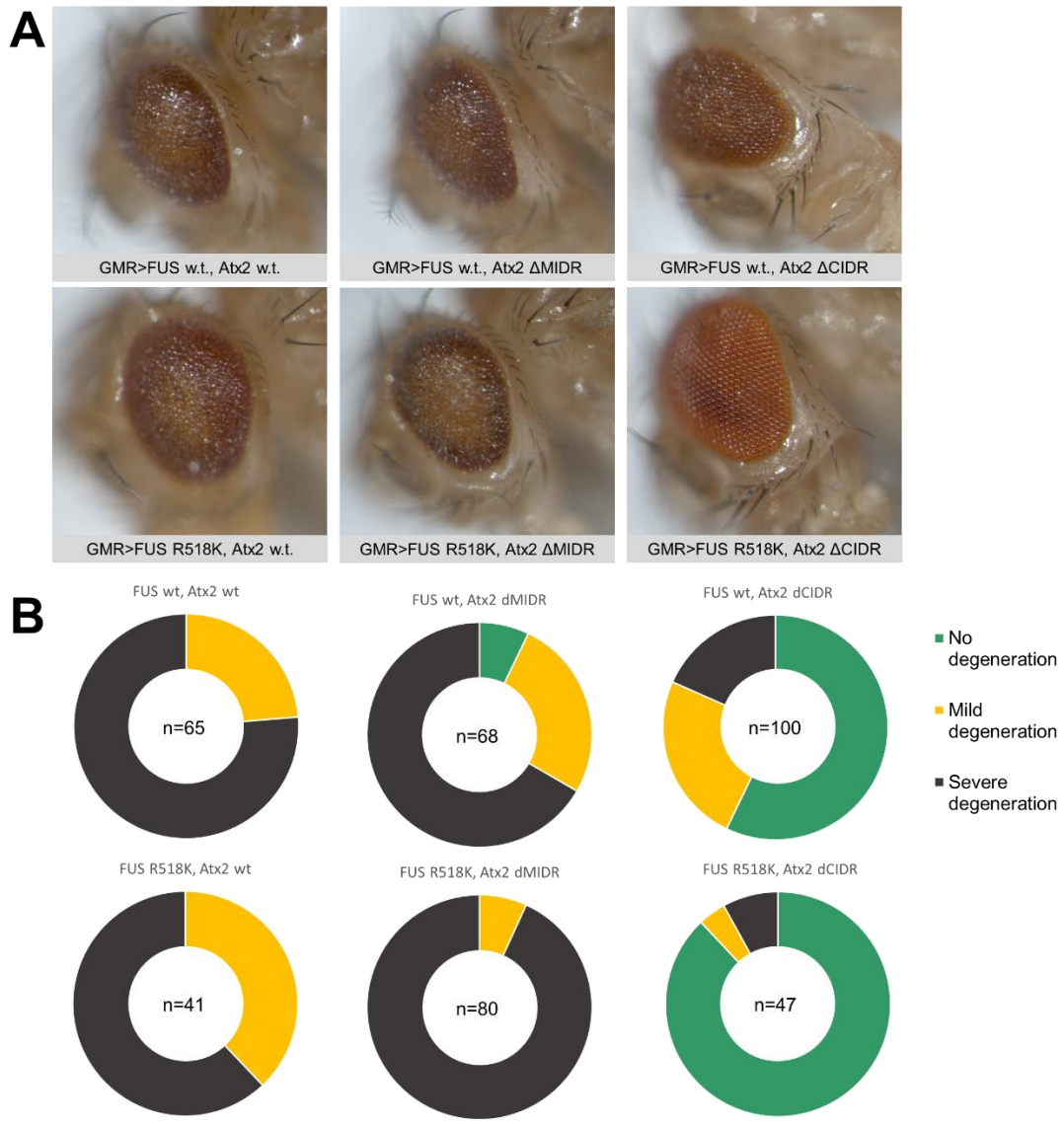

Figure 6: *Drosophila* Atx2 IDRs mediate degeneration in fly models of ALS. (A) Representative images of rough eye phenotype observed in flies expressing w.t. FUS transgene driven in eyes (top row) by the GMR promoter and R518K disease mutant FUS (bottom row) in w.t. Atx2, Atx2ΔMIDR and Atx2ΔCIDR genetic backgrounds. Atx2ΔCIDR background prevents the characteristic discoloration and fused ommatidia of the rough eye phenotype of the degeneration model. (B) Quantification of eye degeneration levels among progeny of stated genotypes. Model used: GMR-Gal4 driven overexpression of UAS-FUS (human) w.t. or ALS associated R518K mutant. Flies raised at 25°C and imaged (frozen) 10 days after ecclosion. No

progressive phenotype was observed. Data collected as part of my thesis work and published as part of Bakthavachalu et al. 2018.

Atx2 CIDR is essential for transfected protein granule forming ability in S2 cells

Following observations of both higher order effects of the IDRs of Atx2 *in vivo* in flies on both long-term habituation and degeneration model progression, and tying in what we know about the expected behaviour of IDRs – confirmed through *in vitro* phase separation assays – it became pertinent to examine if the proposed granule forming ability of these regions is observed and replicable in cells. Firstly, as widely published, we were unable to see Atx2 forming granules in either the neuropils or projections of *Drosophila* neurons in brains or in cells of the developing wing-discs via fluorescent immunostaining and confocal imaging using a variety of anti-Atx2 antibodies (data not shown) (McCann *et al.*, 2011; Sudhakaran *et al.*, 2014; Bakthavachalu *et al.*, 2018). Thus we used *Drosophila* macrophage precursor cultured cells – Schneider 2 or S2R+ cells – to further probe these properties. In untransfected S2 cells, immunostaining for Atx2 yields similar results as in dissected fly tissues: the staining appears generally diffuse and localised to the cytoplasm but not nucleoplasm (data not shown). However, we developed an assay whereby transfecting the cells to overexpress Atx2 (increasing the concentration of the protein in the cytoplasm) consistently yields to the formation of numerous, large granules of Atx2 protein forming in these cells. These granules co-localise with and thus recruit or sequester known SG markers to them, despite media conditions not inducing SGs in adjacent untransfected cells in the same biological preparations (data not shown) (figure 7, table 1).

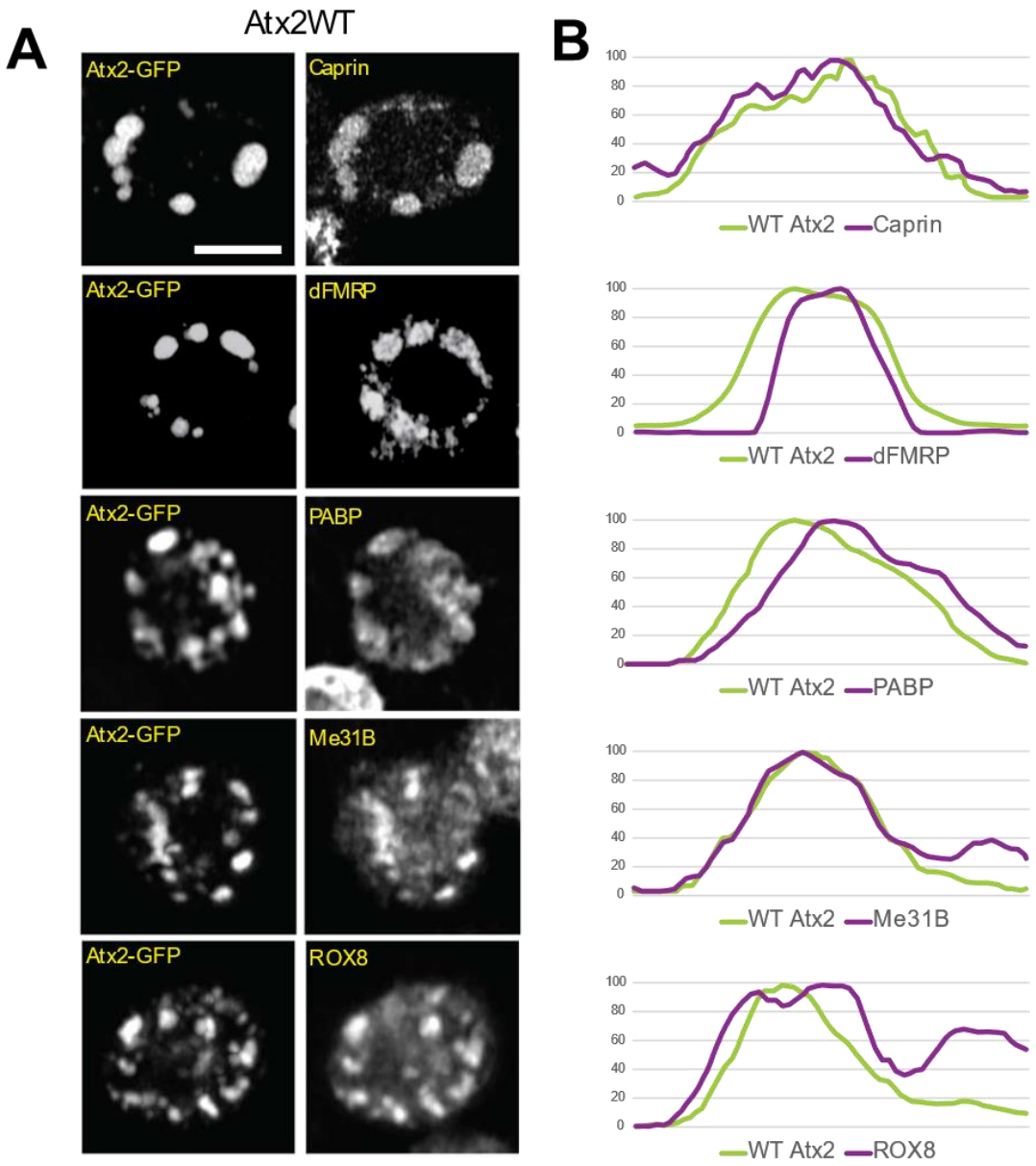

Figure 7: Exogenously overexpressed *Drosophila* Atx2 forms granules in unstressed S2R+ cells that co-localise with stress granule components. (A) Over-expression of Atx2-GFP in unstressed Drosophila S2R+ cells induces the formation of Atx2-GFP granules to which the stress granule components Caprin, dFMR, PABP, Me31B, and ROX8. Scale bar = 5 µm. (B) Normalised profile plots of Atx2-GFP granules in S2R+ cells. Within representative granules of wild type Atx2-GFP (green line), SG components Caprin, dFMRP, PABP, Me31B, and Rox8 show largely overlapping enrichment of fluorescence profile along a line bisecting a granule after immunohistochemistry and imaging (purple line). See table for quantification. Data published in preprint (Petrauskas *et al.*, 2022).

Co-localisation of signal enrichment in granules	
	Atx2-GFP granules
Caprin	100%
dFMRP	100%
PABP	76.7%
Me31B	66.7%
Rox8	100%

Table 1: Quantification of co-localisation for stress granule components to Atx2-GFP granules. From experiments shown in Figure , N = 48-120 images of Atx2-GFP granules were randomly selected for each co-staining with the stress granule components - Caprin, dFMRP, PABP, Me31B, and ROX8 - analysed for signal co-enrichment (see methods) in the case of each component assayed. PABP and Me31B co-localisation was observed as lower due to high background localisation of these proteins known to be diffusely expressed in cells. Data published in preprint (Petrauskas *et al.*, 2022).

This powerful assay was utilised to show the necessity of the Atx2 CIDR specifically for the observed granule inducing ability of Atx2 protein. As previously published and precisely and quantitatively replicated as part of my thesis work, unstressed S2 cells transfected with pAcman-Atx2-ΔCIDR-GFP expression plasmid do not exhibit granules when imaged for GFP fluorescence through confocal microscopy 24H after transfection and expression. Imaged cells 48-72H post-transfection showed diffuse GFP signal in >70% of cells while the remainder contained a smaller number of larger, more-dynamic granules (characterised in Bakthavachalu *et al.*, 2018). This is in contrast to w.t. Atx2 transfection, which forms granules in all transfected cells by 24H post-transfection. Notably, MIDR lacking Atx2 still forms granules in all transfected cells analogously to w.t. Atx2, thus signifying the importance of the CIDR specifically for the observed granule induction (figure 8). This work confirmed that the CIDR of Atx2 modulate not only *in vitro* phase separation of the protein fragments, but that this phenomenon can

be similarly tied to the protein's de-mixing/condensation in cells when its abundance in the cytoplasm is increased, as in the case of our overexpression assay.

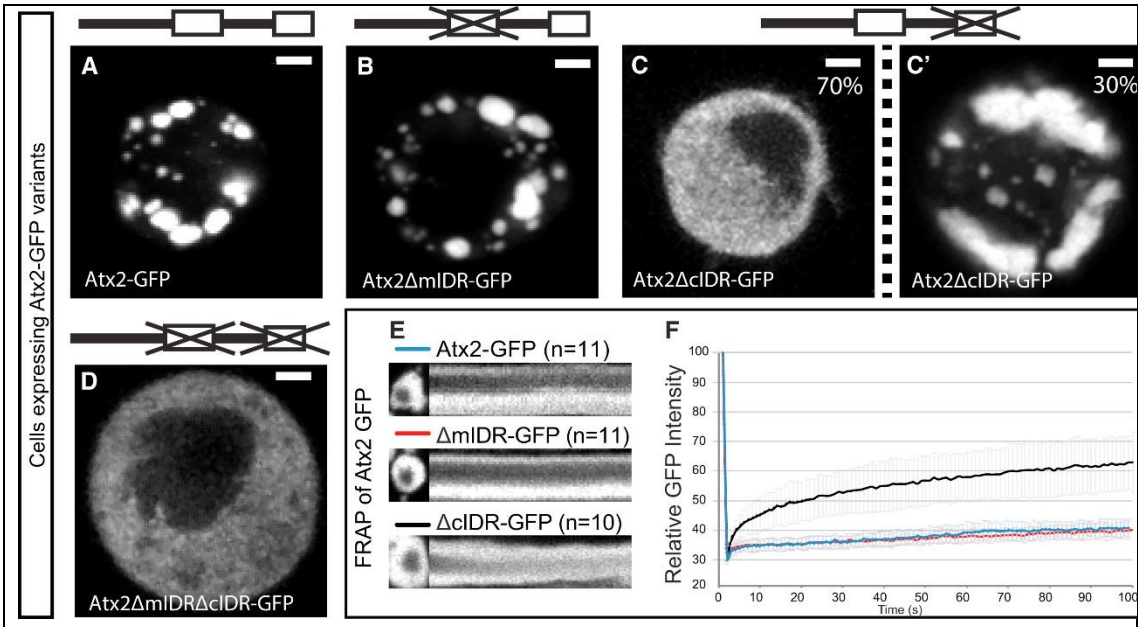

Figure 8: Atx2 CIDR is necessary for the efficient formation of Atx2-GFP granules in the S2R+ cell exogenous overexpression experiments. (A-D) Representative images of S2R+ cells transiently transfected with and expressing various alleles of Atx2-GFP, shown as cartoons above the images. (A) Atx2-GFP efficiently forms granules in transfected cells, while (B) a deletion of the MIDR does not significantly affect the phenotype. (C) The CIDR deletion visibly alters the phenotype of Atx2-GFP in transfected cells, with 70% of the population showing diffuse expression while (C') only 30% of the cells exhibited presence of granules. (D) Deletion of both IDRs in Atx2-GFP leads to the complete loss of granule forming ability in the assay. (E) Kymographs of FRAP experiments on granules in cells transfected with WT Atx2-GFP (blue) ΔMIDR Atx2-GFP (red), and ΔCIDR Atx2-GFP (black) with average fluorescence recovery graphs over 100 seconds (F). ΔCIDR Atx2-GFP displays dramatically faster recovery dynamics compared to the other alleles tested. Scale bars = 1μm. Work was carried out by collaborators in NCBS Bangalore, led by Baskar Bakthavachalu and adapted directly from Bakthavachalu *et al.*, 2018.

Atx2 polyQ expansion does not significantly affect the granule forming ability of Atx2 in S2 cell assays

To further understand the connection of granule forming ability to the IDRs of Atx2, we chose to analyse the poly-glutamine repeat region (included in the MIDR according to our notation) of the protein that is the cause of Ataxin-2 modulated disease in humans

and in mammalian model organisms. Due to this link, we expected that expanding the analogous polyQ tract in *Drosophila* Atx2 would potentially make granules form to a greater extent, more readily or make them less dynamic in our S2 cell overexpression assay. Such an effect would have suggested that disease mimicking expansions affect the ability of Atx2 granules to be the crucibles where pathogenic protein aggregates may form by increasing their viscosity and the packing of proteins within them. W.t. Atx2 23Q tract was expanded to be 60Q long, as direct CAG repeats in the DNA. However, confocal imaging experiments and replicates failed to show any observable difference between w.t. Atx2-GFP granules and Q60-Atx2-GFP granules in our assay (figure 9). The slow dynamics of w.t. Atx2-GFP granules in FRAP experiments meant that we were unable to see a difference in recovery rate in preliminary experiments between the w.t. and Q60 granules due to low dynamic range of the experiment (data not shown). We note that differences may still be detectable in other assays or upon further refinement of methodology for this precise application, however due to a lack of an obvious readout we conclude that the MIDR's effect when deleted or expanded is still secondary to the drastic effect of the CIDR on the phenomenon of Atx2 granule formation.

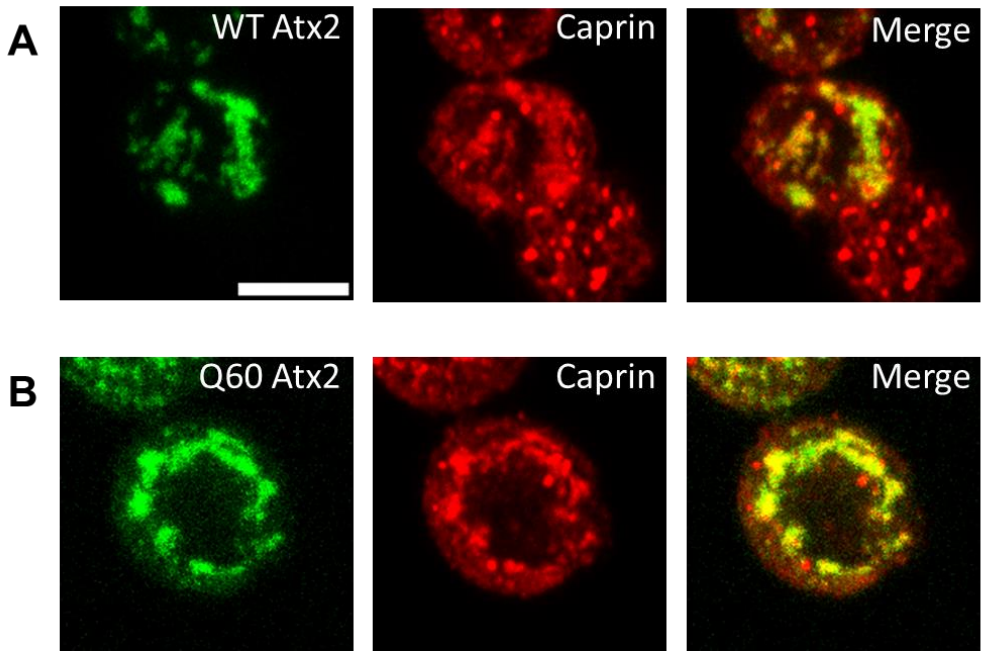

Figure 9: Atx2 poly-glutamine tract expansion does not visibly alter the granule forming phenotype of the protein when overexpressed in a tranfected S2R+ cell assay. (A) Representative fluorescence microscopy images of S2R+ cells transfected with WT Atx2-SNAP plasmids show the expressed protein forming into granules which co-localise with the stress granule marker protein Caprin. Granules

are dynamic and display signs of fusion into larger agglomerates within the cytoplasm. (B) Expansion of the poly-Q tract within the same Atx2-SNAP constructs from 23 to 60 (Q60) does not visibly affect the granular phenotypes of transfected cells or co-localisation with Caprin in this assay. Scale bar = 5μm.

IDR-only construct expression fails to induce granule formation in S2 cell assays

To get a better idea of the forces at play within unstressed, living cells that control whether or not the Atx2 protein condenses into granules as part of its function, the reciprocal experiment to CIDR deletion was carried out. CIDR-only constructs were generated, in the same expression backbone, and transfected into S2 cells. The results showed that the CIDR alone was insufficient to induce granules, irrespective of how long the cells were transfected for (up to 72H were tested), also serving as a control for potential overexpression induced “stress” that might have prompted the forming of SGs in cells that could be mistaken as the cause of w.t. Atx2 granules in our assay (figure 10). The essential region of the protein for granule formation is not sufficient to induce them in living cells, we know the CIDR alone is capable of phase separation *in vitro* however in cells there appears to be additional elements of regulation for condensate formation. While there could be a number of explanations as to why this occurs, for us this naturally reinforced the idea of the phenomena of granule formation or biomolecular condensation within cells as being highly regulated due to its necessity and utility for countless processes and reactions within cells. Together, this implies that elements outside of the CIDR may be necessary for its granule forming ability in the cytoplasm specifically, and more generally for the control of the Atx2 proteins functions in cells.

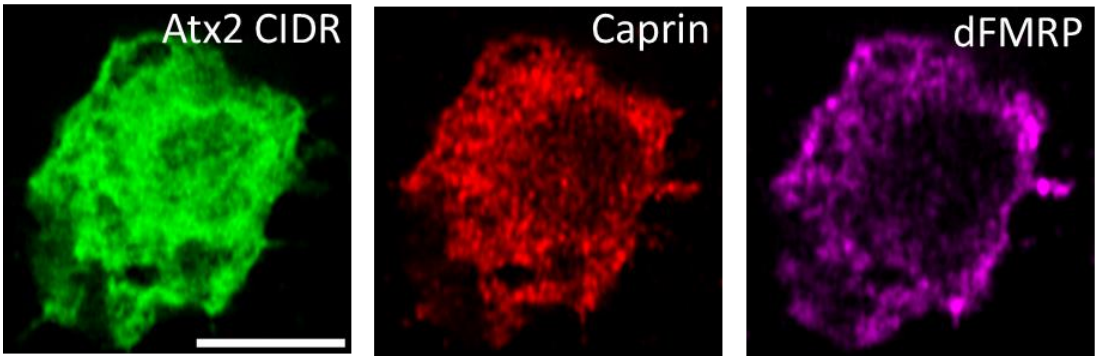

Figure 10: Overexpression of Atx2 CIDR alone is incapable of inducing granule formation in the *Drosophila* S2R+ cell assay. Act-Gal4 driven UAS-Atx2 CIDR expression plasmids were transfected into S2R+ cells for 72H. Confocal fluorescence imaging of the expressed protein fragment could not detect the presence of visible granules in the transfected cell population while the diffuse staining of the stress granule markers Caprin and dFMRP show the absence of hence or otherwise induced stress granules in the same, long-term transfected population of cells. Scale bar = 5μm.

Identifying Human Atxn2 IDRs and showing their functional conservation to fly Atx2 IDRs

Dissecting the discrete functions of the Atx2 protein begins first by examining the amino acid sequence for higher-order trends – that is to say – the fundamental characteristic of forming stable secondary and tertiary structures. As mentioned previously, a particular sequence maybe be identified as either conventionally ordered or intrinsically disordered. Ordered sequences of amino acids are subject to a sum of thermodynamic and steric forces that drive alpha-helix, beta-pleated sheet or other structure formations in order to reduce the Gibbs free energy of the polypeptide into the nearest achievable local energy minimum. This includes hiding and clumping together hydrophobic R groups while exposing polar amino acids to the solvent, an action that automatically drives the formation of globules or pockets of clustered amino acids. In total, this leads to a specific, relatively stable 3D structure that a particular sequence adopts, and from which properties of affinity, active motifs, binding pockets and catalytic action arise. Structure determines function. On the other hand, intrinsically disordered sequences (corresponding to IDRs within the domain structure of a protein) can either have no secondary and tertiary structure or fluctuate between several such structures due to their local energy minima being very close to each other with no major energy hurdles between them. As such, the “function” of disordered regions determines their structure, or conformation at any particular moment in time. These are two completely different paradigms of protein structure, however, in this thesis I will try to show how functions of these disparate domains interact, antagonise and synergise with one another to control the output or role of the protein as a whole. My study focuses on the particular example of the protein Ataxin-2 (Atx2 in *Drosophila* and Atxn2 in mammals, as well as the chordate-only homolog Atxn2-Like/Atxn2L) in order to

describe phenomena which may be transferrable to a more general scale with further work.

Finding candidate IDRs using sequence analysis and computational methods

The initial and fastest method available to our lab in determining regions of potential disorder within the Ataxin-2 protein was digital sequence analysis. Utilizing the databases available through Uniprot (www.uniprot.org), *Drosophila* Atx2 showed the presence of three structured domains, as described earlier, with predicted secondary structure for each domain or in the case of LSm domain – crystal structure from a homolog. The broadest possible analysis would suggest any sequence outside of the denoted "domains" of the proteins would potentially be intrinsically disordered. To narrow down the regions more specifically we used protein sequence alignment between *Drosophila* Atx2 and human and mouse Atxn2 and Atxn2L to show a lack of conservation in regions outside the LSm, LSm-AD and PAM2 domains. This suggested the sequence was potentially functionally conserved and not sequence specifically conserved – a hallmark of regions that do not rely on precise secondary structures and tertiary folding in order to fulfil their roles. Further, making use of prion-like sequence predictive tools (PAPA), disorder predicting tools (D2P2 meta-database, MeDor meta-database) as well as structure *de-novo* modelling tools (iTasser) we were able to confirm the disordered nature of the rest of the protein (figure 11). However, we also identified smaller regions of interest such as the poly-Q and low complexity tracts that had prion-like propensity and in *Drosophila* formed a middle IDR (MIDR) spacer between the LSm-AD domain and the PAM2 domain, as well as identifying a C-terminal tail of Ataxin-2 past the PAM2 domain in both fly and mammalian proteins that was heavily disordered (the CIDR). Confirming these older finding using contemporary tools – notably the AlphaFold database (www.alphafold.ebi.ac.uk) produced the same result of the three confidently structured domains surrounded by the remainder of the amino acid chain as low-confidence looped or disordered regions (figure 11).

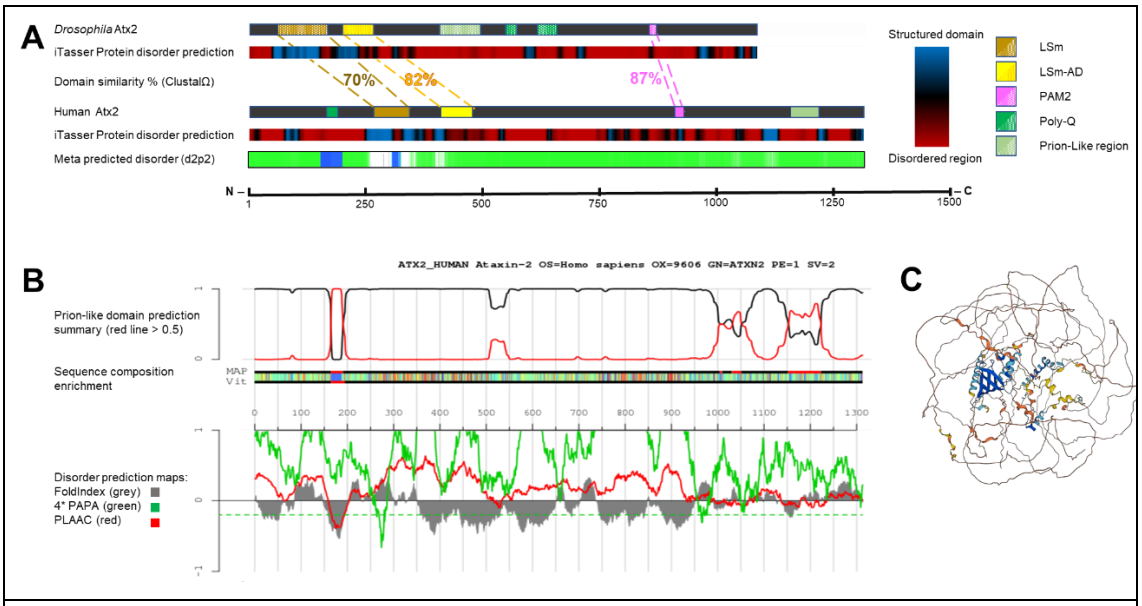

Figure 11: Computational analysis of *Drosophila* and human Ataxin-2 identifies conserved structured domains and predicts candidate IDRs in human Atxn2. (A) Summary of domain structure, homology and predicted disordered protein secondary and tertiary structure in fly and human Ataxin-2. The LSm, LSm-AD and PAM2 domains were identified by three separate approaches – iTasser known structure fitting score (red-blue representing confidence), ClustalΩ homology alignment, and d2p2 disorder prediction meta-tool – as the key conserved structured regions of the protein homologs. All other regions did not show sequence conservation between fly and human as well as low confidence of known structure fitting and high predicted sequence disorder by the d2p2 meta-tool of prediction models. (B) PLAAC web-tool readout of prion-like sequence prediction for human Atxn2. The poly-Q domain as well as a putative CIDR of Atxn2 are flagged as potentially having prion-like propensity. The aggregate score is displayed as a red line on a scale of 0 to 1, with >0.5 above background (black line) being a threshold for likely prion-like behaviour of the sequences based on known prion-like sequence features. Presence of direct or similar amino acid repeats are shown by the sequence composition enrichment bar that highlights the poly-Q domain. Further breakdown of the PLAAC tool output is shown on a 1 to -1 scale as 4*PAPA score (green, lower value predicts prion-like propensity), PLAAC score (red, lower value predicts prion-like propensity), and FoldIndex score (grey, higher value predicts likelihood secondary and tertiary sequence structure). (C) AlphaFold database rendering of human Atxn2 structure, showing a number of small folded domains surrounded by largely unstructured regions of the protein, including the C-terminal. Taken together, these prediction tools based on known structures, amino acid conservation, and features of biologically verified prion-like or disordered sequences support our hypothesis of

Ataxin-2 containing IDRs that flank three structured domains in both *Drosophila* and human. Data shown in (B) was published in collaboration with Joern Huelsmeier in Bakthavachalu *et al.*, 2018.

Exogenous human Atxn2 is incapable of inducing granules in fly S2 cells

The question of Ataxin-2 functional conservation – important for tying the observations made in fruit flies and fly cells to the human cellular and disease context – was first probed in our key assay via exogenous expression in S2 cells. Human Atxn2 protein with the wild-type poly-glutamine tract as well as a disease-associated ~80 glutamine (Q80) expansion were initially used. Q80 constructs were unstable during expression plasmid DNA replication and could lose or pick up additional CAG repeats encoding glutamine amino acids, in a similar way heterogeneity of the expansion within the *Atxn2* locus is observed in sequencing of patients and their parents and offspring. Overexpression of WT UAS-Atxn2 constructs failed to consistently recreate the granule phenotype in S2 cells despite comparable expression levels and transfection efficiencies (figure 12). Instead, in the majority of cases the protein remained diffuse and in approximately 30% of transfected cells the exogenous protein would appear localised to one part of the cytoplasm, akin to a loosely separated mesh. Increasing the number of glutamines – overexpressing the Q80 UAS-Atxn2 construct – did not lead to significant differences in the observed phenotypes (figure 12). The failure of human Atxn2 protein to induce the well-defined Atx2 granules in S2 cells suggested a lack of functional conservation of the protein within our assay, however it was consistent with previously published work (Kaehler *et al.*, 2012) detailing the same observation using overexpression in mammalian cell lines. Interestingly, (Kaehler *et al.*, 2012) showed Atxn2-Like – the paralog of Atxn2 in vertebrates – did indeed form granules when overexpressed in U2OS cells. The absence of a poly-glutamine tract in mammalian Atxn2-Like and the negligible effect of Atxn2 Q80 on granule formation in our assay suggested that the main determinant factor driving this phenomenon may not be the glutamines alone but the IDRs in general of these proteins. Taken together with our previous observations, these data pointed us to further experiments examining the nature of the IDR of Atx2 and its difference from both mammalian forms.

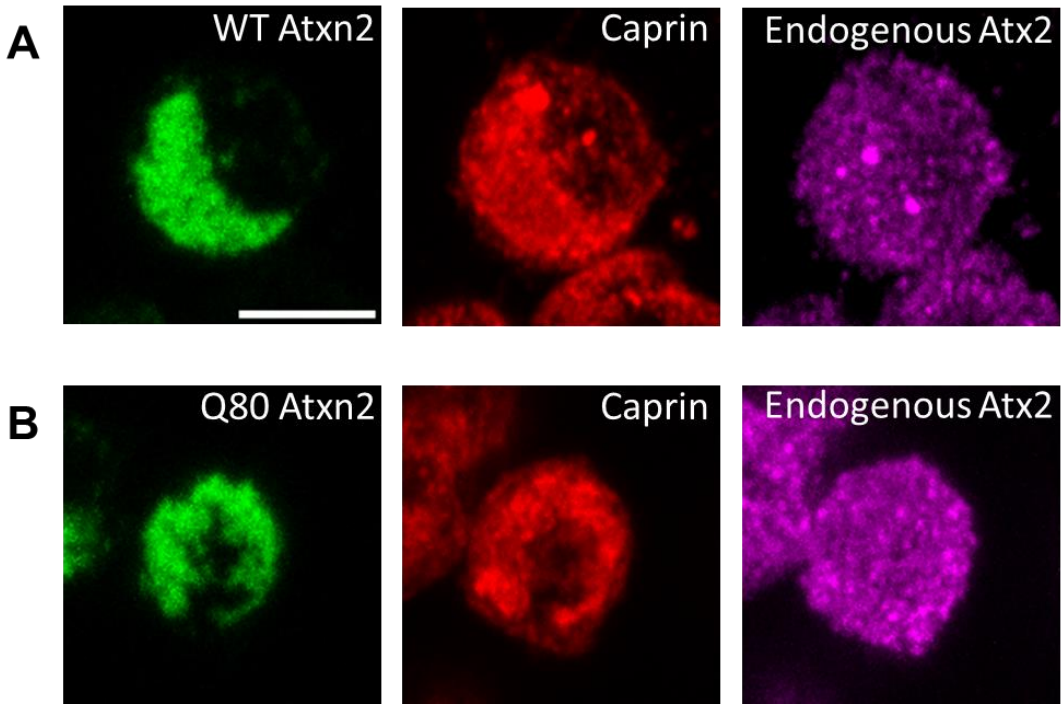

Figure 12: Human Atxn2, or polyQ expanded form, are incapable of inducing distinct granules in the transfected S2R+ cell assay. (A) Whereas fly Atx2 readily forms granules in this assay, fly codon-optimised human Atxn2-SNAP under the same conditions and control of an identical Act-Gal4 > UAS promoter, does not copy this phenotype when overexpressed in S2 cells. The expressed human protein remains cytoplasmic and generally diffuse, or loosely localised to a region of the cytoplasm. (B) The SCA2 associated polyQ expanded allele of Atxn2 (Atxn2Q80) assayed under the same conditions was also incapable of inducing distinct SG-like granules in S2R+ cells. SG marker Caprin and endogenous Atx2 in S2R+ cells were also stained for, but did not show stressed cell signals above or beyond expected staining noise or heterogeneity. Some level of co-localisation between human Atxn2 alleles and Caprin is observed and probed further in subsequent experiments. Scale bar = 5μm.

Identifying the Ataxin-2 paralog IDRs specifically as the controlling factor in their differing granule forming ability in stressed cell assays

Conservation between fly and human Ataxin-2 forms is high in the structured domains (ClustalOmega amino acid similarity between LSm domains = 70%, LSm-AD domains = 82%, and the PAM2 domains = 87%), and we suspected that this would be sufficient to reconstitute correct binding interactions of exogenous overexpressed protein within

the cells in which they were assayed. Due to the inability of the human protein to induce SGs in S2R+ cells, the transfected cells in this assay were subjected to oxidative stress using Sodium Arsenite to form SGs and the recruitment of overexpressed exogenous protein to these granules was tested. Immunohistochemical stainings for SG markers were tested for co-localisation to Ataxin-2 forms in granules. The data (figure 13) conclusively showed that human Atxn2 was able to be recruited to dFMRP positive granules, in the same way fly Atx2 could; and similarly fly Atx2 could be recruited to G3BP1 granules in the same way human Atxn2 could. These findings imply the similarity of the recruitment mechanism of Ataxin-2 to otherwise induced SGs in cells. Tantalisingly, taking into account the previously mentioned differences between the CIDRs of Ataxin-2 orthologs, this data suggests that some conserved elements of Ataxin-2 outside of the IDRs are able to regulate the localisation of the protein to granules in cells. The promiscuous de-mixing ability of IDRs appears to not be enough to always induce granules within cells and be strictly controlled by other factors such as structured domain interactions which we will address subsequently.

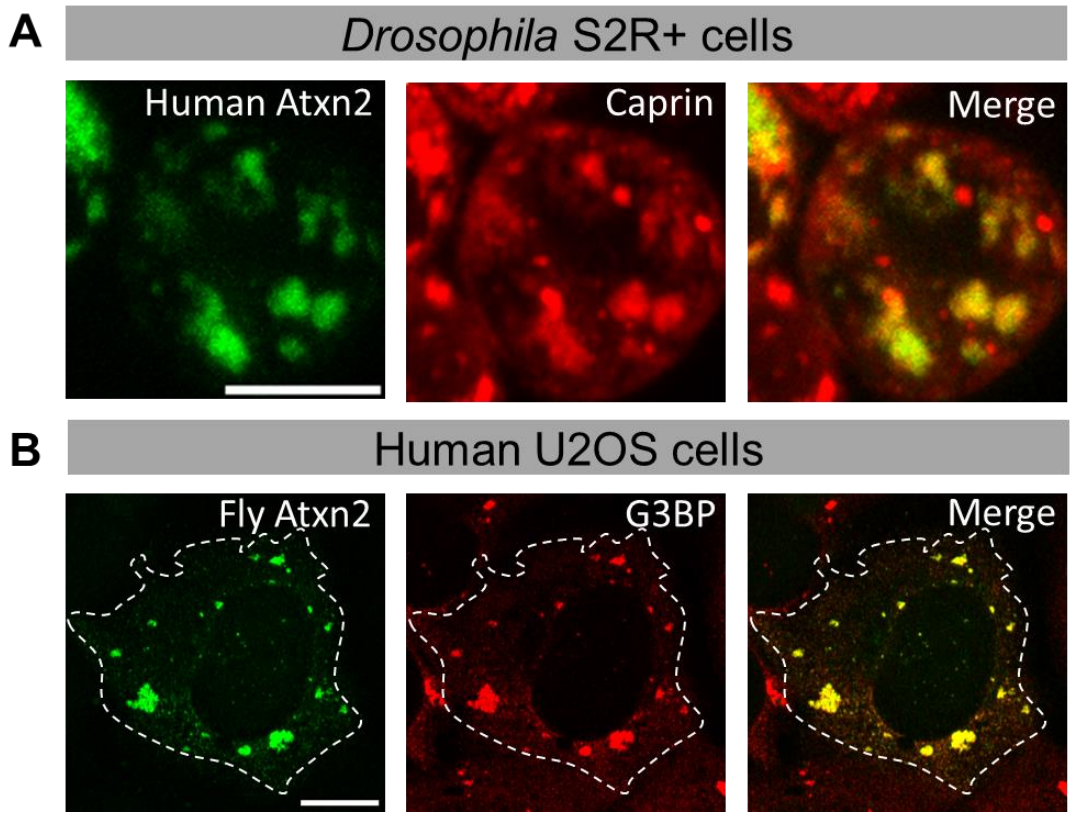

Figure 13: Exogenously expressed Ataxin-2 orthologs are recruited to stress granules and co-localise with stress granule markers in cross-species experiments. (A) Human Atxn2 is recruited to SGs and co-localises with the SG marker G3BP in *Drosophila* S2R+ Sodium Arsenite stressed (3H) transfected cell assay. Act-Gal4

UAS expression system was used. Scale bar for S2 cells = 5μm. (B) Fly Atx2 exhibits the reciprocal phenotype in stressed (1H) human U2OS cells. Co-localisation and thus implied recruitment of the exogenously expressed protein with the SG marker G3BP1/2 is observed. Dashed line indicates the area of a transfected cell. CMV-Gal4 UAS expression system was used in human cells. Scale bar for U2OS cells = 10μm.

A further experiment was carried out utilising domain-swapped Ataxin-2 constructs to independently verify that structured domains were not affecting the granule formation phenotype in the assay. In detail: fly Atx2 cDNA was split into structured domains and disordered regions, as previously described, and the structured domains of LSm, LSm-AD, and PAM2 were swapped out with their homologs from human Atxn2. The reciprocal constructs were also made where human Atxn2 and separately Atxn2-Like cDNAs had their respective structured domains swapped out for *Drosophila* homologous sequences. The entirety of the protein outside of these domains was seen as the IDR for the purposes of this experiment. These constructs were then codon-optimised for expression in *Drosophila* cells or human cells depending on their application. The results after expression in S2 cells (figure 14) were consistent with the previous whole-protein exogenous expression experiments. Notably, the fly IDRs and the Atxn2-Like IDRs were shown to be “stronger” at inducing granule formation compared to human Atxn2 IDRs. The contents of the granules/high protein expression regions remained the same in this experiment as in the previous one, with the exception of fly structured–Atxn2-Like IDR exchanged constructs in S2R+ cells and an observed low co-localisation with human structured–fly Atx2 IDR exchanged constructs in U2OS cells. The former, we speculate is due to some unknown functional domain of Atx2 being not present in the sub-functionalised IDR of Atxn2-Like, while the latter exhibits tell-tale signs of poor antibody penetration in stainings, which in turn may shed light on the tight internal packing of these exogenously induced granules as a result of the “strong” IDRs of fly Atx2. Further analysis of these phenomena may clarify these co-localisation phenomena, however, it was outside of the scope of this thesis.

Altogether, the data point to the activity of the IDRs in the driving of granule formation or at the very least having the potential to promote self-assembly of IDRs and their binding partners into a demixed, concentrated assembly of the protein. We suggest that this is one of the distinct, intrinsic activities of the IDRs of Ataxin-2 and specifically in conjunction with our previous work – the C-terminal IDR. The observed difference in

granule forming ability of the different IDRs in our assays shows us that there is biological control of Ataxin-2 granule formation in cells, as well as evolutionary conservation and sub-functionalisation of this ability among orthologs and paralogs of Ataxin-2.

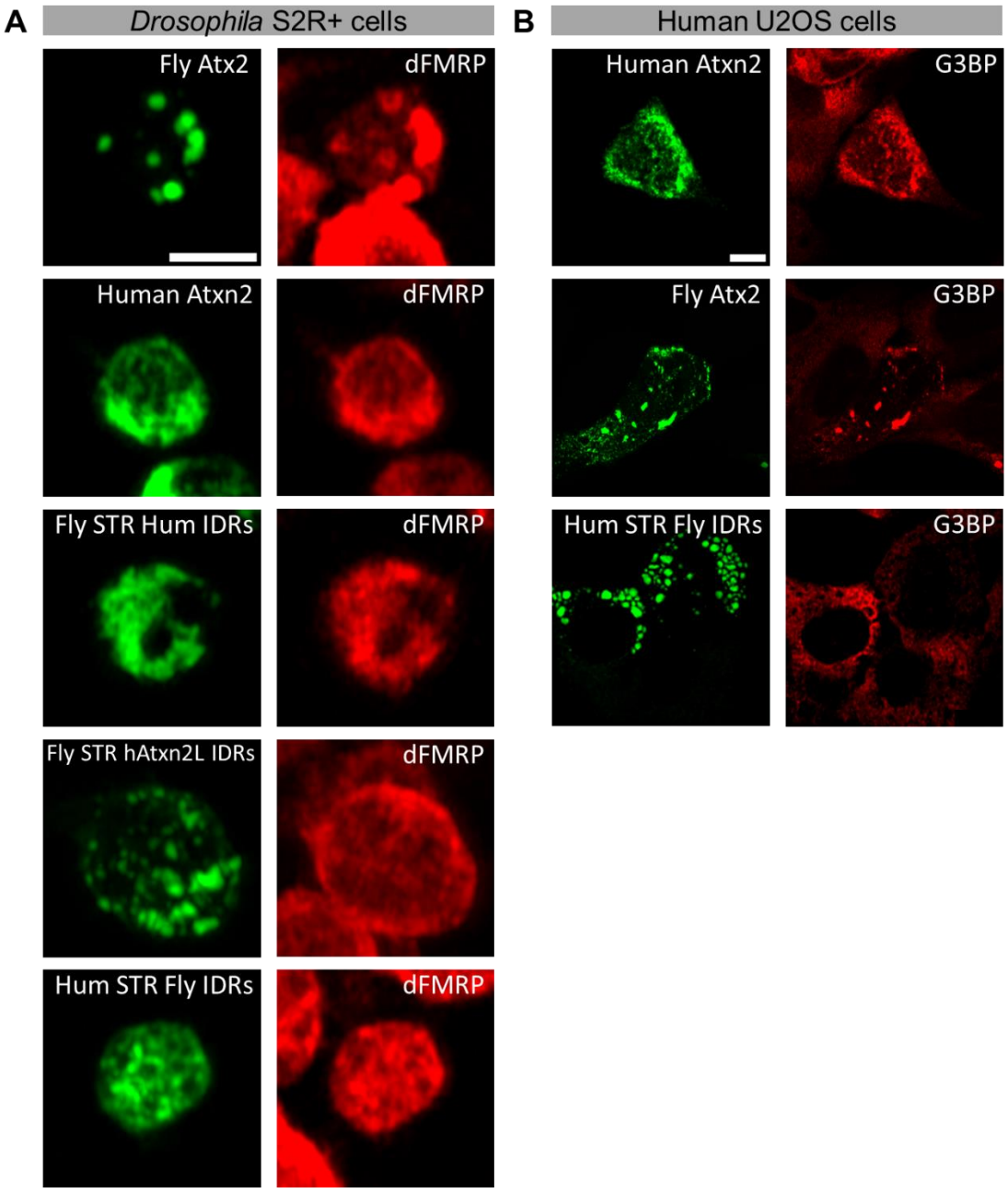

Figure 14: The IDRs of Ataxin-2 paralogs are the main sequences affecting overexpressed exogenous protein granule forming ability in cultured cells. Domain swapped constructs and controls were transfected for 24-48H into unstressed *Drosophila* S2R+ cells and human U2OS cells. Co-staining for the SG marker proteins dFMRP (S2R+ cells) and G3BP1/2 (U2OS cells) were carried out to test

inclusion into granules and untransfected cell health through absence of SGs. (A) Fly Atx2 is able to form granules in the S2R+ cell assay, while human Atxn2 is not under the same circumstances. Similarly, fly Atx2 structured domain sequence – LSm, LSm-AD, and PAM2 – swapped into the corresponding locations of codon-optimised human Atxn2, was unable to restore the construct's granule forming ability ("Fly STR Hum IDRs", where IDRs in this case correspond to the entirety of the unstructured sequences in human Atxn2). The IDRs of the human ortholog Atxn2-Like are observed as sufficient to rescue the granule phenotype in fly cells ("Fly STR hAtxn2L IDRs"), but do not include dFMRP in these granules. The reciprocal exchange of fly Atx2 structured domains for codon-optimised human Atxn2 structured domains ("Hum STR Fly IDRs") allowed for the overexpressed construct to form granules in the assay that co-localise with dFMRP. A diffuse background expression/staining is also seen with this construct. Act-Gal4 driven UAS expression system was used with all S2R+ cell experiments, expression was driven for 24H before staining. Scale bar = 5μm. (B) In human U2OS cells, exogenously overexpressed human Atxn2 does not form granules. However, human codon-optimised fly Atx2 was able to form granules that co-localise with the SG marker G3BP1/2. Similarly, exchanging the structured domains of the same fly Atx2 construct with the human Atxn2 structured domains ("Hum STR Fly IDRs") maintained the construct's ability to form granules in this assay, however full co-localisation with G3BP was not observed, due to antibody penetration issues. CMV-Gal4 driven UAS expression system was used with all U2OS cell experiments, expression was driven for 48H before staining. Scale bar = 10μm. Construct design and cloning was carried out with, and the "Hum STR Fly IDRs" experiment transfection and imaging was carried out by Daniel Fortunati in the Mani Ramaswami lab.

Decoupling the structured domains and interactions of Ataxin-2 from the function Characterising a structured-domain only allele of Atx2 in transgenic flies

Seeing the high functional and sequence homology between fly and human Ataxin-2 structured domains, and knowing that the IDRs are not essential for fly survival, we designed a CRISPR mediated approach to generate *Drosophila* lines with homozygously replaced *Atx2* locus with cDNA controls as well as minimal, structured domain only containing alleles of Ataxin-2. The question to answer was whether the

structured domains alone, which each in turn are lethal (LSm, LSm-AD) or greatly reduce the viability of flies (PAM2) if homozygously deleted in in the *Atx2* locus (Bakthavachalu *et al.*, 2018), carry out a sufficient enough amount of the Atx2 protein's essential functions to allow for the organism's viability. Furthermore, this experiment aimed at establishing if the introns of the *Atx2* gene are necessary for viability (control locus replacement with the cDNA only), if there are sequences outside of the structured domains and MIDR and CIDR of fly Atx2 that are currently unknown as essential for animal viability (replacement with minimal Atx2 construct), as well as to test if the domains of human Atxn2 are sufficiently conserved to carry out the essential functions of the protein in flies (replacement with a human Atxn2 cDNA).

To answer these questions, *Drosophila* control lines were edited with CRISPR/Cas9 and homology-directed repair to replace the Atx2 ORF with a dual, inverted attP integration cassette, modified from the approach used to generate the MiMIC library (Venken *et al.*, 2011). This cassette, marked by the eye-specific 3xP3 driven dsRed signal, allowed for heterozygous CRISPR edited flies to be easily screened and then positive lines to be more specifically and efficiently edited trough site-directed integration. Exchange of the cassette through the co-injection of PhiC31 integrase and a donor plasmid with dual attB sites surrounding an Ataxin-2 allele yielded about 50% chance of cassette replacement by the target construct in the correct orientation. Loss of the dsRed signal in the eyes of the progeny, coupled with PCR screening for orientation as well as sequencing for selected lines efficiently yielded a number of viable and confirmed heterozygous fly lines for the Ataxin-2 constructs that were tested. Self-crossing balanced (on III) lines could yield homozygote animals for the construct being tested if progeny without balancer chromosome markers is selected. These lines and approach remain valuable in the lab due to the capacity to test mutants and constructs of Ataxin-2 *in vivo,* potentially in a background without any endogenous Atx2, and under the control of the native *Atx2* promoter and all cis- and trans-regulatory sequences present in that particular locus.

The minimal Atx2 construct (Mini Atx2) was generated as a means to isolate only the structured domains of Atx2 and to decouple the effect of the IDRs from these domains. Therefore potentially only the essential, structured interactions of the Atx2 protein would be preserved in this mutant and any animal generated with this replacement. Minimal protein constructs have been shown by other groups to be sufficient in reconstituting some of the essential or specific functions of a protein and we were inspired by such results to attempt a similar dissection of functions into ones mediated

by folded domains and IDRs for Atx2 (Tillotson *et al.*, 2017). In detail, the LSm, LSm-AD and PAM2 domains of fly Atx2 as defined by the UniProt database, plus ~4 additional flanking amino acids on each side were combined into a construct using GGGS(3) flexible linkers between the structured domains, a start codon was added, and a SNAP tag with a stop codon separated by a flexible linker was added to the C-terminus. Flexible linkers were used to separate the domains in order to reduce potential steric hinderance or folding interference of the naturally separated structured regions, as is common practice in synthetic biology and protein tagging (Chen, Zaro and Shen, 2013). The Mini Atx2 construct provides a powerful tool, more so than single IDR deletions, and examined further in subsequent experiments, for a clear separation of the proteins specific interactions and the more promiscuous interactions of the IDRs.

Fly cDNA-only replacement of the Atx2 locus is homozygous viable and rescues the lethality of Atx2 null alleles in flies

While homozygous deletion of *Atx2* is lethal in flies, flies generated with the heterozygous and homozygous replacement of the *Atx2* ORF with only the cDNA (longest isoform) of the gene are viable and fertile in both cases. Using the previously described CRISPR/Cas9 and PhiC31 site directed integration gene editing techniques to generate a number of positive and control replacement lines, the cDNA of Atx2 was shown as sufficient for animal survival and fertility and thus allowing us to suggest that for these functions alone the introns of the *Atx2* gene are not required (figure 15, table 2). To further genetically verify these findings, cDNA replaced flies and heterozygote null controls were crossed to other existent Atx2 null alleles in a complementation assay. These alleles, $atx2^{1\text{-}8}$ and $atx2^{X1}$ were independently generated by this lab via either CRISPR/Cas9 (excision of exons 1-8 and non-homologous end joining) or P-element integration and excision (landing in exon 1 and imprecisely excised and repaired to yield an unexpressed gene) respectively (Sudhakaran *et al.*, 2014; Bakthavachalu *et al.*, 2018). Independently generated lines are extremely unlikely to carry secondary mutations or off-target edits that may reduce animal survival in homozygotes with one another as the chances of the same, random, complementary lethal mutations is negligible. Heterozygotes with w.t. *atx2* on balancer chromosomes were, as expected, all viable with the exception of null or backwards integrated controls. However, only the cDNA locus replaced flies were both homozygous viable and able to complement *atx2* null alleles in heterozygotes (figure 15, table 2). Taken together, we can conclude that there are no lethal, off-target mutations on the

generated locus replacement alleles, that Atx2 cDNA is sufficient for the essential functions of the protein thus the introns are dispensable for animal survival, and finally that there may be an essential function performed by hitherto unknown regions of *atx2* outside of the previously implicated LSm, LSM-AD and PAM2 and that this region is insufficiently conserved between fly and human Ataxin-2. Conversely, the very geometry and spacing out of the folded domains of Ataxin-2 within the tertiary structure of the protein may be an essential characteristic of the protein and is a fertile avenue for further experiments.

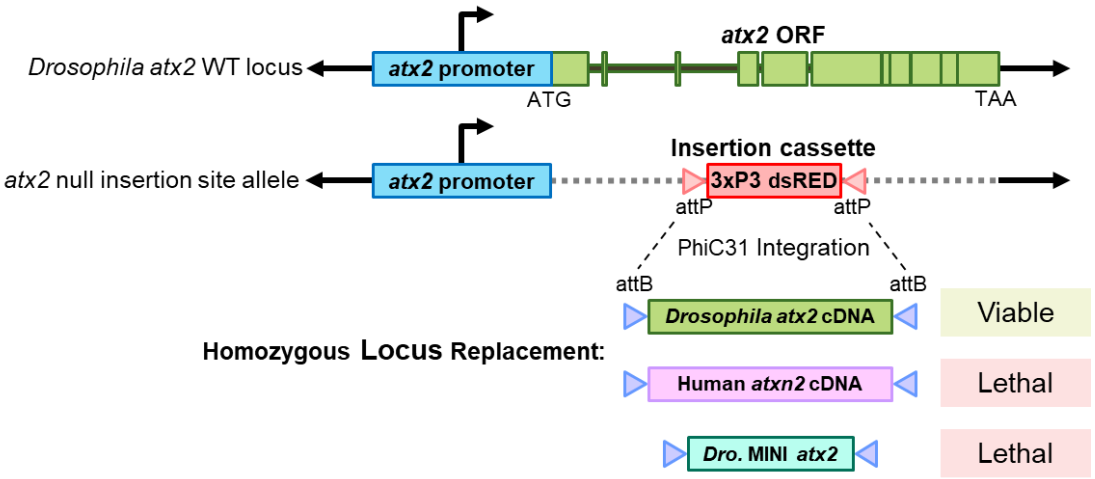

Figure 15: Summary of *Drosophila in vivo atx2* locus replacement experiments and generation strategy. The *atx2* locus was edited with CRISPR/Cas9 to remove the ORF of the gene – including all sequences from the start codon to the final stop codon and all exons and introns. The inserted site-directed integration screenable cassette is then efficiently swapped out using the PhiC31 mediated protocol. Only the fly cDNA replacement of the locus allows for viable (and fertile) homozygote offspring, while both human and fly Mini Atx2 replacements are homozygous lethal. The backwards integrated replacements (not shown) served as controls and were homozygous lethal in all cases.

	Null dsRED Cassette	**Fly cDNA replacement**	**Fly Mini replacement**	**Human cDNA replacement**
***atx2* 1-8 null**	Lethal	Fully Viable	Lethal	Lethal
***atx2* X1 null**	Lethal	Fully Viable	Lethal	Lethal
***atx2* null dsRED cassette**	Lethal	Fully Viable	Lethal	Lethal

Table 2: Summary of *atx2* null allele complementation assays with cDNA, human cDNA and Mini *atx2* locus replacement cassettes *in vivo* in flies. *atx2* 1-8 and X1 are previously generated null alleles of *atx2* made by independent CRISPR/Cas9 exon 1-8 excision + non-homologous end joining, and 1st exon P-element integration and imprecise excision respectively. The *atx2* null dsRED cassette is the null allele generated as part of the work of this thesis and is the parental strain in which site-directed replacement was carried out. As expected, *atx2* nulls are homozygous non-viable, however only the fly cDNA locus replacement is able to successfully complement the null alleles and lead to viable and fertile heterozygote progeny. Neither fly Mini *atx2* nor the *Drosophila* codon-optimised human cDNA are able to complement independently generated null alleles, suggesting that essential functions of the protein are missing or insufficient in both cases for animal survival but also that lethality observed is not a result of off-target or other lethal factors in the generated alleles.

Discussion

The IDRs of Ataxin-2, as previously published by our lab and recently others (Boeynaems *et al.*, preprint 2022), despite their lack of direct sequence conservation appear to clearly be an essential aspect of the protein's biological function and even structurally conserved between *Drosophila* and humans. From higher order phenomena such as long-term habituation to the pathology and cell death in neurodegeneration models, the IDRs and specifically the CIDR of Atx2 play a significant role in facilitating these phenomena via its biological function in self-assembly and granule formation within cells. In the former, deletion of the Atx2 CIDR leads to the loss of long-term habituation that we posit is reliant on mechanisms of synaptic plasticity. Particularly, there is a growing body of evidence that the physical growth and strengthening of signal transduction at activated synapses which makes up the phenomenon we refer to as synaptic plasticity result from localised translation of mRNAs transported to these synapses (McCann *et al.*, 2011; Sudhakaran *et al.*, 2014; Nakayama *et al.*, 2017; Moon *et al.*, 2019; Abouward and Schiavo, 2020). Where Atx2 ties in to this model is due to its ability to bind RNA, bind other RNA-binding proteins, and most importantly de-mix or assemble in the cytoplasm to form mRNPs, thus directly linking the granule forming ability of the IDRs to the mechanisms of long-term potentiation (Nonhoff *et al.*, 2007; Kaehler *et al.*, 2012; Yokoshi *et al.*, 2014;

Bakthavachalu *et al.*, 2018). The sushi-belt model of mRNA transport to active synapses (Doyle and Kiebler, 2011) is consistent with our increasing insight into the precise mechanism of how such transport is controlled in neurons and the role of often IDR containing RNA binding proteins such as Atx2 in it. For example, the Atx2 binding partner and stress granule component Caprin1, which contains IDRs, has been directly shown as essential for mRNA transport to synapses in mouse brains, while work focusing on Annexin A11 has tied the inclusion of this protein in potential mRNPs as critical for hitching these granules to actively retrograde transported lysosomes along microtubules (Nakayama *et al.*, 2017; Liao *et al.*, 2019). Together, we can begin to gain an understanding of how the phenomena of membrane-less organelle formation through liquid-liquid phase separation is utilised in cells to carry out otherwise thermodynamically unfavourable molecule segregation and localisation to distant sites in large cells – such as is required by localised translation in long-term synaptic potentiation that leads to long-term memory.

In the case of neurodegeneration (specifically the cellular degeneration in models of human disease), we observe that the CIDR of Atx2 is critical for facilitating this pathology in fly models (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021). As a broad concept, this is unsurprising given the long history of disordered or more precisely amino acid repeat/low complexity regions of proteins being associated with or causative of human disease such as Huntington's. Particularly Ataxin-2 itself was first identified and characterised as a SCA2 disease causing protein due to the expanded polyQ tract found in patients. However, the key detail we further probe in this thesis' work is the significance of specifically the CIDR – which does not contain the polyQ tract – on the degeneration phenotype. Our experiments examined the effect of polyQ expanded Ataxin-2 alleles on our (admittedly limited) granule formation assay in cells and we found no obvious difference in phenotype. Further experiments isolating the activity of the polyQ tract from the CIDR by adding the segment to the Mini Ataxin-2 construct may shed light on the scale of the effect the region has on granule formation or aggregation. Consistent with our observations, polyQ tract toxicity has been previously reported as being modulated by downstream, C-terminal elements of Ataxin-2 (Ng, Pulst and Huynh, 2007). An understanding emerges from the data that the CIDR's ability to drive the agglomeration of the Ataxin-2 protein is much more powerful than the effects of the MIDR and its constituent polyQ and prion-like domains. The CIDR of Ataxin-2 seems to confer the to the protein most of its granule forming propensity and this function is not only evolutionarily conserved between *Drosophila* and humans but potentially tuned and subfunctionalised in the human case between

the stronger Atxn2L and weaker Atxn2. Altogether, while removing this region lets us detect drastic phenotypes in our assays, under normal conditions and endogenous expression of Atxn2 over the lifetime of a SCA2 patient, the relatively small effect of an extended polyQ tract may critically break down the granule control homeostasis acting upon the protein in the particularly susceptible Purkinje neurons. As recent work suggests aromatic-rich peptides may be able to form stable prion-like assemblies (LARKS) it is possible an expanded polyQ is simply the most common mutation (due to CAG repeat genomic instability) that leads to more tight binding between Atxn2 molecules within condensates (Hughes *et al.*, 2018). A relatively small upset to a well tuned system could eventually lead to the formation of permanent aggregates or the right conditions to nucleate the aggregation of other pathogenic proteins.

The archaic idea of unstructured regions of proteins being more or less unnecessary in protein function is further challenged by our work. There may be little or no sequence conservation between disordered regions of homologous proteins, however the intrinsically disordered nature of these regions and – as in the case of Ataxin-2, the general domain architecture – appears to be evolutionarily significant and thus conserved. The model emerges of a biological activity – the ability to self-assemble into a reversible granule – being utilised for a variety of cellular functions that are so beneficial to animal survival (ability for long-term habituation for example) that the negative selective pressure of the most common failure state – cell death due to aggregate formation or promotion of aggregates – is insufficient for granule forming domains to be lost through natural selection (Ramaswami, Taylor and Parker, 2013; Bakthavachalu *et al.*, 2018).

We propose the existence of a hitherto unknown essential feature of the Ataxin-2 protein for fly survival. Mini Atx2, nor human Atxn2 cDNA for that matter, were unable to rescue homozygous Atx2 null lethality in flies despite theoretically possessing all previously known essential domains of Atx2. This unknown feature of Atx2 cannot be an upstream or downstream regulatory element of the gene as the mutant alleles were inserted into the same locus and confirmed to express protein. It could not be an intronic sequence either due to the cDNA replacement cassette being viable and fertile, and must instead be a short sequence, domain or simply the architecture of the domains that was missing from the mini construct (such as the necessity to space out antagonistically acting domains). Further study of the behavioural (such as long-term habituation and sleep) and degeneration modulation activity of mini Atx2 will be important as well as a piece-wise complementation of the regions not included in the

mini construct to identify what essential feature of the protein is missing from mini as well as to see if the structured domains may be able to carry out their essential roles within the cell independently of one another. However, evidence from our lab suggests w.t. Atx2 has a dominant effect on these pathways and therefore future testing of the homozygous lethal Mini Atx2 will have to rely on conditional removal of w.t. protein or gene in adult animals. Such a dissection of the activity of the domains may begin to shed light on the great question of how different isoforms of *atx2* tie into the protein's essential functions in animals and the cellular control framework for different functions of the protein.

Finally, synthesising all the findings of this chapter, we can conclude that the implied generic granule forming ability of Ataxin-2 IDRs is in fact tightly regulated due to the problems such an inherently promiscuous phenomenon may cause to the otherwise well organised cytoplasm filled with phase-separated compartments and membraneless organelles (Banani *et al.*, 2017; Verdile, De Paola and Paronetto, 2019). Not only does the CIDR, and thus Ataxin-2 granules, have distinct functions in stress granule formation and mRNA transport in long-term habituation, it does not form granules by itself alone, and outside of these defined roles unless in pathologic contexts of overexpression or disease models. The natural question arose and is examined in chapter 2 of this thesis, inspired by the work of Protter *et al.*, 2018, at the time, and more recently Sanders *et al.*, 2020 and Boeynaems *et al.*, preprint 2022: do the structured domains of Ataxin-2 contribute to or affect the promiscuous interactions and activity of the IDRs?

2. The structured domains of Ataxin-2 modify and direct the other activities of the protein, both normal and pathologic

In *Drosophila melanogaster,* there is one genetic ortholog to *Atxn2* and *Atxn2L* referred to as *Atx2* and located on chromosome III (Satterfield, Jackson and Pallanck, 2002). Homozygous deletion of the Atx2 gene in Drosophila has been reported as being lethal during embryo development (no larvae are reported hatching from eggs) from the moment the first mutants were generated. Flybase records further detail on the viability of other published Atx2 nulls and mutants, (*FlyBase Gene Report: Dmel\Atx2*) with the consensus being the same. The $Atx2^{1\text{-}8}$ allele, generated by the Ramaswami lab for Bakthavachalu *et al.*, 2018 by excising exons 1 to 8 using CRISPR/Cas9 technology and non-homologous end joining similarly, independently shows the essential nature of the Atx2 gene in flies. Notably, Atx2 has two homologs in humans and in vertebrates in general: Atxn2 and Atxn2 like (Atxn2L). Homozygous deletions of Atxn2 are not lethal in mice due to apparent redundancy of function, while Atxn2L knockout mice show mid-gestation death – more similarly to the expectation from the fly case (Lastres-Becker *et al.*, 2008; Fittschen *et al.*, 2015; Key *et al.*, 2020). Recalling the data previously shown and discussed in Chapter 1, we know that the structured domains of Atx2 carry out some of these essential functions of the protein in at least some cells or during some stages of cell or organismal development that lead to the lethal phenotypes of domain-deleted mutant homozygotes (Bakthavachalu *et al.*, 2018). Furthermore, dissection of the granule-forming phenotypes of the Ataxin-2 CIDRs in cell assays further implied that these same structured domains may modulate the activity of the IDRs. The biological question that this chapter will attempt to answer therefore becomes clear: do the structured domains of Ataxin-2 affect the normal and pathogenic activities of the IDRs, and how?

Ataxin-2 has three distinct, folded domains that are well conserved between *Drosophila* and human homologs and paralogs. These are the LSm, LSmAD and PAM2 domains (in order of location from N-terminal to C-terminal) (Jiménez-López and Guzmán, 2014). There are other sequences that are directly conserved between distant homologs of Ataxin-2, such as the poly-glutamine tract which is found both in Human Atxn2 and fly Atx2, however, its nature and proposed activity (briefly examined in the previous chapter) as a low-complexity domain is closer to that expected of a typical

IDR. While the glutamine tract contains the same sequence (CAG repeats with one or two CAA alternative codons), the location and length of the tract are different between fly and human. Therefore we will focus on the three folded domains and their structured interactions in these experiments to specifically tie in Ataxin-2 into the generalised, hypothetical model of aggregate formation presented in (Protter *et al.*, 2018).

We will dissect the structured domains, first in terms of granules (i.e. they do not form in our cell assays due the properties of the CIDR alone, yet the CIDR is essential for their formation), the specific binding partners of the domains (previously identified through IP-WBs, mass spectrometry, and IP-RNAseq), and finally test how these structured domains affect the higher order phenomena that Ataxin-2 is involved in – particularly modulation of neurodegeneration models. The subsequent work from our lab has already utilised Targets of RNA-Binding Proteins Identified by Editing (TRIBE) technology to identify the subset of mRNAs targeted by Atx2 in the *Drosophila* brain and S2 cells (Singh *et al.*, 2021). The observed trend that Atx2 interacts with AU-rich elements in 3′UTRs of its targets already highlights the interaction with the Poly-A tail Binding Protein (PABP) as potentially being important for the selectivity of what gets included in proposed cytosolic Atx2 granules in healthy cells. These granules could be essential to Atx2's function in modulating the stability and translation of its target mRNAs as subsequent analyses of Atx2 domain deletions demonstrated the CIDR and therefore granule forming ability is essential for Atx2 to interact with target mRNAs within mRNPs. The LSm domain was identified as having an antagonistic effect to the CIDR and reduce granule formation, and thus Atx2 must have essential functions outside of mRNPs (Singh *et al.*, 2021). Finally, we identify the specific interaction between the PAM2 domain of Atx2 and PABP as essential for defining the contents and thus identity of any mRNPs that Atx2 forms. The nature of the granules – their content, and not granules in general will be shown as key to their downstream promotion of the pathology in *Drosophila* degeneration models (Petrauskas *et al.*, 2022, BioRxiv preprint). Altogether, we add to the understanding of neuronal translational control mechanisms and even posit a potential target for subtracting the effect of Ataxin-2 from granules in which it may catalyse the pathogenic processes in neurodegenerative disease.

Atx2 structured domains alone are insufficient to induce granule formation in fly S2 cells

Within the context of the activity of the Ataxin-2 proteins IDRs, the follow-up question arose: is this granule forming propensity controlled or modified by the structured domains, and if so, how? We know from work carried out in our lab and others, and also replicated as part of my own experiments for this thesis, that the Atx2 construct without the CIDR cannot form granules efficiently in our S2 cell overexpression assay nor can the CIDR form granules without additional domains (Bakthavachalu *et al.*, 2018). In the case of the former experiments, approximately 30% of Atx2ΔCIDR transfected cells contained granules, which were morphologically different and exhibited much more liquid-like molecular dynamics compared to WT Atx2 overexpression granules. However, the granules that did form showed the same, expected protein components colocalising, implying they were still compositionally similar to WT Atx2 granules and SGs.

To further narrow down the effects of the structured domains alone on granule formation in our assay, we designed radically truncated “minimal” Atx2 expression constructs (mini Atx2). Specifically, pUASt plasmids containing only the LSm, LSm-AD, and PAM2 domains, linked together with flexible glycine-serine spacer sequences that form loops as a secondary amino acid structure and containing a codon-optimised C-terminal SNAP tag were generated and transfected into S2R+ cells, driven by the actin promoter Gal4-UAS system. The resulting data showed the complete inability of these constructs to induce granule formation in this assay even after prolonged (72H) overexpression, strengthening our previous conclusions about the effect being driven primarily by the IDRs and the CIDR most significantly (figure 16). Taking into account the CIDR-only construct’s inability to form granules, an additive effect between the structured domains and the IDR must take place. The alternative explanation is that additional IDRs – defined as any sequence outside of the three structured domains of Atx2 – may be required for the protein to exhibit granule formation in this specific assay, despite the intrinsic ability of the CIDR to phase separate by itself in *in vitro* assays.

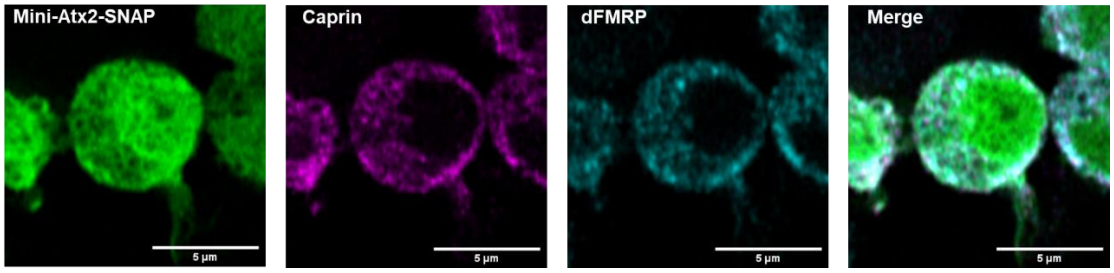

Figure 16: The structured domain only construct of Atx2 is incapable of forming granules when overexpressed in S2R+ cells. *Drosophila* S2R+ cultured cells were transfected with Act-Gal4 driven UAS Mini Atx2 plasmids. The protein product was overexpressed for up to 72H however showed a diffuse distribution throughout the cytoplasm – including diffusing within the nucleus due to the small size of the protein. Co-immunostaining and imaging of the Atx2 partner proteins and SG markers Caprin and dFMRP show that the construct overexpression and general methodology of this assay did not induce SGs. Scale bar = 5µm.

Structured domains of Atx2 are sufficient for the proteins' recruitment to oxidative SGs in fly S2 cells

We made efforts to distinguish granule induction ability in cells from protein recruitment to stress granules, assayed by transfecting cells with the required constructs and subjecting the cells to oxidative stress by sodium arsenite, before immunofluorescent staining and confocal imaging. As described before, Ataxin-2 IDRs are the necessary domains of the protein for granule induction in unstressed cells. However, we observed that mini-Atx2 transfected cells were firstly able to form SGs, but secondly that the SGs recruited and concentrated the expressed mini-Atx2 protein to them. After transfection, expressed mini-Atx2 protein remained diffuse throughout the cell (including nucleus) however upon stress, a portion of the protein would concentrate into puncta of high fluorescence intensity (therefore implying high density of protein) with a descending gradient of intensity towards the background, diffuse fraction of mini-Atx2 protein. These areas of high intensity (puncta) co-localised with and shared analogous profile plots with stainings for stress granule components (dFMRP and Caprin) – indicating that the construct was indeed recruited to SGs and not other or non-specific puncta (figure 17). This finding allowed us to posit that, while insufficient for granule induction or self-assembly by themselves, the structured domains – or at least one of them – were sufficient for Atx2 recruitment to SGs. The implication of this for our model was that structured domains may affect or influence the activities Ataxin-2 protein that are

also reliant on its IDRs, i.e. the structured domains play a role in SG inclusion, and the IDRs induce and contribute to their assembly.

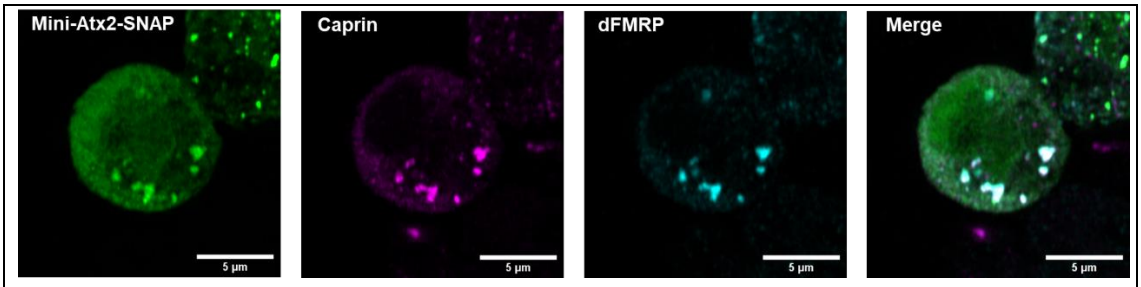

Figure 17: The structured domains alone are sufficient for mini Atx2 recruitment to oxidative stress induced SGs in S2R+ cells. *Drosophila* S2R+ cultured cells were transfected with Act-Gal4 driven UAS Mini Atx2 plasmids for and subjected to oxidative stress by the addition of sodium arsenite (final concentration 0.5mM) for the last 3H of the 72H transfection. The overexpressed mini Atx2 protein formed into visible puncta in the cytosol that were significantly brighter than the background diffusely distributed fraction of the protein. Co-immunostaining and imaging of the Atx2 partner proteins and SG markers Caprin and dFMRP confirm that SGs were formed in transfected stressed cells and that mini Atx2 puncta co-localise with these SGs, and thus must be recruited to them under these conditions. Scale bar = 5µm.

The PAM2 domain is specifically necessary and sufficient for Atx2 recruitment to oxidative SGs in S2 cell assays

Naturally, our follow-up experiments aimed to figure which one or what combination of structured domains of Atx2 are sufficient for recruitment to SGs. Single domain deleted mini Atx2 constructs were generated and tested in the same S2 cell assay as w.t. mini-Atx2 described earlier. All mutants behaved identically to w.t. mini Atx2 in transfected, unstressed cells. In arsenite stressed cells, LSm and separately Lsm-AD deleted mini-Atx2 constructs retained their ability to be recruited to SGs and showed identical morphology and co-localisation profiles to the SG marker Caprin as w.t. mini Atx2. The PAM2 domain deleted mini Atx2 did not get recruited to SGs after stress and remained diffuse in the same way as in unstressed cells (figure 18). This exciting finding identified the short PAM2 domain – the PABP, and through it, mRNA poly-adenosine tail binding region (Deo *et al.*, 1999) – as the essential feature Atx2 that is sufficient for the proteins' recruitment to SGs. Furthermore, point mutants in the mini Atx2 PAM2 domain designed to specifically block the interaction with PABP replicate the loss of

recruitment to SGs seen with the whole domain deletion (figure 18). F866A and L859A substitutions in the PAM2 domain (as well as double substitutions) exchanged the previously described and modelled amino acids that were key to binding the MLLE domain of PABP (Jiménez-López and Guzmán, 2014). This additional experiment helps exclude potential other effects of the full PAM2 domain and isolate the specific role of PABP binding as the key mechanism behind the observed Atx2 recruitment to SGs. It also implicated Atx2 proteins' association with mRNA poly-A tails as the general pathway that leads to its inclusion in SGs, which was significant in countering the potential model that IDRs alone might bring Atx2 to granules or that another protein partner was mainly responsible. While likely not the exclusive activity driving this observation, we suggest for Atx2, the PAM2-PABP interaction may be the main and fundamental interaction driving recruitment to and in turn formation and maintenance of SGs in S2 cells, in which Atx2 plays a key role in their formation and maintenance.

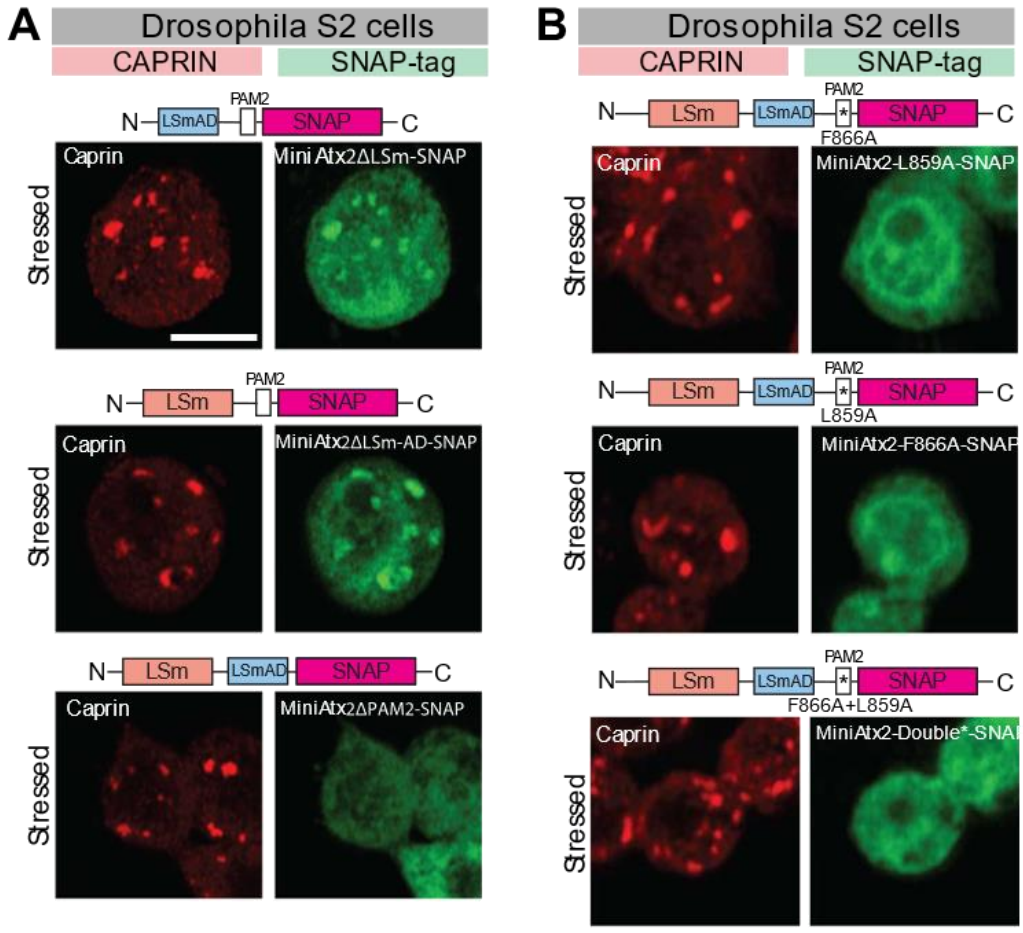

Figure 18: The PAM2 domain, and specifically its interaction with PABP is necessary for mini Atx2 recruitment to SGs in oxidatively stressed *Drosophila* S2R+ cell assay. (A) Single domain deleted mini Atx2 constructs were tested for recruitment to sodium

arsenite induced SGs in S2R+ cells. Only the mini Atx2 ΔPAM2 protein failed to show puncta in transfected stressed cells, despite unimpaired SGs formation in these cells as evidenced by the staining for the SG marker Caprin. Thus it was shown that that the PAM2 domain specifically may be the key structured interaction that recruits Atx2 to SGs during their formation. (B) Single amino acid substituted point mutants in the PAM2 domain that disrupt the binding with PABP were sufficient to replicate the loss of mini Atx2 recruitment to SGs seen with the full PAM2 domain deletion. F866A, L859A and the double mutant forms of mini Atx2 would be unable to interface with the MLLE domain of PABP, and thus this data shows the importance of this interaction specifically for the recruitment of Atx2 to SGs in fly cells and implicitly its subsequent functions therein. Scale bar = 5μm. Point mutant constructs were designed and imaged in collaboration with Daniel Fortunati. The figure is adapted from (Petrauskas *et al.*, 2022).

The PAM2 domain determines recruitment of Ataxin-2 to SGs in mammal cells and the mechanism is conserved between *Drosophila* and human proteins and cells

To strengthen the interpretability of the model and make it more general to all Ataxin-2 proteins, we repeated analogous experiments in human cell lines. The short length and very high conservation of the PAM2 domain allowed us to speculate that the mechanism of recruitment to SGs, shown here by staining for the SG marker G3BP1/2, will be similarly conserved between flies and humans despite the phylogenetic distance between the species and observed differences in the IDR “strength” between the proteins. Human mini Atxn2 constructs under the control of CMV-Gal4>UAS expression system were generated and transfected into U2OS human-derived cell lines. In unstressed conditions, the expressed construct remained completely diffuse throughout both cell types. Similarly to fly mini Atx2, after oxidative stress was induced in cells with sodium arsenite, transfected cells formed SGs and recruited mini Atxn2 protein to them, as was expected. Using single-domain mutant mini Atxn2 constructs we further confirmed it was indeed the same structured interaction, reliant on the PAM2 domain, that recruited mini Atxn2 to SGs as the PAM2 deleted mutant was the only completely diffuse stained construct in arsenite stressed conditions (figure 19).

Similarly, we replicated the homologous point mutants in the PAM2 domain of human mini Atxn2 to determine if the implied interaction with PABPC1 (cytoplasmic homolog to fly PABP) is the only structured interaction of the domain required for Atxn2 recruitment to SGs. L914A, (homologous to L859A in fly Atx2) and F921A (homologous to F866A) single amino acid substitutions, as well as and a triple mutant that included the Q928A substitution which was identified and modelled as key to binding the MLLE domain of human PABPC1 but did not have a clear homolog in fly Atx2, all prevented the constructs' recruitment to SGs. Once again, this single structured interaction between the PAM2 domain and PABP is identified as essential for the mini Atxn2 construct's recruitment to SGs in this assay. PABPC1 itself binds the poly-A tails of mRNAs – an immensely large set of targets – and thus allows for Ataxin-2 to enter into proximity to these molecules and any other associated proteins. While other domains, such as the LSm and the CIDR have intrinsic RNA binding ability, we suggest and will examine further in subsequent sections, that the PAM2:PABP binding is the major method of endogenous Ataxin-2 localisation within the cytosol as part of the normal functions of cells (Yokoshi *et al.*, 2014; Singh *et al.*, 2021).

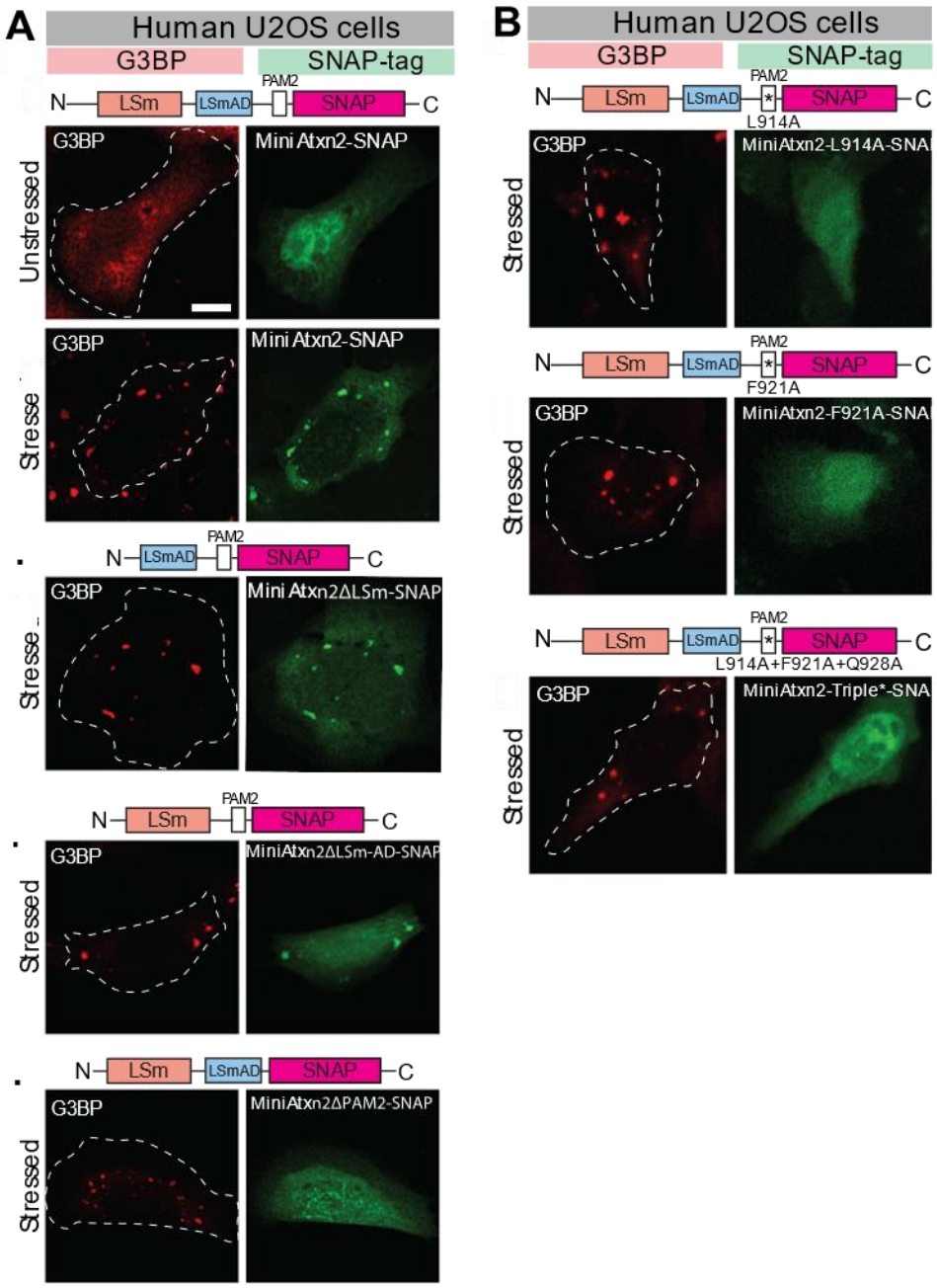

Figure 19: The PAM2 domain, and specifically its interaction with PABP is analogously necessary for human mini Atxn2 recruitment to SGs in oxidatively stressed human U2OS cell assay. (A) In an analogous fashion to the *Drosophila* experiments, human mini Atxn2 constructs expressed for 48H using the CMV-Gal4>UAS system were not able to induce granules in unstressed U2OS cells, but were recruited to SGs after they were induced for 1H with sodium arsenite (0.5mM). Single domain deleted human mini Atxn2 constructs were tested to confirm the necessity of the PAM2 domain for this recruitment and confirmed the conservation of this mechanism between humans and flies. Mini Atxn2 ΔPAM2 was the only domain deleted construct that failed to show granules in transfected stressed cells, despite

unimpaired SGs formation in these cells as evidenced by the staining for the SG marker G3BP. (B) Single amino acid substituted point mutants homologous to fly Atx2 in the human Atxn2 PAM2 domain that disrupt the binding with PABPC1 were sufficient to replicate the loss of mini Atxn2 recruitment to SGs seen with the full PAM2 domain deletion. L914A, F921A and a triple mutant including the non-conserved Q928A interacting amino acid substituted forms of mini Atxn2 failed to be recruited to SGs. Thus, in the same way to fly Atx2, removing the interaction with the MLLE domain of human PABPC1 disrupts the recruitment of human Atxn2 (and potentially Atxn2L which shares a near-identical PAM2 domain) to SGs in human cells and implicitly its subsequent functions therein. Dashed lines indicate the outlines of the same cells in different stainings displayed. Scale bar = 10μm. Point mutant constructs were designed and imaged in collaboration with Daniel Fortunati. The figure is adapted from (Petrauskas *et al.*, 2022).

The PAM2 is the shortest and most conserved domain between human and *Drosophila* homologs of Ataxin-2. It is also a well characterised domain known to primarily bind the MLLE domain of PABP – itself one of the most highly conserved proteins between eukaryotes (Xie, Kozlov and Gehring, 2014). The similarities of these domains led us to posit that the interaction is necessary for Ataxin-2 recruitment to SG specifically, and potentially granules in general, is a highly important and thus conserved one. We tested if the function of the PAM2 domains was retained between species in the same manner the sequence is conserved and thus we exogenously expressed *Drosophila* mini Atx2 in human U2OS and HEK293T cells, as well as human mini Atxn2 in fly S2 cells. Transfected cells of both species, when under oxidative stress conditions, formed SGs and recruited the homologous mini Ataxin-2 construct to them in all cases (figure 20). We therefore know that not only is the PAM2 domains' function in SG recruitment conserved in humans, but that the domains themselves are conserved sufficiently to display inter-species compatibility in our particular assay.

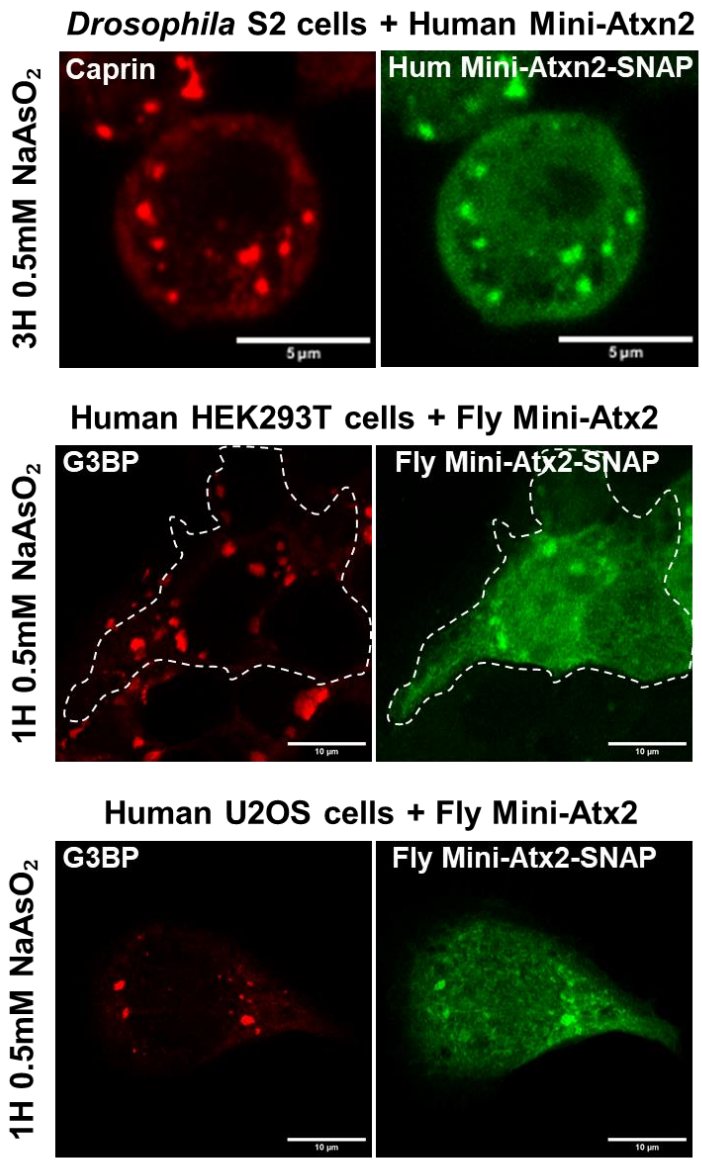

Figure 20: Mini Ataxin-2 recruitment to induced stress granules by the PAM2 domain is conserved between humans and *Drosophila* in cross-species transfected cell assays. Human mini Atx2, transfected into *Drosophila* S2R+ cells was recruited to sodium arsenite induced SGs, shown by the Caprin marker protein staining. Similarly, fly mini Atx2 was recruited to induced SGs in human HEK293T and U2OS cells – shown by the staining against the marker G3BP. Taken together and in light of previous experiments identifying the highly conserved PAM2 domain as essential for this recruitment, we posit that this function of the PAM2 domain is highly conserved in animals and thus crucial to survival and cellular and organismal fitness. Mini Ataxin-2 constructs were codon optimised for the expression host species. Sodium arsenite stress was carried out for 3H in S2 cells and 1H in HEK293T and U2OS cells at a final concentration of 0.5mM. Gal4-UAS expression system was

used in all experiments with Gal4 expression in fly cells driven by the Actin promoter, and in human cells – by the CMV promoter. S2 cell scale bar = 5μm; human cell scale bars = 10μm.

Structured interactions with PABP and LSM12 are maintained by mini Ataxin-2 even outside of granules

With the generation of mini Ataxin-2 constructs, we were able to separate the functions of the structured domains of the proteins from the functions of the IDRs and in-turn granules. However, having identified recruitment to SGs to be facilitated by the PAM2 domain, we still sought to verify this domain's direct association with PABP, and also to determine when this interaction takes place. Potentially, it could have occurred only in granules and during the formation of SGs, yet the data from IP-WB experiments carried out using mini Ataxin-2 constructs in mammalian and fly cultured cells suggested that it takes place even outside of granules, when Ataxin-2 is diffusely distributed. Mini Ataxin-2 is able to pull-down PABP in transfected, unstressed human and fly cells (figure 21). Under these conditions, the construct is diffusely distributed throughout the cytoplasm yet still interacts with PABP, as well as LSM12 – a partner protein in an translational activation complex (Lee *et al.*, 2017). Furthermore, utilising single-domain deleted mini Ataxin-2 constructs, we showed that the PAM2 domain is expectedly necessary for PABP binding and that point mutants previously shown to disrupt the recruitment to SGs expectedly disrupt the pulldown of PABP by mini Ataxin-2 from the cytoplasm. The LSm domain, similarly was identified as necessary for the pulldown of LSM12. Together, these data further verify our model of the PAM2 domain's localising activity on the protein, and show that PABP and LSM12 interactions are independent of granules in our assay. This finding in-turn may suggest a plausible order in which Ataxin-2 recruitment to granules takes place, since PABP binding drives it yet is already present in unstressed conditions and without Ataxin-2 granules present, suggesting it takes place first and foremost in the process.

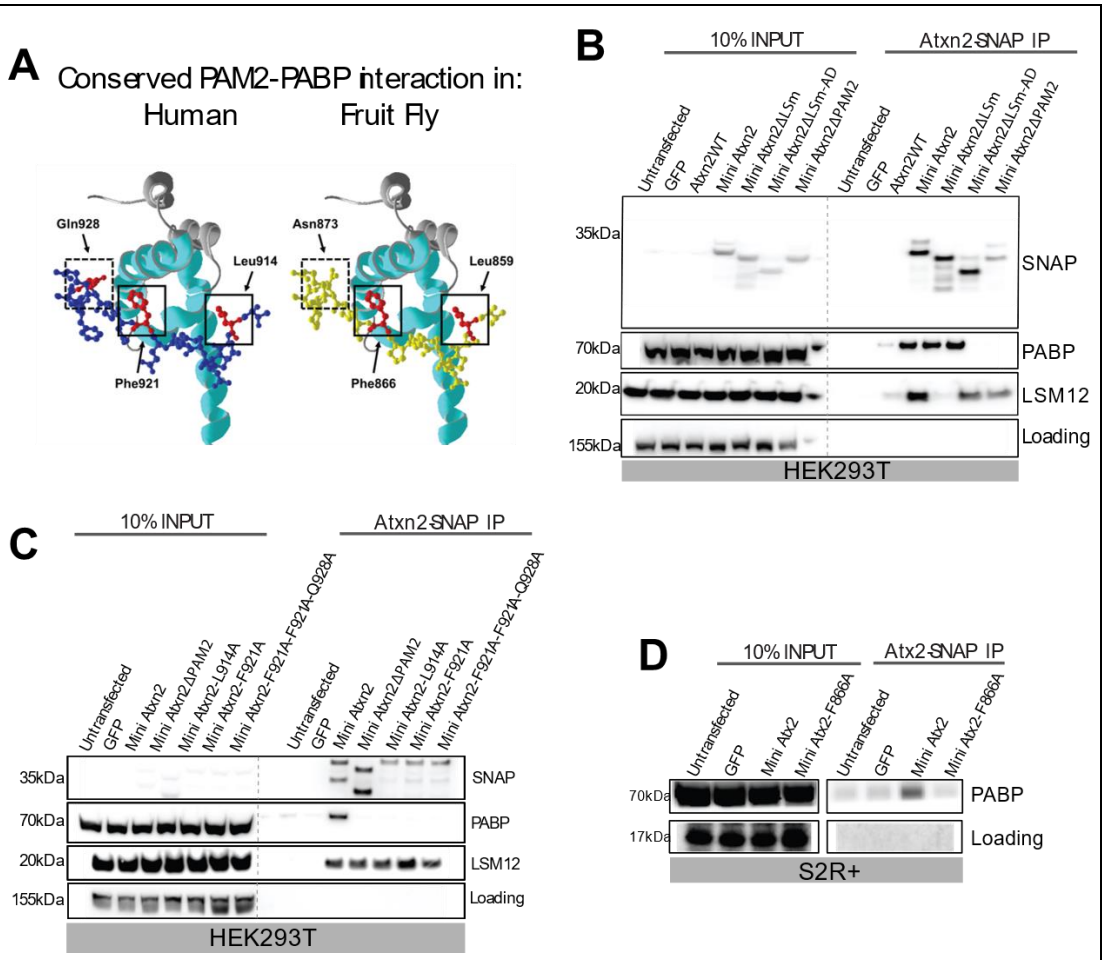

Figure 21: A minimized Ataxin-2 construct containing only the known structured domains maintains the ability to interact with PABP and LSM12. (A) Structural model of the PABP MLLE domain (ribbon) showing the near-perfect structural similarity of the human Atxn2 PAM2 domain (blue, uniprot ID: Q99700) with the Drosophila Atx2 PAM2 domain (yellow, uniprot ID: Q8SWR8). The key interacting residues are highlighted in red. (B) Human minimized Atxn2 SNAP IP-WB from HEK293T cells probing for PABP and LSM12 showing the effects of different domain deletions. The PAM2 domain is necessary and sufficient for the Atxn2-PABP interaction, while the LSm domain is necessary and sufficient for the Atxn2-LSM12 interaction. (C) Point-mutations targeting key interacting residues of the PAM2 domain were predicted to replicate the effect of a full PAM2 deletion in the minimized Atx2 construct. Human minimized Atxn2-SNAP IP WB from HEK293T cells showing that mutating either of the key hydrophobic residues L914 or F921 in the PAM2 domain is sufficient to prevent its interaction with PABP. The interaction with LSM12 is unaffected by the point mutations. (D) *Drosophila* minimized Atx2-SNAP IP-WB from S2 cells. An analogous PAM2 domain point mutation on F866 blocks the Atx2-PABP interaction.

Point mutant data was produced with Daniel Fortunati. Figure is adapted from (Petrauskas *et al.*, 2022).

The PAM2 domain affects the protein composition of Ataxin-2 granules in cells

Having carried out characterisation of the PAM2 domain outside of Ataxin-2 granules in the context of the structured domain only containing mini Ataxin-2 constructs, the question naturally followed of how the domain affects the formation and nature of such granules in cells. We used the S2 cell overexpression assay with full length PAM2 deleted Ataxin-2 constructs to, firstly show that the PAM2 domain is not necessary for granule formation in this assay, but secondly that the granules formed without the domain lack a number of typically present partner proteins (figure 22). dFMRP, Caprin and of course PABP are missing from these atypical granules, while Me31B and ROX8 remain included. What this suggests is that the IDRs of Atx2 have a generic granule forming ability but that structured interactions with PABP and in turn its target polyadenylated mRNAs define the internal environment, contents and possibly functions of granules formed. Recent work suggests that the contents of biological condensates define their identity from otherwise useless generic granules and mRNPs and thus allow for specific activities to take place within them (Sanders *et al.*, 2020). In the case of Ataxin-2, this could be the remodelling of a translational activation complex into a translational repression complex, potentially even on the same associated mRNA as it goes through its lifespan in the cytosol (Lee *et al.*, 2017). Finally, the CIDR is once again shown as the key feature driving the phenomenon of granule formation as pairwise deletion of it and the PAM2 or LSm domains removes this ability, even in the case of the latter where the domain antagonistically acts against granule forming propensity as published by our lab (Singh *et al.*, 2021). Critically, we verify our model of how the activity of IDRs in general and the Ataxin-2 CIDR specifically, must and is tightly controlled by structured interactions for normal biological function of the protein in cells – via PAM2 for localisation and LSm for inhibition.

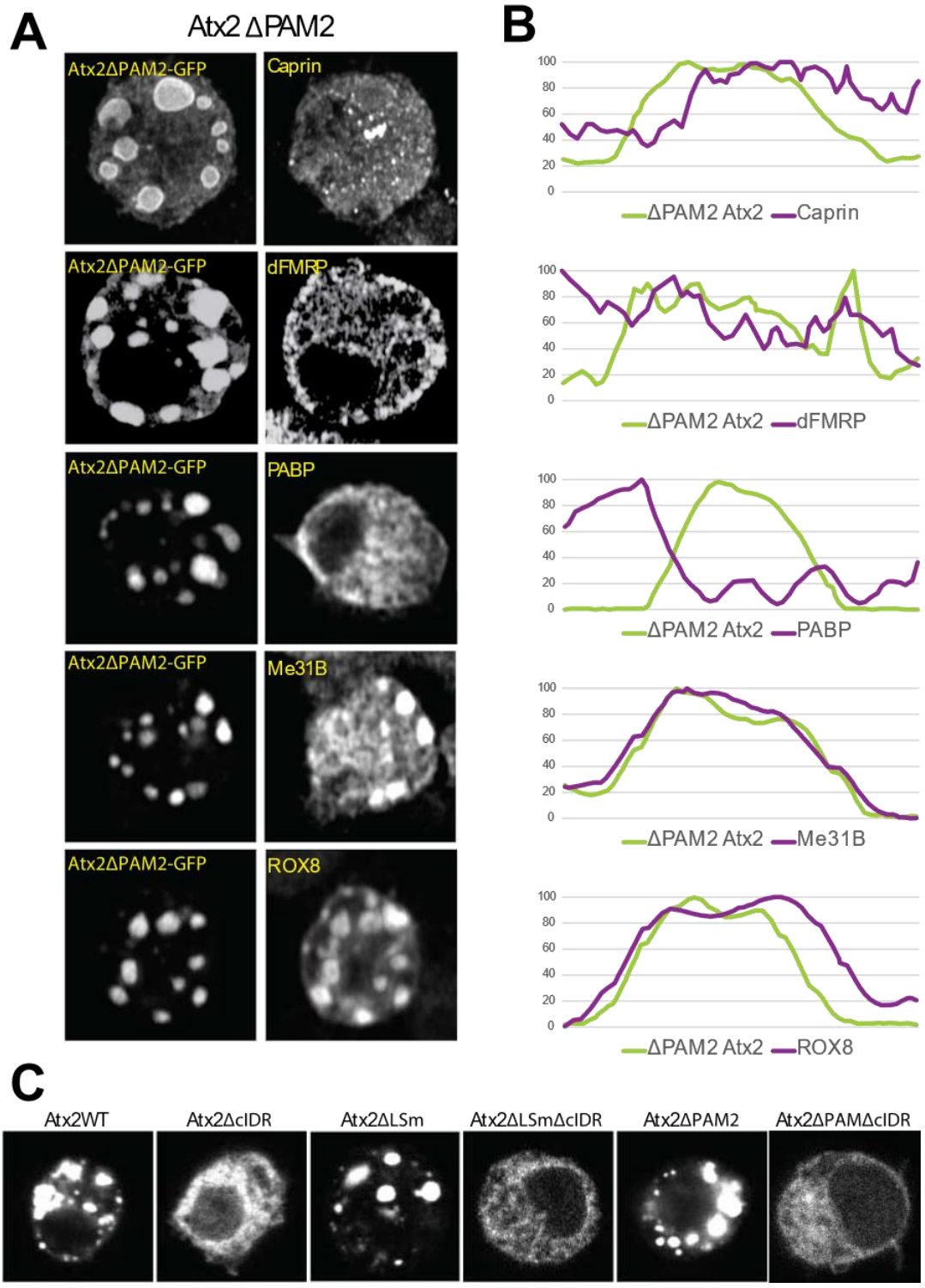

Figure 22: Presence of the PAM2 domain affects the protein composition of Atx2-GFP granules in S2 cells. (A) While w.t. Atx2 overexpression generates granules that co-localise with stress granule markers, deletion of the PAM2 domain affects the Atx2-GFP granule composition. Over-expression of Atx2ΔPAM2-GFP in S2R+ cells still induces the formation of granules, but some SG markers fail to co-localize in these, notably dFMR, Caprin and PABP. (B) Normalised profile plots of fluorescent

signal co-localisation with ΔPAM2 granules for each of the components tested for and images displayed in (A). See table 3 for quantification. (C) Atx2-GFP granule formation in S2 cells relies primarily on the CIDR. Deletion of the CIDR in Atx2WT, Atx2ΔPAM2 and Atx2ΔLSm, removes their ability to form granules upon overexpression. Design and imaging of constructs in (C) was carried out by collaborators in Baskar Bakthavachalu lab and Aman Singh. Scale bar = 5 μm.

Co-localisation of signal enrichment in granules

	ΔPAM2 Atx2-GFP granules
Caprin	7.1%
dFMRP	2.4%
PABP	14.6%
Me31B	52%
Rox8	100%

Table 3: Quantification of co-localisation for stress granule components to ΔPAM2 Atx2-GFP granules. From experiments shown in Figure 22, N = 48-120 images of Atx2-GFP granules were randomly selected for each co-staining with the stress granule components – Caprin, dFMRP, PABP, Me31B, and ROX8 – and analysed for signal co-enrichment (see methods) in the case of each component assayed. Caprin, dFMRP and PABP co-localisation was observed as significantly reduced compared to w.t. Atx2-GFP granules due to the lack of the PAM2 domain. Me31B co-localisation was impaired but had at a lower w.t. co-localisation baseline due to high background levels of this protein known to be diffusely expressed in cells. Data published in preprint (Petrauskas *et al.*, 2022).

The PAM2 domain defines the majority of the mRNA targets of Ataxin-2 in granules in *Drosophila* neurons

Given the PAM2 domain was shown as affecting some of the protein contents of Atx2 granules in our overexpression induced granule assay, we aimed to determine if the same can be true with regards to the mRNA contents and binding partners of the protein. The broad binding of the PAM2 partner PABP to polyadenylated mRNAs suggested that this very likely to be true. Our lab's and collaborators' recent publication used Targets of RNA-Binding Proteins Identified by Editing (TRIBE) technology to identify mRNAs associated with Atx2 in the Drosophila adult brain (McMahon *et al.*, 2016; Singh *et al.*, 2021). Briefly, Atx2 was fused with the catalytic domain of *Drosophila* ADAR (ADARcd) – an RNA-editing enzyme – and fly lines carrying this fusion protein under the control of the UAS promoter were generated using CRISPR/Cas9 genomic insertion. This allowed for *in vivo* editing of mRNAs associated with the Atx2 fusion protein to take place in adult fly brains using elav-Gal4 and $Gal80^{ts}$ to control special and temporal expression. After RNA extraction and sequencing, these edits on specific mRNAs could be identified using computational methods and thus a set of 256 Atx2 target mRNAs was found. Further, the editing of mRNAs was shown to be dependent on the CIDR of Atx2 and thus some form of granules in fly neurons, while mutants lacking the LSm domain, both edited Atx2 TRIBE target RNAs and formed mRNP granules in S2R+ cells more efficiently than the wild-type. Taken together, Atx2-ADARcd editing of target mRNAs was shown to occur in and be reflective of mRNP granule assembly. We thus aimed to replicate these experiments with PAM2 deleted constructs to determine if it was responsible for selecting the mRNA targets of Atx2.

In contrast to Atx2 forms lacking LSm or LSm-AD domains (Singh et al., 2021), Atx2ΔPAM2-ADARcd edited significantly fewer RNA targets than wild-type Atx2-ADARcd (108 genes and 165 edits vs 256 genes and 317 edits, figure 23). More strikingly, the cohort of mRNAs edited by the ΔPAM2 mutant form differed extensively from the largely overlapping cohorts edited by either wild-type forms of Atx2. Of the 108 genes edited by Atx2ΔPAM2-ADARcd, 36 were also targets of w.t. Atx2-ADARcd, while the remaining 72 were unique. 50 edit sites were common between the Atx2ΔPAM2 and Atx2WT targets. Those sites were edited with much lower efficiency in Atx2ΔPAM2 as compared to w.t. Atx2. Meanwhile, the location of edits made by Atx2ΔPAM2-ADARcd also differed dramatically as to where they occurred relative to the coding sequences of the target mRNAs. While edits made by wild-type and ΔLSm

forms of Atx2-ADARcd were greatly enriched in the 3'UTR of the mRNAs, Atx2ΔPAM2 targets were edited indiscriminately all along the mRNA length.

Taken together, these data identify the PAM2 motif as necessary for Atx2 engagement with its correct mRNA targets. This is unsurprising given that the PAM2 motif interacts with PABP, which binds polyA tracts at the 3' end of mRNAs (Deo *et al.*, 1999). Therefore, the data point to a role for the structured PAM2:PABP interaction in guiding the association of Atx2 with mRNAs and their other bound proteins, and for subsequent inclusion of these mRNAs in Atx2-containing granules. This single structured domain and its binding is the major determinant of Atx2 localisation, protein and mRNA partners, and the utilisation of its CIDR and granule forming ability for biological functions.

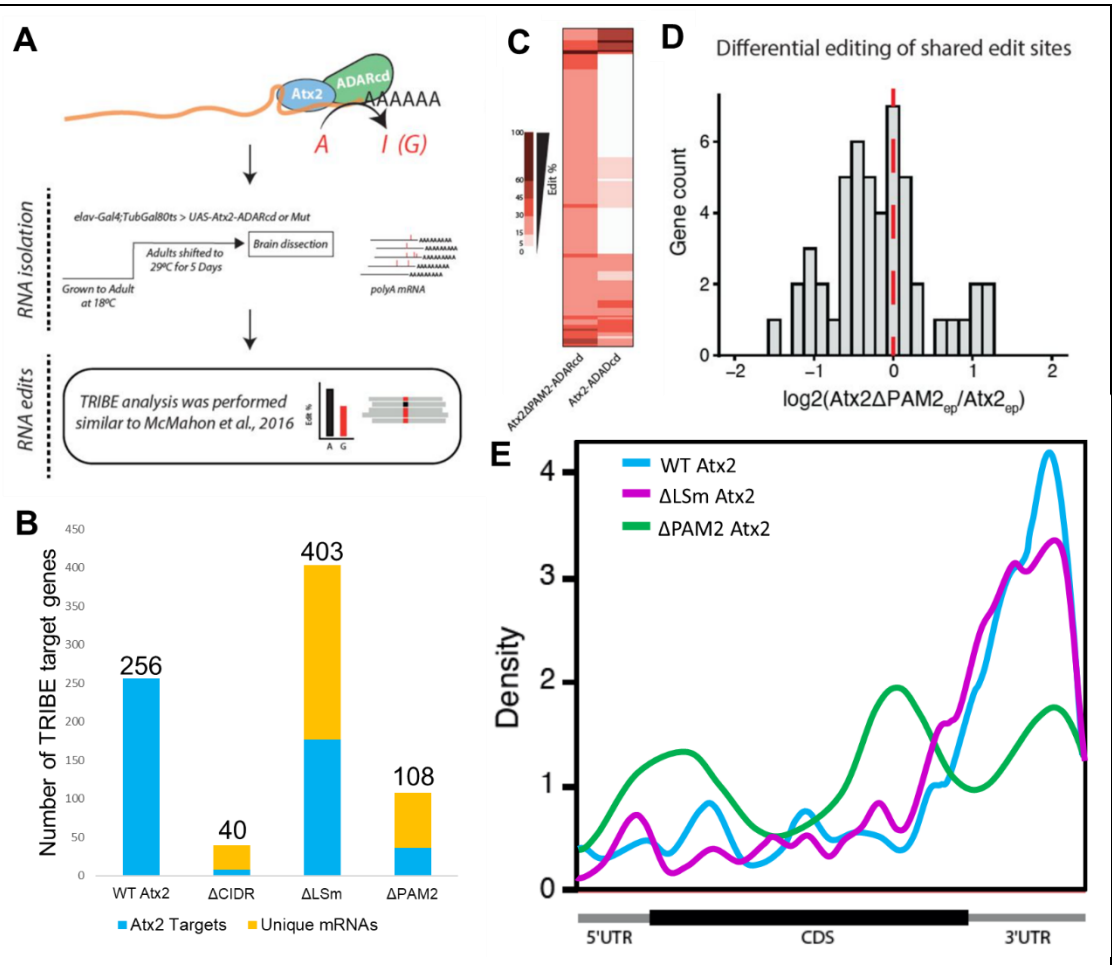

Figure 23: The PAM2 domain facilitates the selection of Atx2 RNA targets.
(A) Flowchart depicting the TRIBE analyses pipeline. Atx2ΔPAM2-ADARcd was expressed in adult Drosophila brains. Total brain RNA was isolated and RNA edits were identified and compared to Atx2-ADARcd, similar to Singh *et al.* 2021. (B)

Comparisons of genes identified by TRIBE between w.t. Atx2-ADARcd and Atx2ΔCIDR-ADARcd, Atx2ΔLSm-ADARcd, and Atx2ΔPAM2-ADARcd targets, showing the proportions of recurrent target mRNAs and unique hits in each case. N= total number of genes edited and identified above threshold. (C) Most Atx2ΔPAM2 targets identified by TRIBE are unique and not edited in w.t. Atx2, suggesting that these new targets bound by Atx2ΔPAM2 are not native Ataxin-2 granule targets. (D) Comparisons of the editing efficiency ratio of common edits between w.t. Atx2 vs Atx2ΔPAM2 show a much lower editing efficiency in Atx2ΔPAM2 compared to w.t. Atx2. (E) PAM2 deletion results in loss of 3'UTR specificity seen in w.t. Atx2 and LSm deletion TRIBE target mRNAs. W.t. Atx2 and Atx2ΔLSm-ADARcd data are extracted from (Singh *et al.* 2021). All work in this figure was carried out by collaborators in NCBS Bangalore: Baskar Bakthavachalu, Aman Singh, and Arvind Reddy Kandi. Figure is adapted from (Petrauskas *et al.*, 2022).

The IDR and PAM2 domains promote and the LSm domain inhibits cytotoxicity in Drosophila neurodegeneration models

Recalling the significance of the PAM2 domain for the control of Atx2 granule content and thus identity, what remains to be seen is the effect of this domain and therefore the effect of the aforementioned granule identity on their higher order functions. Of such functions, we have previously identified Atx2 granules as necessary for long-term olfactory habituation and neurodegenerative model severity in flies (Bakthavachalu *et al.*, 2018). The former remains an exciting avenue for future research, however we focused on the effects non-typical ΔPAM2-Atx2 granules might exhibit in the context of *Drosophila* degeneration models. As in the case of SCA2 in humans, the model used relied only on dysregulation of Atx2 by utilising generated fly lines with UAS-Atx2-ADARcd or mutant form insertions. Under the control of the elav promoter driven Gal4, UAS-Atx2-ADARcd alleles could be selectively overexpressed in neurons, while mef2-Gal4 allowed for overexpression in muscle cells. In both of these models, such overexpression is proposed to increase Atx2 granule formation in targeted cells (analogously to the induction of granules by Atx2 overexpression in cultured S2R+ cells) and leads to progressive degeneration phenotypes in adult flies (figure 24). Neuronal degeneration was tested using the positive geotaxis/fly climbing paradigm that relies on healthy fly brain function and signal transduction along motor neurons, while micro CT scanning was utilised to directly image loss of adult flight muscle fibers within the thorax of flies at different timepoints. $Gal80^{ts}$ under the control of the tubulin

promoter was paired with the expression system to inhibit Gal4 expression at 18°C and thus allow healthy development of eggs and larvae at this temperature, with the inhibitor being degraded at 29°C – at which temperature the adult flies were kept from the time of ecclosure to initiate the tissue degeneration.

In climbing assays, w.t. Atx2 overexpression showed a progressive loss of fly climbing ability from day 1 post ecclosure to day 15. Similarly, the lack of the LSm domain in UAS-Atx2-ADARcd did not affect the observed climbing deficits and thus neuronal degeneration in this model. However, as shown in previous work from our lab, ΔCIDR Atx2-ADARcd alleviated the effects of the degeneration at day 15, with the majority of animals performing only slightly worse than the baseline at day 1 due to the inability of this mutant to form granules efficiently when overexpressed (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021). Interestingly, ΔPAM2 Atx2-ADARcd mimicked the lack of degeneration seen with the ΔCIDR form. The same result was seen in the muscle cell degeneration model, where w.t. Atx2-ADARcd leads to lesions in and loss of flight muscle fibers by day 20 post ecclosure, while ΔCIDR and ΔPAM2 protein expression does not induce such cell death (ΔLSm showed slightly increased level of degeneration compared to w.t. by day 20, in concordance with the known granule suppressing function of this domain). The PAM2 domain's modulation of degeneration however, must not be due to granule formation phenotypes but instead granule contents, localisation and identity, as ΔPAM2 Atx2 is still able to efficiently form granules in our cultured cell assays. As per our model discussed throughout, the PAM2 domain is necessary for the creation of the correct mRNPs and other Atx2 containing granules, with the internal microenvironments of which may, in disease contexts, facilitate the nucleation of pathologic aggregates. This data is consistent with previously published observations of the human Atxn2 PAM2 domain and human PABP involvement in modulating the severity of exogenous TDP43 protein disease model in flies (Kim *et al.*, 2014). However, where our work breaks new ground compared to this previous study is in examining endogenous fly Atx2-PABP interactions – not a completely exogenous system from humans where we have previously shown the CIDR of human Atxn2 being unable to form granules in our assay, human Atxn2 cDNA being unable to rescue Atx2 null lethality in flies, and human structured domain swapped fly Atx2 being unable to co-localise expected partner proteins to granules. Furthermore, we break down the specific activities of the Atx2 domains on our modelled cell degeneration. Both the CIDR and the PAM2 domains are necessary to drive the pathology in this model, but for different intrinsic reasons within the same disease processes.

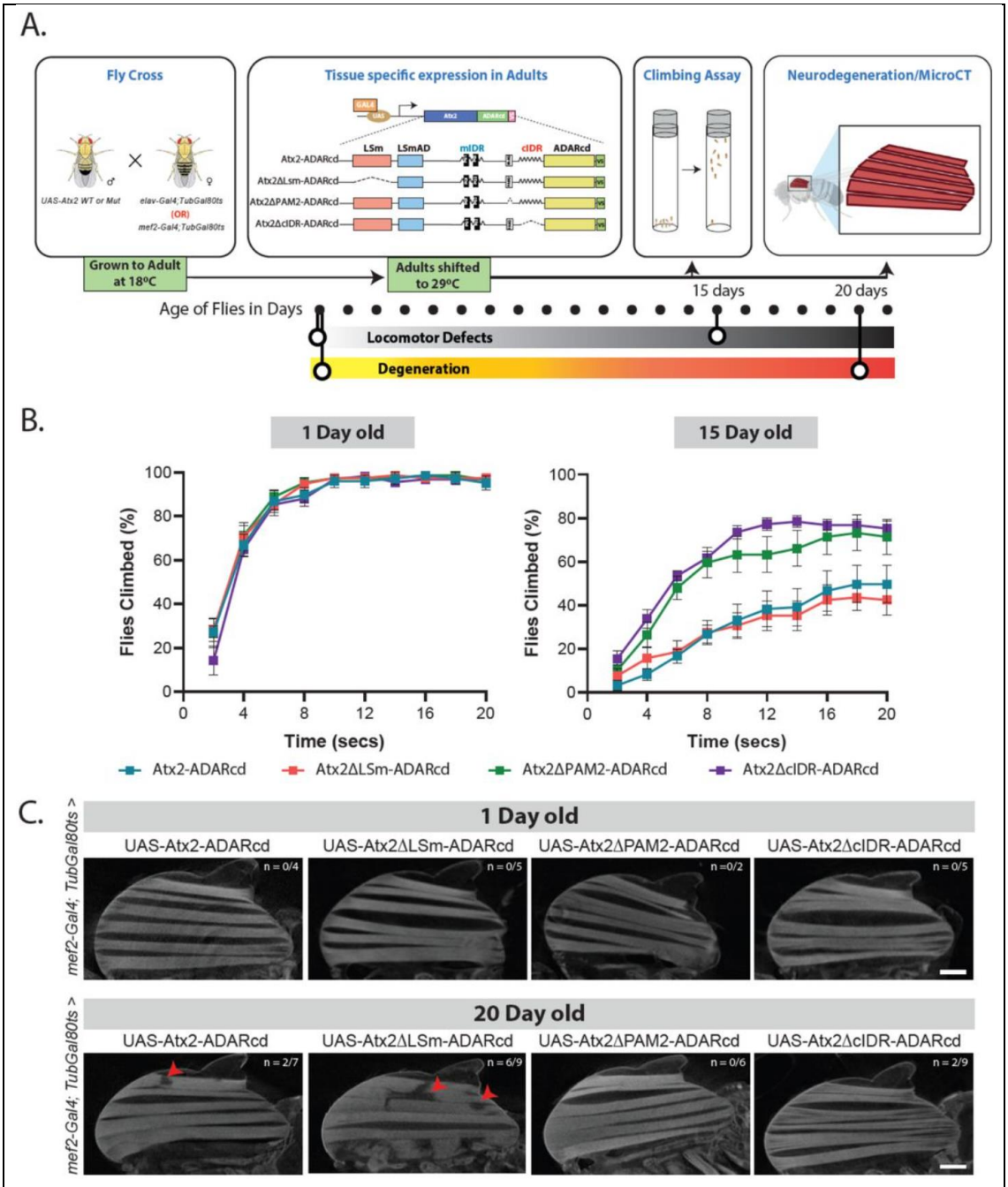

Figure 24: The IDR and PAM2 domains promote and the LSm domain inhibits neurodegeneration in Drosophila. (A) A schematic of the experimental design is shown. UAS-transgenes were crossed with elav-Gal4; tub-Gal80ts or mef2-Gal4, tub-Gal80ts and kept at 18°C till the adult flies emerged. The flies were shifted to 29 °C for days shown with dots under the experimental design. Fly climbing or indirect flight muscle cytotoxicity was studied. (B) Drosophila climbing behavior assay was performed by driving UAS-transgene (w.t. Atx2, Atx2ΔCIDR, Atx2ΔPAM or Atx2ΔLSm) with elav-Gal4. A graph was plotted with number of flies (Y-axis) that crossed the 20ml mark at a given time (X-axis). (C) Cellular toxicity was measured by driving UAS-transgene (w.t. Atx2, Atx2ΔCIDR, Atx2ΔPAM or Atx2ΔLSm) with

mef2-Gal4. Fly indirect flight muscles were imaged using micro-CT and the loss of muscle fibers is shown with solid red arrowheads. Figure and legend is taken directly from Petrauskas *et al.*, 2022, and the experiments were carried out by collaborators in NCBS Bangalore, and Baskar Bakthavachalu lab.

Conclusions and Outlook

The findings of my PhD research, publications and this thesis can be taken together to make a number of conclusions and inferences about the functions of Ataxin-2 domains and particularly how these functions complement or antagonise one another in normal cell function. Firstly, that there are sequential protein-protein interactions in the cytoplasm that Ataxin-2 is involved with and that in turn direct and control the different translational states of a particular mRNA target. Secondly, the findings highlight how the Ataxin-2 biomolecular condensate is made up by the generic aggregation ability of the IDRs as well as distinct structured interactions that positively and negatively modulate and direct this activity. These mRNP granules are necessary for both the normal function of the protein but also the pathways that promote degeneration models, and thus the antagonistic roles of the domains promote or protect against neurodegeneration, suggesting that potential therapeutics against the whole Ataxin-2 protein may have both downstream negative effects and also subtract the protein's potential protective effect from the LSm domain. Finally, this work identifies a specific, structured molecular mechanism relying on the PAM2-PABP interaction that contributes to the assembly of the correct mRNP granules and their function in cells.

Molecular mechanisms of Ataxin-2 function

Some RNA-binding proteins can remain associated with mRNAs across multiple stages: RNA processing, transport, translation, or translational control (Maniatis and Reed, 2002; Hachet and Ephrussi, 2004; Lin *et al.*, 2015; Harlen and Churchman, 2017; Formicola, Vijayakumar and Besse, 2019; Gomes and Shorter, 2019). Ataxin-2 is likely to be one such protein. It is a translational activator of the *Drosophila period* mRNA, a repressor of several different miRNAs and mRNAs (such as *CamKII*), a facilitator of neuronal mRNP-granule and stress-granule formation as well as being generally required for extending the lifetime of associated mRNAs (Nonhoff *et al.*, 2007; McCann *et al.*, 2011; Lim and Allada, 2013; Zhang *et al.*, 2013; Sudhakaran *et al.*, 2014; Yokoshi *et al.*, 2014; Bakthavachalu *et al.*, 2018; Inagaki *et al.*, 2020). These different functions could represent distinct Ataxin-2 complexes with specific activities regulated by the different target mRNAs it engages with. However, our work linking the recruitment of Atx2 to its mRNA targets via the PAM2 domain – PABP – polyA tail interaction suggests a different model of activity control that remains consistent with previous research. Sequential interactions mediated by different protein regions during

mRNP creation in the cytoplasm may allow Ataxin-2 to contribute in multiple ways to the translational control of a single mRNA – via both the promiscuous granule-forming ability of the IDR as well as distinct, potentially opposite acting protein binding partners mediated through the structured domains.

Previous work has shown that Atx-2 enhances *period* mRNA translation through a mechanism requiring LSm-domain interactions with a complex of LSM12 and TYF (Twenty Four) proteins associated with the 5'cap of the translating mRNA (Lim and Allada, 2013; Zhang *et al.*, 2013; Lee *et al.*, 2017). Given considerable supportive evidence for direct binding between the LSm-domain and LSM12, we postulate that LSm-domain-LSM12 interactions occur in translating polysomes (Satterfield and Pallanck, 2006) and contribute to increased efficiency of translation. This proposal is consistent with the observation that the LSm domain opposes the formation of mRNP granules, which usually contain translationally repressed mRNAs (Singh *et al.*, 2021).

However, the LSm domain must also contribute to LSM12-independent functions, because while LSm domain deletions from *Drosophila* Atx2 cause lethality and LSM12 null mutants, while arrhythmic, are viable and fertile (Lee *et al.*, 2017). Furthermore, KD of LSM12, contrary to what may be expected given the "granule-opposing" effect of the LSm domains in our assays, reduces the number and rate of assembly of SGs in oxidatively stressed cells, implying that LSM12 has a function in facilitating granule formation in this context (Lee *et al.*, 2020). One synthesis of the data is that LSm domains additionally contribute, perhaps indirectly, to interactions with the DEAD-box helicase Me31B/DDX6 in a translational repressor complex (Lee *et al.*, 2017; Brandmann *et al.*, 2018). Thus, we suggest that in the case of actively translating mRNAs, the Atx2 function is driven by LSm-domain association with LSM12 and translational initiators within a specific microenvironment, and that LSM12 disengages from a translational initiation complex as the mRNA transitions into a repressed state driven by Me31B, becoming a different cytosolic compartment with a new function.

While polyA tails and PABP are known to support translation and the Ataxin-2 PAM2 domain is involved in targeting the protein to polysomes (Satterfield and Pallanck, 2006), existing data do not directly address how Ataxin-2 PAM2 motif interactions contribute to translational activation. One possibility, supported by observations on the *period* mRNA is that the PAM2-domain guides Ataxin-2 to the 3'UTR of its target mRNAs (Lim and Allada, 2013). Our observation that PABP co-immunoprecipitates with mini Ataxin-2, show that the PAM2:PABP interactions occur independently of and

prior to mRNP granule formation. Recent findings that this association antagonizes the Ataxin-2 condensation (Boeynaems *et al.*, 2021) are consistent with a model in which the Ataxin-2 PAM2 domain interacts with PABP in translating mRNAs to support efficient translation driven by the LSm-LSM12 complex. However, in addition to supporting translation, PABP is also known to associate with translational repressors that could drive either mRNA deadenylation and/or storage (Machida *et al.*, 2018; Yoshida *et al.*, 2006). Our data support such a dual role for Ataxin-2 associated with PABP in translational repression. First, when Ataxin-2 target mRNAs are not actively translated, then the mRNP through Me31B/DDX6 and PABP may recruit deadenylases to transition into either a translationally dormant or degradative state (Machida *et al.*, 2018; Yoshida *et al.*, 2006; Lee *et al.*, 2017; Yi *et al.*, 2018). Second, Atx2 associated mRNAs may move into or transition to become part of mRNP granules whose formation is facilitated by Atx2 IDR-mediated condensation. We postulate that mRNAs in such assemblies are stored in a form that is protected from degradation. All throughout these activities of Ataxin-2, the IDRs could contribute to acting as crowding agents generating a microenvironment where disparate partner proteins are recruited to, bind and are replaced by one another in response to stimuli that drive the control of target RNAs throughout their lifespans. While the above model, shown in figure 25 is consistent with all our data, we acknowledge that it needs extensive and rigorous testing in the context of the life cycle of a single Ataxin-2 target mRNA.

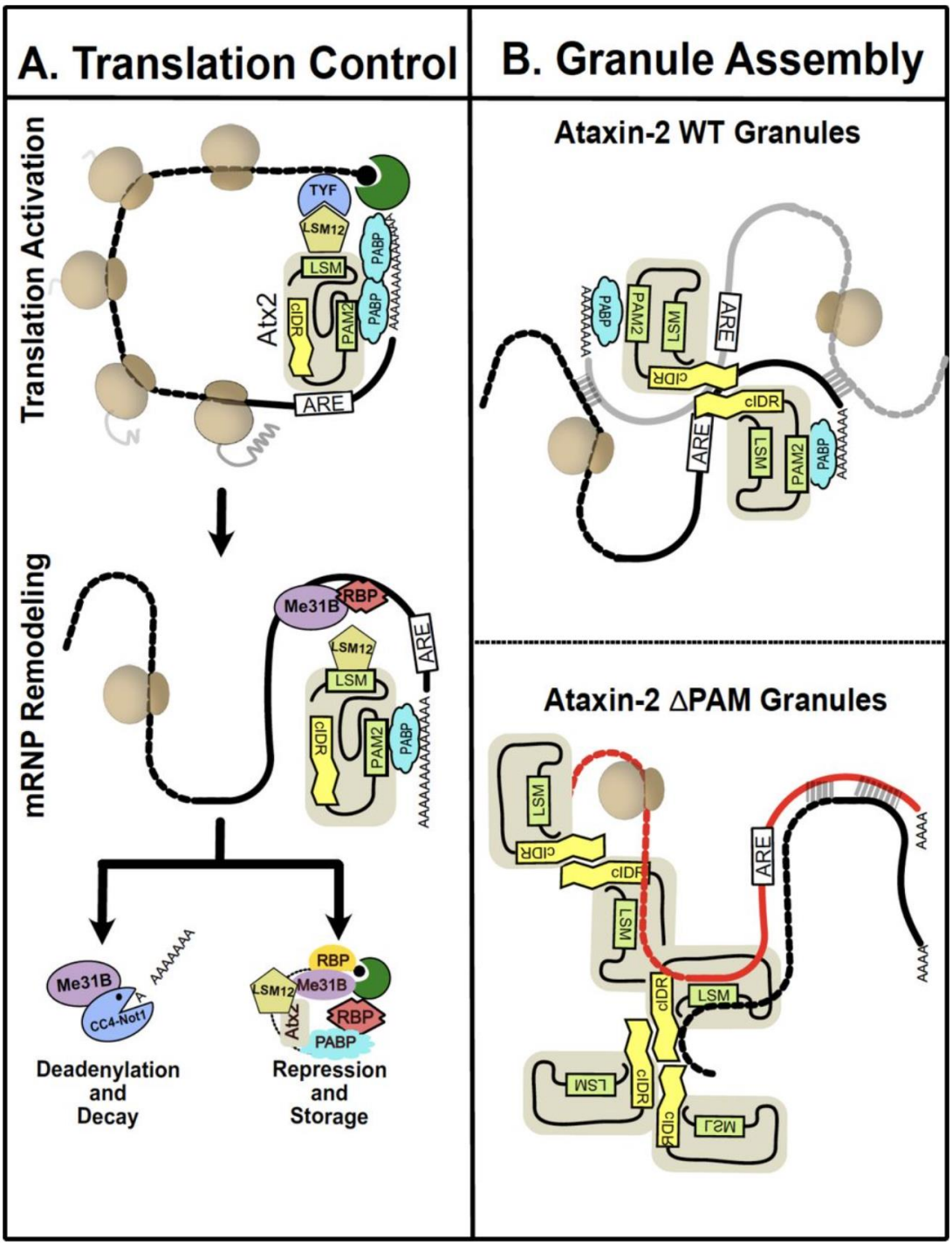

Figure 25: A model for Ataxin-2 RNP dynamics and the role of PAM2 domain in determining its RNP composition and mRNA selection.

(A) Ataxin-2 is recruited to mRNAs by RBPs during different stages of the mRNA life cycle. Ataxin-2 activates translation of subsets of mRNA by recruiting LSM12, TYF and other translation activation complexes. Under specific conditions, mRNP remodelling exposes Ataxin-2 cIDR that mediates multivalent interactions and RNP granule assembly. Ataxin-2 recruits Me31B and CCR4-NOT1 complexes that lead to deadenylating and/or translation repression. It is possible that LSM12/TYF continue

to associate with RNA but are probably not part of repressor complexes. RNA deadenylation can lead to degradation or translation repression and storage in RNP granules. (B) Ataxin-2-PAM2 domain determines protein and RNA partners of the RNP granules. PAM2 domain is essential for recruitment of Ataxin-2 to stress granules that also contains other RBPs (eg. Me31B, FMRP, Rox8, Rin and Caprin). Ataxin-2-cIDR along with RNA-RNA interaction stabilise the stress induces RNP condensation. In the absence of the PAM2 domain, Ataxin-2 fails to recruit specific target mRNA and proteins. Remodelling of Ataxin-2 exposes the cIDR to induce phase separation and aberrant RNP condensation. The Ataxin-2ΔPAM2 granules are non-toxic and lack several known stress granule proteins (eg.FMRP, Caprin and PABP). Figure and legend is taken directly from Petrauskas *et al.*, 2022, and was made in collaboration with Baskar Bakthavachalu, Joern Huelsmeier, Amanjot Singh and Mani Ramaswami.

Implications for Ataxin-2 as a therapeutic target

Antisense Oligonucleotide (ASO) based therapeutic strategies that lower levels of Atxn2 are being developed for the treatment of ALS and SCA2 (Becker *et al.*, 2017; Scoles *et al.*, 2017). These therapeutics aim at reducing the numbers, size or internal density of potentially pathogenic Atxn2 granules – modelled in cultured cell assays using the proxy of SGs – and more broadly aim to reduce the contribution of Atxn2 to the cells' capacity to form biological condensates. Our experiments provide a finer dissection of the activities of Ataxin-2, suggesting that the function of the LSm domain may be protective against granules, while IDR mediated assembly mechanisms alone are not sufficient to form the granules that drive disease model pathology in cells and animals. Structured interactions, specifically PAM2:PABP are required for the right type of biological condensate, phase separated microenvironment, or simply granule to form wherein the Ataxin-2 protein is able to contribute to a particular pathway of degeneration.

Our previous work showed that *Atx2* mutants lacking the cIDR required for Ataxin-2 granule formation in Drosophila neurons and S2R+ cells, were resistant to neurodegeneration as assessed in *Drosophila* disease models (Bakthavachalu *et al.*, 2018; Huelsmeier *et al.*, 2021). We further showed that the LSm-domain antagonizes Ataxin-2 granule formation (Singh *et al.*, 2021). Here we advance the latter observation

by demonstrating that Ataxin-2 forms lacking the LSm domain may more effectively cause cytotoxicity than the wild-type or IDR-deficient forms (figure 24, C). These observations independently confirm our original conclusions and provide further support for a model in which the efficiency of mRNP assembly correlates with the speed and severity of neurodegenerative processes in *Drosophila*.

The importance of the PAM2 domain in promoting degeneration has been previously observed by experiments showing that heterologous expression of a pathogenic form of human Ataxin-2 lacking its PAM2 domain, but not the full-length form, suppresses cytotoxicity in *Drosophila* expressing human TDP-43 (Kim *et al.*, 2014). Our observations that expression of Atx2ΔPAM2 is far less toxic than expression of wild-type Atx2 is consistent with this. In addition, by showing that Atx2ΔPAM2 forms compositionally different Ataxin-2 granules, we highlight the importance of specific granule components, and not granules *per se,* in neurodegenerative pathologies. Thus, while liquid-liquid transitions mediated by disordered domains could be a shared requirement for the formation of multiple types of mRNP granules, we speculate that each granule type, with distinctive composition, could preferentially support one or other type of proteinopathy (Vogler *et al.*, 2018; De Graeve and Bessé, 2018). The implication of these findings for the development of future therapeutics against Atxn2, such as the aforementioned ASOs, is significant. Generic KD of Atxn2 may prove insufficient at completely removing the right molecular conditions where pathogenic interactions between proteins may take place in a disease patient's cells. Furthermore, the specific, structured interaction between PAM2:PABP opens up as a potential therapeutic target that may have the capacity to be drugged with small molecules, thus possibly removing Atxn2 protein from the disease-driving condensate context.

Structured interactions may determine mRNP granule composition

Many lines of evidence argue that specific molecular interactions, such as mediated by structured domains of the P-body component Edc3 or the stress-granule components G3BP and Caprin, contribute to the mRNP granule formation (Decker, Teixeira and Parker, 2007; Kedersha *et al.*, 2016). In engineered systems, the condensation of RNA-binding proteins and mRNAs into granules has been clearly shown to depend on both traditional protein-protein interactions and on more promiscuous interactions between intrinsically disordered regions (Protter *et al.*, 2018). Our work now identifies the interactions between Ataxin-2's PAM2 motif and PABP as a critical contributor to

the assembly of Ataxin-2 containing mRNP granules. This suggests a mechanism by which the interaction helps select mRNA and protein components of mRNP granules. We thus put forward an interpretation that Ataxin-2, guided by PAM2:PABP interactions and LSm domain interactions, recruits target mRNAs and associated proteins into translating mRNPs (Satterfield and Pallanck, 2006). Under conditions where the translation is arrested, LSm-domain interactions are altered (Lee *et al.*, 2017), and transcripts are released from stalled ribosomes. Base-pairing interactions between exposed mRNA side chains, as well as interactions between Ataxin-2's now more accessible intrinsically disordered regions, contribute to the assembly of these mRNPs into granules. This logical sequence of events is consistent with: (a) TRIBE data showing a reduced number of edits of native Ataxin-2 target mRNAs by Atx2ΔPAM2-ADARcd; (b) the inability of ΔPAM2-miniAtx2 constructs to associate with stress granules; and (c) the aberrant protein composition of granules induced by Atx2ΔPAM2 in S2 cells. The additional observation that Atx2ΔPAM2-ADARcd expression results in a large number of non-native mRNA edits, indicates that the PAM2:PABP interaction not only selects correct target mRNAs but also prevents Ataxin-2 engagement with incorrect mRNA target regions.

Outlook

Our conclusion that Ataxin-2-PAM2:PABP interactions are involved in the selection of mRNA components of RNP granules is superficially inconsistent with the observation that RNA components of native stress granules can be predicted with remarkable accuracy on the basis of mRNA size. This argues for a primary role for RNA-RNA interactions in the stress granule assembly (Jain and Vale, 2017; Van Treeck *et al.*, 2018; Matheny *et al.*, 2021). However, we note that experiments presented here do not address mechanisms by which mRNAs are selected into stress granules. Instead, the TRIBE data address how Atx2-target mRNAs are selected into neuronal mRNP granules that exist in non-stressed cells *in vivo*, and microscopic studies analyse protein components of mRNP granules formed following Atx2 expression in S2 cells. Our experiments and observations therefore point to fundamental differences in mechanisms by which the assembly of neuronal granules, or granule types found in unstressed cells, may differ from those involved in stress-granule assembly. The regulation and composition of the former class could well rely extensively on specific protein-protein and protein-mRNA interactions, which may be revealed by future analyses of mechanisms by which such mRNP assemblies are formed *in vivo*. Furthermore, this model fits well with contemporary interpretations on the role of phase

separation as an organising principle *in vivo*. Notably, the older idea of homotypic interactions driving the assembly and contents of membraneleless organelles, puncta, compartments or even any multi-protein complexes has given way to an understanding that both promiscuous and specific interactions add up and that neither type alone is sufficient or necessary for the correct self-assembly of the aforementioned structures (Korkmazhan, Tompa and Dunn, 2021). Our findings regarding Ataxin-2 in the distinct granule types studied, while generalisable to some extent, also highlight this additive model in the few observations of content differences between, for example, neuronal granules and SGs (and thus implied differences in the process of the granules' assembly).

The question of the precise nature and variety of granules or condensates in which Ataxin-2 is involved in remains. Notably, the PAM2:PABP interaction adds a potentially very broad set of mRNAs as partners of Ataxin-2, however, only a subset of all polyadenylated mRNAs are seen by us as hits in our TRIBE analysis. In light of the emergent understanding of the heterogeneity found among mRNPs and granules previously classed as performing uniform functions, there could be a layering of broad selectivity by various proteins in an interaction network that overlap in distinct ways to confer high levels of selectivity and specificity of contents through the intersection of their otherwise broad selectivities. If such a hypothesis is tested further through well designed experiments, the precise granule where Ataxin-2 may modulate pathways that lead to proteinopathies can be identified. SGs, which are commonly suggested as being nucleation sites for aggregate initiation may be better off seen as proxies for the various activities being examined by specific experiments. SGs are not always found in affected cells of patients or in the affected cells of animal models of those same diseases where they are proposed to play a role, which is even true for SCA2 where Ataxin-2 is shown not just to modulate degeneration through granules but to form aggregates itself (Huynh *et al.*, 1999, 2003). However, we are yet to successfully image distinct mRNPs or granules containing Atx2 in healthy or disease model *Drosophila* neurons despite all the evidence suggesting they should exist. This issue needs to be resolved in the future, potentially through the use of super-resolution or single-molecule imaging or through the design of advanced reporters that may boost the signal from any mRNPs that form.

Real-time single-molecule imaging of Ataxin-2 and mutant behaviour during granule formation also has the potential to embellish the model itself with new insights on the mechanisms of action. Particularly, questions of how Ataxin-2 transitions from one

complex and state to another, as well as what happens after protein aggregation remain. In the case of the latter, two non-exclusive competing theories exist and could be tested implicating aggregation causing cellular arrest and subsequent death due to resource sequestration in granules and the aggregates themselves, or alternatively a break-down of stress signalling through a stress granule feedforward loop that leads to the same cellular death pathways (Nihei, Ito and Suzuki, 2012; Markmiller *et al.*, 2018; Wolozin and Ivanov, 2019; Hetz, Zhang and Kaufman, 2020).

Work in ALS or SCA2 patient derived IPSCs could the effects of human Atxn2 CIDR on degeneration surrogate phenotypes and thus test the model itself in the closest system to human disease as is possible (Rowe and Daley, 2019). A reciprocal experiment testing PAM2 domain deletions in the same context could build the foundation for a future small molecule screen to find inhibitors of PAM2:PABP binding that might provide a pharmacological means to mimic this deletion and its potential neuroprotective effects. A structured protein domain usually undergoes a metabolic or binding function that is determined by the steric nature of that region – a physical interaction. And a physical interaction can be potentially blocked by a small molecule; an inhibitor/competitor can be developed to bind the active domain. For example, in our mini Ataxin-2 SG recruitment assays, we have a clear read-out of the effect of the PAM2 domain on this phenomenon by staining for the construct's presence in granules. There is potential to automate such an assay for rapid and extensive screening of small molecules that may block the PAM2 domain, thus giving a diffuse read-out. Small molecules with such a hypothetical activities may prove to be the inspiration for future human therapeutics for degenerative diseases where Ataxin-2 plays a role in promoting their pathology by being recruited to granules where its IDRs can create the right microenvironment for aggregate formation. A drug that subtracts the effects of the Atxn2 IDRs from these processes could potentially reduce the likelihood of downstream degeneration.

Altogether, the work I carried out as part of this thesis improves our understanding of the forces at work controlling the sometimes antagonistic functions of Ataxin-2 and its domains. We flesh-out our understanding of the Atx2 cIDR and its functions, showing that while such a domain is intrinsically multi-valent and promiscuous in its binding, this activity is tightly controlled in and utilised by cells, and that this is achieved through the structured domains and their specific binding partners. We propose a model whereby Ataxin-2 complexes are remodelled based on context during control of mRNAs at different stages of its lifespan, and how structured interactors may change to facilitate

opposite functions to the previous complex. Finally, our work sheds light on how these normal functions of Ataxin-2 are mis-utilised in the pathology of proteinopathies, modelled in *Drosophila.* Data from these experiments astonishingly show how the cIDR and the structured PAM2 domain both have the same effect on Atx2 mediated degeneration, but for completely different underlying reasons that contribute to the same pathologic context. For cellular degeneration in the models we tested to take place, not only IDR generated granules are necessary but also these granules must have the specific composition determined by the PAM2 domain in order to act as crucibles for aggregation. All being said, I will take the final words of my thesis to reflect on the topic with which I started – the human cost of neurodegenerative disease and the urgent need for effective therapeutics. While we sincerely hope for it, it remains to be seen if current trials for Atxn2 ASOs will be successful in translating findings from the lab to better patient outcomes. But it is also my hope that this work may prove useful for informing our current understanding of Ataxin-2 role in proteinopathies as well as suggesting more targeted therapeutic approaches to be investigated, such as screening for small molecules that may disrupt the PAM2:PABP interaction.

Supplementary Information

Primer list

Name	Direction	T_m	Description	Sequence
AP1	rev	59.9	Plan B 3' Gibson context fragment amplification rev	TCGCATCTTCATCTCCCTAGGTTTTTGATT
AP2	for	59.4	Plan B 3' Gibson context fragment amplification for	ACCACCAACAGCAGTTGTAAGACAC
AP3	for	59.9	Plan B 5' Gibson context fragment amplification for	GAGCGGTCCGTTATCGATGGC
AP4	rev	61.8	Plan B 5' Gibson context fragment amplification rev	GCTGCGGCCGCTGAGC
AP5	for	60.7	Gibson homology addition: overlap 5' hum Atxn2 long cDNA, addition 5' fly UTR context	CACATCCAATCGACGTTATTATGCGCTCAGCGGCCG
AP6	rev	59.6	Gibson homology addition: overlap 5' fly UTR context, addition 5' Atxn2 long cDNA	GCTGCGGCCGCTGAGCGCATAATAACGTCGATTGGATGTGGATCGC
AP7	for	60	Gibson fragment forward for fly 5' UTR context (CRISPR scenario)	ACTGCTCCTTCTGGCCTCCC
AP8	for	65.1	Atxn2 cDNA amplification primer	ATGCGCTCAGCGGCCGCA

			for, Also a hum Atxn2 cDNA seq primer (1)	
AP9	rev	64.3	Atxn2 cDNA amplification primer rev, Also a hum Atxn2 cDNA seq primer (8)	TTACAACTGCTGTTGGTGGTGGGCTTGT
AP10	rev	59.4	Gibson homology addition: overlap 3' hum Atxn2 long cDNA, addition 3' fly UTR context	ATTCCCCAGCTGGCGGTGTCTTACAACTGCTGTTGGTGGTGGG
AP11	for	59.8	Gibson homology addition: overlap 3' fly UTR context, addition 3' Atxn2 long cDNA	ACCACCAACAGCAGTTGTAAGACACCGCCAGCTGGGG
AP12	rev	60.5	Gibson fragment forward rev fly 3' UTR context (CRISPR scenario)	CGTAACGTAGCTCCCACTCGCT
AP13	rev	59.4	Hum Atxn2 cDNA seq primer (1.5), behind poly-Q	CAGTCCTTTGTTACTGTTTCGACCTCTG
AP14	for	59.4	Hum Atxn2 cDNA seq primer (2)	CAGAGGTCGAAACAGTAACAAAGGACTG
AP15	for	56.5	Hum Atxn2 cDNA seq primer (3)	GTGTAGTGTCTACGTATGATAGCAGTTT
AP16	for	60.7	Hum Atxn2 cDNA seq primer (4)	CGGGCAGCCACCCCTACA
AP17	for	59.9	Hum Atxn2 cDNA seq primer (5)	TAGTGAGGCTAAAGATTCCAGGCTTCAAG
AP18	for	59.6	Hum Atxn2 cDNA seq primer (6)	GGAAATCAACATTGAATCCCAATGCAAAGG

AP19	for	60.6	Hum Atxn2 cDNA seq primer (7)	ATGGCACCACCAACACACGC
AP20	for	~62.0 See web site	CRISPR sgRNA template DNA 5' Packman Atx2 forward for SP6 in vitro transcription http://www.taejoonlab.org/index.php/CRISPR_protocol	ATTTAGGTGACACTATAGTTTGTGTCCGATTGCCGGGGTTTTAGAGCTAGAAATAGC
AP21	for	~62.0 See web site	CRISPR sgRNA template DNA 3' Packman Atx2 forward for SP6 in vitro transcription	ATTTAGGTGACACTATAAACGTAGCTCCCACTCGCTAGTTTTAGAGCTAGAAATAGC
AP22	rev	~62.0 See web site	CRISPR sgRNA TEMPLATE DNA universal reverse primer (includes tracrRNA complement sequence)	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAAC
AP23	rev	59.3	Hum Atxn2 cDNA seq primer (9) - revised	CTTCAGCGACATGGTGAGGGG
AP24	for	59.6	Hum Atxn2 cDNA seq primer (10) - revised	GCTGCCAATGTCCGCAAGC
AP25	for	59.4	[Plan B] 3' fragment with Gibson homology to 3' of Human Atxn2 (no stop codon) and primer for sfGFP tag from Baskar's HR Atx2 constructs	ACCACCAACAGCAGTTGgaagtgcataccaatcaggacccg

AP26	rev	59.3	[Plan B] 3' fragment with Gibson homology to fly Atx2 3' UTR and a primer for sfGFP (that ends with a V5 tag and stop codon)	CCCCAGCTGGCGGTGTCtcaGGGTGCTATCCAGGCCCAGC
AP27	for	~60.0	[Plan B] Fly 3' context forward with a stop codon at the start	TgAGACACCGCCAGCTGGGG
AP28	rev	~61.0	Modified (non-stop codon) Atxn2 cDNA amplification primer rev - a version of AP9, Also a hum Atxn2 cDNA seq primer (8.5) . USE ME INSTEAD OF AP9	CAACTGCTGTTGGTGGTGGGCTTGT
AP29	for	60	Hum Atxn2 cDNA seq primer (11) - revised, "overflow" 5'	CCCAGCAGCACAACAGACTGTC
AP30	rev	60.2	Hum Atxn2 cDNA seq primer (12) - revised, "overflow"	GGAGAAGGAGGACGACGAAGGG
AP31	for	~60	Atxn2 cDNA amplification primer forward - adjusted for point mutation from sequencing, Also a hum Atxn2 cDNA seq primer (1.5.2) Use me instead of AP8	TGCGGCCGCAGAGCGCAT

AP32	rev	60.7	Reverse compliment of ap16, for 2 part amplification of human cDNA for Gibson assembly	TGTAGGGGTGGCTGCCCG
AP33	rev	59.3	[plan C] Like AP26, but with gibson homology to pUASattb backbone digested with XbaI and a stop codon before the homology	cttcacaaagatccttcaGGTGCTATCCAGGCCCAGC
AP34	for	~60	[plan C] Same as Ap31 but with gibson homology to pUASattB digested with EcoRI	gaatagggaattgggATGCGCTCTGCGGCCG
AP35	for	~60	[plan C] Same as Ap31 but with gibson homology to pUASattB digested with XbaI	ggctcgagggtacctATGCGCTCTGCGGCCG
AP36	for	60.5	[Fly cDNA] at the start forward	atgaacaacaatagcaagcggaaaacc
AP37	rev	59.5	[Fly cDNA] at the end reverse	tcactgtggctgatgctgctg
AP38	rev	60.1	[Fly cDNA] at the end reverse non-stop	ctgtggctgatgctgctgca
AP39	for	59.5	[Fly cDNA] middle forward	caaaatagcaactcgccgccg
AP40	rev	59.5	[Fly cDNA] middle reverse	cggcggcgagttgctattttg
AP41	for	60.5	[Fly cDNA] gibson with pUAST 5' XbaI overhang	aacagatctgcggccgcggctcgagggtacctatgaacaacaatagcaagcggaaaaccc

AP42	rev	59.5	[Fly cDNA] gibson with pUAST 3' XbaI overhang (stopped)	acagaagtaaggttccttcacaaagatcct tcactgtggctgatgctgctg
AP43	rev	60.5	[Fly cDNA] at the start reverse for seq	GGGTTTTCCGCTTGCTATTGT TGTTCAT
AP44	for	59.5	[Fly cDNA] at the end forward for seq	cagcagcatcagccacagtga
AP45	rev	60.1	[Fly cDNA] at the end reverse but with 3x Flag, stop and gibson homology to pUASt 3' XbaI	acagaagtaaggttccttcacaaagatcct tcaCTTATCATCATCGTCCTTG TAATCCTTGTCATCGTCGTCC TTGTAGTCTTTATCATCATCA TCTTTGTAGTCGCCCGATCCA CCctgtggctgatgctgctgca
AP46	for	60.4	[Fly cDNA] [snap] SNAP forward but with gibson homology to unstoppered 3' of Fly cDNA	cagatcgtgatgcagcagcatcagccac agATGGACAAAGACTGCGAAA TGAAGCG
AP47	rev	61.7	[Fly cDNA] [snap] SNAP reverse but with stop and gibson homology to pUASt 3' Xbal	acagaagtaaggttccttcacaaagatcct tcaACCCAGCCCAGGCTTGCC
AP48	rev	60	[pUASt] amplify the vector at XbaI cut but provide overlap to 5' of synthesised human cDNA fragment 1 Q23!!!	CGAGCGTGGGGCGGCGGCG GCGGAGCGCATctagaggtaccct cgagccgc
AP49	for	59.3	[pUASt] amplify the vector at XbaI cut but provide overlap to 3' of synthesised human cDNA fragment 2 while	TGTGCAGGCCCATCACCAGC AGCAGCTGAAaGGTGGATCG GGCGACTACAAAGATGATGA TGATAAAGACTACAAGGACG ACGATGACAAGGATTACAAG

			including 3x Flag tags plus stop codon	GACGATGATGATAAGTGAaggatctttgtgaaggaaccttacttctgt
AP50	for	59.3	[pUASt] amplify the vector at XbaI cut but provide overlap to 3' of synthesised human cDNA fragment 2 with stop but untagged	AGTGTGCAGGCCCATCACCAGCAGCAGCTGtgaaggatctttgtgaaggaaccttacttctgt
AP51	for	60.6	[pUASt] [planB] Forward primer for 5' fragment of pUASt that starts at AatII digestion site (to give gibson homology to digested plasmid)	catttccccgaaaagtgccacctg
AP52	for	58.4	[human cDNA] [snap] SNAP forward but with gibson homology to unstoppered frag 2 of hum. cDNA	AGTGTGCAGGCCCATCACCAGCAGCAGCTGaaaATGGACAAAGACTGCGAAATGAAGC
AP53	rev	59.6	[pAcman hum planB] human frag 1 Q23 homology, primer for 5' context	CGAGCGTGGGGCGGCGGCGGCGGAGCGCATAATAACGTCGATTGGATGTGGATCGC
AP54	for	61	[pAcman hum planB] human frag 2 homology, primer for 3' context with stop	AGTGTGCAGGCCCATCACCAGCAGCAGCTGtgaGACACCGCCAGCTGGGGA
AP55	for	61	[pAcman hum PlanB] human frag 2 homology, primer for	TGTGCAGGCCCATCACCAGCAGCAGCTGAAaGGTGGATCGGGCGACTACAAAGATGATGA

			3' context with 3x flag and stop	TGATAAAGACTACAAGGACG ACGATGACAAGGATTACAAG GACGATGATGATAAGtgaGAC ACCGCCAGCTGGGGA
AP56	rev	59.6	[pAcman fly cDNA planB] fly 5' homology, primer for 5' context	tcgggttttccgcttgctattgttgttcatAAT AACGTCGATTGGATGTGGAT CGC
AP57	for	61	[pAcman fly cDNA planB] fly 3' homology, primer for 3' context with stop	atcgtgatgcagcagcatcagccacagtg aGACACCGCCAGCTGGGGA
AP58	for	61	[pAcman fly cDNA planB] fly 3' homology, primer for 3' context with 3x flag and stop	cagatcgtgatgcagcagcatcagccac agGGTGGATCGGGCGACTAC AAAGATGATGATGATAAAGAC TACAAGGACGACGATGACAA GGATTACAAGGACGATGATG ATAAGtgaGACACCGCCAGCT GGGGA
AP59	rev	61.7	[pAcman] [snap] SNAP reverse but with stop and gibson homology to fly 3' context (the 5' of the 3' context)	CCCGTGCGCCATTCCCCAGC TGGCGGTGTCtcaACCCAGCC CAGGCTTGCC
AP60	for	59.3	[sfGFP] forward but with gibson homology to 3' of human cDNA fragment 2 non-stop	AGTGTGCAGGCCCATCACCA GCAGCAGCTGGAATTCATGG TGTCCAAGGGCGA
AP61	for	59.3	[sfGFP] forward but with gibson homology to 3' of fly cDNA non-stop	cagatcgtgatgcagcagcatcagccac agGAATTCATGGTGTCCAAGG GCGA

AP62	rev	59.3	[sfGFP] reverse but with stop and gibson homology to pUASt digested with XbaI	acagaagtaaggttccttcacaaagatcct tcaCTCGAGCTTGTACAGCTC ATCCATG
AP63	rev	60	[pUASt] amplify the vector at XbaI cut but provide overlap to 5' of synthesised human cDNA fragment 1 Q80!!!	GGAGCGTGGGGCGGCGGCG GCACTGCGCATctagaggtaccctc gagccgc
AP64	rev	59.6	[pAcman hum planB] human frag 1 Q80 homology, primer for 5' context	GGAGCGTGGGGCGGCGGCG GCACTGCGCATAATAACGTC GATTGGATGTGGATCGC
AP65	for	57.5	[pUASt] Forward for downstream of XbaI digestion with no overlaps	aggatctttgtgaaggaaccttacttct
AP66	rev	60.5	[Part 2] [human cDNA] regular reverse primer for part 2 synthesised	CAGCTGCTGCTGGTGATGGG
AP67	for	71.7	[Part 2] [human cDNA] Gib adapter primer for synthesised Q23 frag1	TCGCTGCCGCCACGCGCCG CCACCCCAACCCGCCCACCA AGCCGCCCGCCC
AP68	for	71.7	[Part 2] [human cDNA] Gib adapter primer for synthesised Q80 frag1	AGCCTGCCCCCACGCGCCG CCACACCCACGCGCCCACCA AGCCGCCCGCCC
AP69	for	60.5	[Seq] Part 1 Q23 forward for confim PCR of two-part assembly	TATGGACAGCTCCTACGCCA AGC

AP70	for	60.8	[Seq] Part 1 Q80 forward for confim PCR of two-part assembly	TATGGATTCGTCCTACGCCAAGCG
AP71	rev	60	[Seq] Part 2 reverse for confim PCR of two-part assembly	CTTATTGCCGGCTGGCGAATTCT
AP72	rev	70.7	[Part 2] [human cDNA] regular reverse primer for part 2 synthesised, alt version of AP66 with higher Tm	CAGCTGCTGCTGGTGATGGGCCTGCACACT
AP73	rev	61	Hum Atxn2 cDNA seq primer (13) - revised, "overflow" with 143 buffer sequence	GCGGCGGAGGGATACGGT
AP74	for	61.2	Hum Atxn2 cDNA seq primer (14) - revised, "overflow" with 66 buffer sequence	GCACTACAGCCCATTCCAGTCTCG
AP75	rev	61.4	fly Atx2 cdna seq 1 - rev	CCAAAACAGTCTGAGGCGGAGGG
AP76	rev	59	fly Atx2 cdna seq 2 - rev	TTGTTACTGTTTCGACCTCTGCCC
AP77	rev	60	fly Atx2 cdna seq 3 - rev	TGGTTTGCCCTTGCTTCCCG
AP78	rev	59.7	fly Atx2 cdna seq 4 - rev	GAGGCCCTTCTGAAGACATGCG
AP79	rev	50	fly Atx2 cdna seq 5 - rev	TAGTAGTTGATCCATAGATTCAGA
AP80	rev	61.1	fly Atx2 cdna seq 6 - rev	GCTTGATTCACTGGCATGGGCG

AP81	rev	61.6	fly Atx2 cdna seq 7 - rev	GGCCTGGTGCTGATGGTGCT
AP82	for	60.2	fly Atx2 cdna seq 8 - for	GCCCTCGCTCAAAGTGCACTAC
AP83	for	56.4	crispr integration cassette confirm #1 - for	agaagcggttttcgggagtag
AP84	rev	56.7	crispr integration cassette confirm #2 - rev	ccgacatgacacaaggggttta
AP85	for	59	10x FLAG cassette confirm - for	Ggactacaaggacgacgatgacaag
AP86	for	59.9	HALO cassette confirm - for	GGCAGAAATCGGTACTGGCTTTCC
AP87	rev	56.4	crispr integration cassette confirm #1 prime - rev	ctactcccgaaaaccgcttct
AP88	for	56.7	crispr integration cassette confirm #2 prime - for	taaacccttgtgtcatgtcgg
AP89	for	61	5' human Atxn2 CIDR + 5' AscI adapter - from pUASt hum Atxn2 w.t.	ggcgcgccaTTCTCCCAGCCCAAGCCATCG
AP90	rev	60.2	3' human Atxn2 CIDR + 3' AvrII adapter - from pUASt hum Atxn2 w.t.	cctaggCAGCTGCTGCTGGTGATGGG
AP91	for	66.8	5' fly Atx2 CIDR + 5' AscI adapter - from pUC19 CIDR only	ggcgcgccaCCGGCGGCCAACCAGCCGAT
AP92	rev	65.3	3' fly Atx2 CIDR + 3' AvrII adapter - from pUC19 CIDR only	cctaggCTGTGGCTGATGCTGCTGCATCACGA

AP93	for	60	In Atx2 locus, 5', reading into the Atx2 UTR from ~1kb away	actggaaatgtcaacggcaccg
AP94	rev	60	In Atx2 locus, 3', reading into the Atx2 UTR from ~0.5kb away	agaagaggcggagaagtggcaa
AP95	for	60	In Atx2 locus, 5', reading into the Atx2 UTR from ~100bp away	agtgcgcaggcgataaccattc
Ap96	rev	60.4	In Atx2 locus, 3', reading into the Atx2 UTR from ~100bp away	tcgacttacaggacttgggcgg
Ap97	for	48.9	pET28a - hMini Atxn2/ dPAM2	AGCTTGCGGCCGCACatgAATATGCGAATGGTACA
Ap98	for	53.8	pET28a - hPAM2	AGCTTGCGGCCGCACatgGTCCGAAAGTCCACGCT
Ap99	for	49.6	pET28a - hFUS	AGCTTGCGGCCGCACatggcctcaaacgattatac
Ap100	rev	53.2	SNAP - pET28a	tggtggtggtggtgctcTCCCAGTCCGGGCTT
Ap101	for	52.2	hFUS - SNAP	aaccgcagcagtggtATGGATAAGGACTGCGAGAT
Ap102	rev	51.1	hMini Atxn2 - hFUS	atcgtttgaggccatGCTAAACGACCGTGGAT
Ap103	rev	50.3	h dPAM2 Mini Atxn2 - hFUS	atcgtttgaggccatGGAGTTCCTCTGCACC
Ap104	na	55	Mycoplasma test 1	GGGAGCAAACACGATAGATACCCT
Ap105	rev	52.6	hPAM2 (flex) - hFUS	atcgtttgaggccatcctagggtcgacG
Ap106	for	55.8	hFUS for	atggcctcaaacgattataccca
Ap107	rev	55.9	hFUS rev no stop codon	accactgctgcggttgta

Ap108	for	52.2	SNAP for	ATGGATAAGGACTGCGAGAT
Ap109	na	55	Mycoplasma test 2	TGCACCATCTGTCACTCTGTTAACCTC
Ap110	for	60	#1 oli1, #2 oli1 -- Takara cloning oligos	TCGCGGATCCGAATTATGAATATGCGAATGGTACACATTC
Ap111	rev	60	#1 oli2 (final), #2 oli2 (final)	GACGGAGCTCGAATTTCCCAGTCCGGGCTTTCC
Ap112	for	60	#3 oli1	ATGGGTCGCGGATCCGAATTATGGCCTCAAACGATTATACCC
Ap113	rev	60	#3 oli2, #4 oli4, #5 oli4, #6 oli2	CCTTATCCATACCACTGCTGCGGTTGTAAC
Ap114	for	60	#3 oli3, #4 oli5, #5 oli5, #6 oli3	CAGCAGTGGTATGGATAAGGACTGCGAGATGAAGC
Ap115	rev	60	#3 oli4, #4 oli6, #5 oli6, #6 oli4, #7 oli2	TTGTCGACGGAGCTCGAATTTCCCAGTCCGGGCTTTCC
Ap116	for	60	#4 oli1, #5 oli1	ATGGGTCGCGGATCCGAATTATGAATATGCGAATGGTACACATTC
Ap117	rev	60	#4 oli2	TTGAGGCCATGCTAAACGACCGTGGATTAAACTCC
Ap118	for	60	#4 oli3	GTCGTTTAGCATGGCCTCAAACGATTATACCC
Ap119	rev	60	#5 oli2	TTGAGGCCATGGAGTTCCTCTGCACCGC
Ap120	for	60	#5 oli3	GAGGAACTCCATGGCCTCAAACGATTATACCC
Ap121	for	60	#6 oli1	ATGGGTCGCGGATCCGAATTATGGTCCGAAAGTCCACGCTGAATCCCAACGCTAAGGAGTTTAATCCACGGTCGTTTAGCATGGCCTCAAACGATTATACCC
Ap122	for	60	#7 oli1	ATGGGTCGCGGATCCGAATTATGGTCCGAAAGTCCACGCTGAATCCCAACGCTAAGGAGT

				TTAATCCACGGTCGTTTAGCATGGATAAGGACTGCGAGATGAAGC
Ap123	for	52	Snap for with GATC overhang	GATCATGGATAAGGACTGCGAGAT
Ap124	rev	53.4	Snap rev with CTAG overhang	CTAGTTATCCCAGTCCGGGCT
Ap125	for	52.1	Hum mini for with GATC overhang	GATCATGAATATGCGAATGGTACACA
Ap126	for	53.7	Human cDNA for with GATC overhang	GATCATGCGCTCCGCCG
Ap127	for	53.5	Hum lsm-AD for with GATC overhang	GATCatgtatggagttgtctcgacc
Ap128	for	60	pCMV3.1 seq for	GGCGTGGATAGCGGTTTGACTC
Ap129	rev	60	pCMV3.1 seq rev	AGCTGGTTCTTTCCGCCTCAGA
Ap130	rev	59.9	Hum Atx2 genomic insertion 5' rev	CCATTTCCATTGGAGCCGAAGCA
Ap131	rev	59.8	Fly mini Atx2 genomic insertion 5' rev	GGCGATATCGAAGTTGCCGGAAA
Ap132	rev	60.3	Fly cDNA Atx2 genomic insertion 5' rev	ctgcaccagccttgttgttagtgg
Ap133	for	60.3	SNAP tag genomic insertion 3' for	TGCCCATTCTGATTCCCTGCCA
Ap134	for	59.6	Atx2 HR arm 5' locus for	caggctgggaatttcactaggtgc
Ap135	rev	59.8	Atx2 HR arm 5' locus rev	ccgaatgtcaagataggcggcc
Ap136	for	60.1	Atx2 HR arm 3' locus for (AP94 is 3' locus rev)	ccagaagcaaagcgacagcga

Ap137	for	60	Atx2 locus 5' UTR for (AP12 is 3' UTR rev)	GTAGCAGCCGAAGCAAGAGAGC
Ap138	rev	59.8	Integration Cassette 5' end rev	GTCGCTCCGTAGACGAAGCG
Ap139	for	60.6	Integration Cassette 3' end for	CCACACCTCCCCCTGAACCTG
AP140	rev	60.1	Atx2 locus 3' UTR rev (replace AP12, use with AP137)	AGCCGCGCGTTTACTCAAAAGT
AP141	for	59.9	Serrate locus 406-427 FWD 1	GCGCGCCTGGGTTACACTAAAT
AP142	rev	60	Serrate locus 1057-1078 REV 1	CGGAGCTCACAGATTCTGCTGC
AP143	for	60	Human Atxn2 replacement cassette 5'ORF (LSM domain) for	CAACATGCGCATGGTGCACATC
AP144	rev	60	Human Atxn2 replacement cassette 5'ORF (LSM domain) rev	ATGTCATTGGGgTCCCAGCCAT
AP145	for	60.9	Hum Atxn2 rep cass 5' For overlap Att+Met for PCR for pJet	atctctagaATGCGCTCCGCCG
AP146	for	61.5	Mini Atx2 rep cass 5' For overlap Met for PCR for pJet	aATGCATTCCGCCACAGCCCT
USE AP36	for	60.5	Fly Atx2 rep cass 5' For overlap Att+Met for PCR for pJet (AP36)	
AP147	rev	61.2	General SNAP rep cass 3' Rev with	GTTTATCCCAGTCCGGGCTTTCCC

			stop cod for PCR for pJet	
AP148	rev	~50	mini adar gibb frag rev 35mer	attggtgaaggggcggccgATCCATcctaggcgc
AP149	for	~50	mini not1 adar gibb frag for 35mer	atcctttacttcaggcggccATGCATTCCGCCACA
AP150	for	~50	mini bgl2 adar gibb frag for 35mer	aagagaactctgaatagatcATGCATTCCGCCACA
AP151	rev	~50	adjusted mini adar gibb frag rev 35mer	caccattggtgaaggggcgATCCATcctaggcgc
AP152	for	~50	adjusted mini not1 adar gibb frag for 35mer	tcttatcctttacttcaggcATGCATTCCGCCACA
AP153	for	~50	adjusted mini bgl2 adar gibb frag for 35mer	caagaagagaactctgaataATGCATTCCGCCACA
AP154	rev	60	cut adapter Not1 rev	ggccgCCtaggcgcgccTGAGC
AP155	for	60	cut adapter Bgl2 for	gatcATGCATTCCGCCACAGCCC
AP156	for	60	cut adapter Not1 for	ggccATGCATTCCGCCACAGCCC
AP157	for	60	sequencing mini forward	ATGCATTCCGCCACAGCCC
AP158	rev	60	sequencing rev ADAR start	gcccgctaatacctttcgacgc
AP159	for	50.5	ADAR plasmid NOTI adapters	GCGGCCGCTccttcaccaatggtggt
AP160	rev	50.9	ADAR plasmid NOTI adapters - rev in mhc intron	gcggccgcctgaagtaaaggataagaattaggg
AP161	for	60.2	Mini Atx2 NOTI adapters	gcggccgcATGCATTCCGCCACAGCCC
AP162	rev	60	Mini Atx2 NOTI adapters	gcggccgccctaggcgcgccTGAGC

AP163	rev	50.6	ADAR plasmid NotI adapter, rev HSP70 promot, replace 160	gcggccgccagagttctcttcttgtattcaat
AP164	for	60	for Mini Atx2 primer in LSM, bglII digest + spacer	ttactcAGATCTATGCATTCCGCCACAGCCC
AP165	rev	60	rev Mini Atx2 primer in spacer, notI digest + spacer	AGTTATgcggccgccctaggcgcgccTGAGC
AP166	rev	60	blunt rev mini Atx2	cctaggcgcgccTGAGC
AP167	for	60	blunt for in LSM of mini Atx2	ATGCATTCCGCCACAGCCC
AP168	rev	60	blunt rev dPAM mini Atx2	cctagggtcgacGGAGCCG
AP169	rev	60	rev Mini dPAM2 Atx2 primer in spacer, notI digest + spacer	AGTTATgcggccgccctagggtcgacGGAGCCG
AP170	for	60	screening common mini dLSM-AD for	TGGACCTGGAGAATGGCGACG
AP171	rev	60	screening common Adar rev	TGATTGCACTGAAGGTCCAGCTGT
AP172	rev	60.3	screening common LSM outwards rev	CAGCTCCAGGGCGATATCGAAGT

Publications and Presentations

Research paper – Bakthavachalu B, Huelsmeier J, Sudhakaran IP, Hillebrand J, Singh A, **Petrauskas A**, Thiagarajan D, Sankaranarayanan M, Mizoue L, Anderson EN, Pandey UB, Ross E, VijayRaghavan K, Parker R, Ramaswami M. RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration. Neuron. 2018 May 16;98(4):754-766.e4. doi: 10.1016/j.neuron.2018.04.032. PMID: 29772202.

Poster presentation – Neurofly 2018, Krakow, Poland

Presentation – Zurich Image and Data Analysis School 2019, ETH Zurich, Switzerland

Poster Presentation – Wellcome Campus “Molecular Neurodegeneration” winter school 2019, Cambridge, England

Research paper – Singh A, Hulsmeier J, Kandi AR, Pothapragada SS, Hillebrand J, **Petrauskas A**, Agrawal K, Rt K, Thiagarajan D, Jayaprakashappa D, VijayRaghavan K, Ramaswami M, Bakthavachalu B. Antagonistic roles for Ataxin-2 structured and disordered domains in RNP condensation. Elife. 2021 Mar 10;10:e60326. doi: 10.7554/eLife.60326. PMID: 33689682; PMCID: PMC7946432.

Research paper (in review) – **Petrauskas A**, Fortunati DL, Singh A, Kandi AR, Pothapragada SS, Agrawal K, Huelsmeier J, Hillebrand J, Brown G, Chaturvedi D, Lee J, Lim C, Auburger G, K. VijayRaghavan, Ramaswami M, Bakthavachalu B. Structured and disordered regions of Ataxin-2 contribute differently to the specificity and efficiency of mRNP granule formation. bioRxiv 2022.02.15.480566; doi: https://doi.org/10.1101/2022.02.15.480566

Short Talk – Neurofly 2022, Saint Malo, France

Bibliography

Abouward, R. and Schiavo, G. (2020) ‘Walking the line: mechanisms underlying directional mRNA transport and localisation in neurons and beyond’, *Cellular and Molecular Life Sciences 2020 78:6*. Springer, 78(6), pp. 2665–2681. doi: 10.1007/S00018-020-03724-3.

Alberti, S. *et al.* (2017) ‘Granulostasis: Protein quality control of RNP granules’, *Frontiers in Molecular Neuroscience*. Frontiers Research Foundation, 10. doi: 10.3389/FNMOL.2017.00084/FULL.

Alberti, S., Gladfelter, A. and Mittag, T. (2019) ‘Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates.’, *Cell*, 176(3), pp. 419–434. doi: 10.1016/j.cell.2018.12.035.

Albrecht, M. and Lengauer, T. (2004) ‘Survey on the PABC recognition motif PAM2’, *Biochemical and Biophysical Research Communications*, 316(1), pp. 129–138. doi: 10.1016/j.bbrc.2004.02.024.

Auburger, G. *et al.* (2017) ‘Efficient Prevention of Neurodegenerative Diseases by Depletion of Starvation Response Factor Ataxin-2’, *Trends in Neurosciences*, 40(8), pp. 507–516. doi: 10.1016/j.tins.2017.06.004.

Babinchak, W. M. and Surewicz, W. K. (2020) ‘Liquid–Liquid Phase Separation and Its Mechanistic Role in Pathological Protein Aggregation’, *Journal of Molecular Biology*. Academic Press, 432(7), pp. 1910–1925. doi: 10.1016/J.JMB.2020.03.004.

Bakthavachalu, B. *et al.* (2018) ‘RNP-Granule Assembly via Ataxin-2 Disordered Domains Is Required for Long-Term Memory and Neurodegeneration’, *Neuron*. Cell Press, 98(4), pp. 754-766.e4. doi: 10.1016/J.NEURON.2018.04.032.

Banani, S. F. *et al.* (2017) ‘Biomolecular condensates: organizers of cellular biochemistry’, *Nature Reviews Molecular Cell Biology 2017 18:5*. Nature Publishing Group, 18(5), pp. 285–298. doi: 10.1038/nrm.2017.7.

Becker, L. A. *et al.* (2017) ‘Therapeutic reduction of ataxin-2 extends lifespan and reduces pathology in TDP-43 mice’, *Nature*. Nature Publishing Group, 544(7650), pp. 367–371. doi: 10.1038/nature22038.

Biology, M. and Fundaci, F. C. (2014) ‘Stable s2 protocols’, (May). doi: 10.1007/978-1-59745-257-1.

Boeynaems, S. *et al.* (no date) ‘Poly (A)-binding protein is an ataxin-2 chaperone that emulsifies biomolecular condensates’, *biorxiv.org*. doi: 10.1101/2021.08.23.457426.

Brandmann, T. *et al.* (2018) ‘ Molecular architecture of LSM 14 interactions involved in the assembly of mRNA silencing complexes ’, *The EMBO Journal*. EMBO, 37(7). doi:

10.15252/EMBJ.201797869.
Buchan, J. R., Muhlrad, D. and Parker, R. (2008) ‘P bodies promote stress granule assembly in Saccharomyces cerevisiae’, *Journal of Cell Biology*, 183(3), pp. 441–455. doi: 10.1083/JCB.200807043.
Ceriani, M. F. (2007) ‘Basic Protocols for Drosophila S2 Cell Line’, in, pp. 415–422. doi: 10.1007/978-1-59745-257-1_33.
Chen, X., Zaro, J. L. and Shen, W.-C. (2013) ‘Fusion protein linkers: property, design and functionality.’, *Advanced drug delivery reviews*. NIH Public Access, 65(10), pp. 1357–69. doi: 10.1016/j.addr.2012.09.039.
Decker, C. J., Teixeira, D. and Parker, R. (2007) ‘Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae’, *Journal of Cell Biology*, 179(3), pp. 437–449. doi: 10.1083/JCB.200704147.
Deo, R. C. *et al.* (1999) ‘Recognition of polyadenylate RNA by the poly(A)-binding protein’, *Cell*. Elsevier B.V., 98(6), pp. 835–845. doi: 10.1016/S0092-8674(00)81517-2.
Doyle, M. and Kiebler, M. A. (2011) ‘Mechanisms of dendritic mRNA transport and its role in synaptic tagging’, *The EMBO Journal*. John Wiley & Sons, Ltd, 30(17), pp. 3540–3552. doi: 10.1038/EMBOJ.2011.278.
Drost, J. *et al.* (2013) ‘Ataxin-2 modulates the levels of Grb2 and Src but not ras signaling’, *Journal of Molecular Neuroscience*, 51(1), pp. 68–81. doi: 10.1007/S12031-012-9949-4.
Elden, A. C. *et al.* (2010) ‘Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS.’, *Nature*. NIH Public Access, 466(7310), pp. 1069–75. doi: 10.1038/nature09320.
Fittschen, M. *et al.* (2015) ‘Genetic ablation of ataxin-2 increases several global translation factors in their transcript abundance but decreases translation rate’, *Neurogenetics*. Springer Verlag, 16(3), pp. 181–192. Available at: https://link.springer.com/article/10.1007/s10048-015-0441-5 (Accessed: 22 March 2022).
FlyBase Gene Report: Dmel\Atx2 (no date). Available at: http://flybase.org/reports/FBgn0041188 (Accessed: 31 August 2022).
Formicola, N., Vijayakumar, J. and Besse, F. (2019) ‘Neuronal ribonucleoprotein granules: Dynamic sensors of localized signals’, *Traffic*. Blackwell Munksgaard, 20(9), pp. 639–649. doi: 10.1111/TRA.12672/.
Franks, T. M. and Lykke-Andersen, J. (2008) ‘The Control of mRNA Decapping and P-Body Formation’, *Molecular cell*. NIH Public Access, 32(5), p. 605. doi: 10.1016/J.MOLCEL.2008.11.001.
Gan, L. *et al.* (2018) ‘Converging pathways in neurodegeneration, from genetics to

mechanisms', *Nature Neuroscience*. Springer US, 21(October). doi: 10.1038/s41593-018-0237-7.
Gebauer, F. *et al.* (2020) 'RNA-binding proteins in human genetic disease', *Nature Reviews Genetics 2020 22:3*. Nature Publishing Group, 22(3), pp. 185–198. doi: 10.1038/s41576-020-00302-y.
Glenner, G. G. and Wong, C. W. (1984) 'Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein', *Biochemical and biophysical research communications*. Biochem Biophys Res Commun, 120(3), pp. 885–890. doi: 10.1016/S0006-291X(84)80190-4.
Gomes, E. and Shorter, J. (2019) 'The molecular language of membraneless organelles', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 294(18), pp. 7115–7127. doi: 10.1074/JBC.TM118.001192.
De Graeve, F. and Bessé, F. (2018) 'Neuronal RNP granules: From physiological to pathological assemblies', *Biological Chemistry*. Walter de Gruyter GmbH, 399(7), pp. 623–635. doi: 10.1515/HSZ-2018-0141/HTML.
Groth, A. C. *et al.* (2000) 'A phage integrase directs efficient site-specific integration in human cells.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 97(11), pp. 5995–6000. doi: 10.1073/pnas.090527097.
Hachet, O. and Ephrussi, A. (2004) 'Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization', *Nature*, 428(6986), pp. 959–963. doi: 10.1038/NATURE02521.
Harlen, K. M. and Churchman, L. S. (2017) 'The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(4), pp. 263–273. doi: 10.1038/NRM.2017.10.
Hart, M. P. *et al.* (2012) 'Distinct TDP-43 pathology in ALS patients with ataxin 2 intermediate-length polyQ expansions.', *Acta neuropathologica*. NIH Public Access, 124(2), pp. 221–30. doi: 10.1007/s00401-012-0985-5.
He, W. and Parker, R. (2000) 'Functions of Lsm proteins in mRNA degradation and splicing.', *Current opinion in cell biology*, 12(3), pp. 346–50. doi: 10.1016/S0955-0674(00)00098-3.
Hentze, M. W. *et al.* (2018) 'A brave new world of RNA-binding proteins', *Nature Reviews Molecular Cell Biology 2018 19:5*. Nature Publishing Group, 19(5), pp. 327–341. doi: 10.1038/nrm.2017.130.
Hetz, C., Zhang, K. and Kaufman, R. J. (2020) 'Mechanisms, regulation and functions of the unfolded protein response', *Nature Reviews Molecular Cell Biology*. Nature

Research, 21(8), pp. 421–438. doi: 10.1038/S41580-020-0250-Z.
Van den Heuvel, D. M. A. *et al.* (2014) 'Taking a risk: a therapeutic focus on ataxin-2 in amyotrophic lateral sclerosis?', *Trends in Molecular Medicine*. Elsevier, 20(1), pp. 25–35. doi: 10.1016/j.molmed.2013.09.001.
Hirokawa, N. (2006) 'mRNA Transport in Dendrites: RNA Granules, Motors, and Tracks', *The Journal of Neuroscience*. Society for Neuroscience, 26(27), p. 7139. doi: 10.1523/JNEUROSCI.1821-06.2006.
Huelsmeier, J. *et al.* (2021) 'A C-terminal ataxin-2 disordered region promotes Huntingtin protein aggregation and neurodegeneration in Drosophila models of Huntington's disease', *G3: Genes|Genomes|Genetics*. Oxford University Press, 11(12). doi: 10.1093/G3JOURNAL/JKAB355.
Hughes, M. P. *et al.* (2018) 'Atomic structures of low-complexity protein segments reveal kinked b sheets that assemble networks', *Science*, 359(6376), pp. 698–701. doi: 10.1126/science.aan6398.
Huynh, D. P. *et al.* (1999) 'Expression of Ataxin-2 in Brains from Normal Individuals and Patients with Alzheimer's Disease and Spinocerebellar Ataxia 2'.
Huynh, D. P. *et al.* (2003) 'Expansion of the polyQ repeat in ataxin-2 alters its Golgi localization, disrupts the Golgi complex and causes cell death', *Human Molecular Genetics*, 12(13), pp. 1485–1496. Available at: https://academic.oup.com/hmg/article/12/13/1485/554900 (Accessed: 19 March 2019).
Imbert, G. *et al.* (1996) 'Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats.', *Nature genetics*. Presse Dienstleistungsgesellschaft mbH und Co. KG, 14(3), pp. 285–91. doi: 10.1038/ng1196-285.
Inagaki, H. *et al.* (2020) 'Direct evidence that ataxin-2 is a translational activator mediating cytoplasmic polyadenylation', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 295(47), pp. 15810–18525. doi: 10.1074/JBC.RA120.013835.
Jain, A. and Vale, R. D. (2017) 'RNA phase transitions in repeat expansion disorders', *Nature*. Nature Publishing Group, 546(7657), pp. 243–247. doi: 10.1038/NATURE22386.
Jiménez-López, D. and Guzmán, P. (2014) 'Insights into the evolution and domain structure of ataxin-2 proteins across eukaryotes', *BMC Research Notes*, 7(1), p. 453. doi: 10.1186/1756-0500-7-453.
Kaehler, C. *et al.* (2012) 'Ataxin-2-like is a regulator of stress granules and processing bodies.', *PloS one*. Edited by J. Ma. Public Library of Science, 7(11), p. e50134. doi: 10.1371/journal.pone.0050134.

Kametani, F. and Hasegawa, M. (2018) 'Reconsideration of Amyloid Hypothesis and Tau Hypothesis in Alzheimer's Disease', *Frontiers in Neuroscience*. Frontiers Media SA, 12(JAN). doi: 10.3389/FNINS.2018.00025.
Karran, E. and Strooper, B. (no date) 'The amyloid hypothesis in Alzheimer disease: new insights from new therapeutics'. doi: 10.1038/s41573-022-00391-w.
Kedersha, N. *et al.* (2005) 'Stress granules and processing bodies are dynamically linked sites of mRNP remodeling.', *The Journal of cell biology*. The Rockefeller University Press, 169(6), pp. 871–84. doi: 10.1083/jcb.200502088.
Kedersha, N. *et al.* (2016) 'G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits.', *The Journal of cell biology*, 212(7), pp. 845–60. doi: 10.1083/jcb.201508028.
Kedersha, N. and Anderson, P. (2007) *Mammalian Stress Granules and Processing Bodies*, *Methods in Enzymology*. Academic Press. doi: 10.1016/S0076-6879(07)31005-7.
Kedersha, N., Ivanov, P. and Anderson, P. (2013) 'Stress granules and cell signaling: more than just a passing phase?', *Trends in Biochemical Sciences*. Elsevier Current Trends, 38(10), pp. 494–506. doi: 10.1016/J.TIBS.2013.07.004.
Kedersha, N. L. *et al.* (1999) 'RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules.', *The Journal of cell biology*. The Rockefeller University Press, 147(7), pp. 1431–42. doi: 10.1083/jcb.147.7.1431.
Key, J. *et al.* (2020) 'Mid-Gestation lethality of Atxn2l-Ablated Mice', *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI), 21(14), pp. 1–24. doi: 10.3390/IJMS21145124.
Khong, A. *et al.* (2017) 'The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules', *Molecular Cell*. Cell Press, 68(4), pp. 808-820.e5. doi: 10.1016/J.MOLCEL.2017.10.015.
Khong, A. and Parker, R. (2020) 'The landscape of eukaryotic mRNPs.', *RNA (New York, N.Y.)*. Cold Spring Harbor Laboratory Press, 26(3), pp. 229–239. doi: 10.1261/rna.073601.119.
Kiehl, T. R., Shibata, H. and Pulst, S. M. (2000) 'The ortholog of human ataxin-2 is essential for early embryonic patterning in C. elegans', *Journal of Molecular Neuroscience*, 15(3), pp. 231–241. doi: 10.1385/JMN:15:3:231.
Kim, H.-J. *et al.* (2014) 'Therapeutic modulation of eIF2α phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models.', *Nature genetics*. Nature Publishing Group, 46(2), pp. 152–60. doi: 10.1038/ng.2853.
Kim, H. J. and Taylor, J. P. (2017) 'Lost in Transportation: Nucleocytoplasmic

Transport Defects in ALS and Other Neurodegenerative Diseases', *Neuron*. Cell Press, 96(2), pp. 285–297. doi: 10.1016/J.NEURON.2017.07.029.
Korkmazhan, E., Tompa, P. and Dunn, A. R. (2021) 'The role of ordered cooperative assembly in biomolecular condensates', *Nature Reviews Molecular Cell Biology 2021 22:10*. Nature Publishing Group, 22(10), pp. 647–648. doi: 10.1038/s41580-021-00408-z.
Kozlov, G. *et al.* (2001) 'Structure and function of the C-terminal PABC domain of human poly(A)-binding protein'. National Academy of Sciences, 98(8), pp. 4409–4413. doi: 10.1073/pnas.071024998.
Kufel, J. *et al.* (2003) 'Lsm Proteins are required for normal processing and stability of ribosomal RNAs.', *The Journal of biological chemistry*, 278(4), pp. 2147–56. doi: 10.1074/jbc.M208856200.
Lastres-Becker, I. *et al.* (2008) 'Insulin receptor and lipid metabolism pathology in ataxin-2 knock-out mice', *Human Molecular Genetics*, 17(10), pp. 1465–1481. doi: 10.1093/HMG/DDN035.
Lastres-Becker, I. *et al.* (2016) 'Mammalian ataxin-2 modulates translation control at the pre-initiation complex via PI3K/mTOR and is induced by starvation.', *Biochimica et biophysica acta*, 1862(9), pp. 1558–69. doi: 10.1016/j.bbadis.2016.05.017.
Lee Id, J. *et al.* (2020) 'LSM12-EPAC1 defines a neuroprotective pathway that sustains the nucleocytoplasmic RAN gradient'. doi: 10.1371/journal.pbio.3001002.
Lee, J. *et al.* (2017) 'LSM12 and ME31B/DDX6 Define Distinct Modes of Posttranscriptional Regulation by ATAXIN-2 Protein Complex in Drosophila Circadian Pacemaker Neurons', *Molecular Cell*. Cell Press, 66(1), pp. 129-140.e7. doi: 10.1016/J.MOLCEL.2017.03.004.
Liao, Y.-C. *et al.* (2019) 'RNA Granules Hitchhike on Lysosomes for Long-Distance Transport, Using Annexin A11 as a Molecular Tether', *Cell*. Elsevier, 179(1), pp. 147-164.e20. doi: 10.1016/j.cell.2019.08.050.
Lim, C. and Allada, R. (2013) 'ATAXIN-2 activates PERIOD translation to sustain circadian rhythms in Drosophila.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 340(6134), pp. 875–9. doi: 10.1126/science.1234785.
Lin, Y. *et al.* (2015) 'Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins.', *Molecular cell*. Howard Hughes Medical Institute, 60(2), pp. 208–19. doi: 10.1016/j.molcel.2015.08.018.
Liu, J. *et al.* (2009) 'Deranged calcium signaling and neurodegeneration in spinocerebellar ataxia type 2.', *The Journal of neuroscience : the official journal of the Society for Neuroscience*. NIH Public Access, 29(29), pp. 9148–62. doi: 10.1523/JNEUROSCI.0660-09.2009.

van de Loo, S. *et al.* (2009) ‘Ataxin-2 associates with rough endoplasmic reticulum’, *Experimental Neurology*. Academic Press, 215(1), pp. 110–118. doi: 10.1016/J.EXPNEUROL.2008.09.020.
Machida, K. *et al.* (no date) ‘Dynamic interaction of poly (A)-binding protein with the ribosome’, *nature.com*. Available at: https://www.nature.com/articles/s41598-018-35753-1 (Accessed: 30 September 2022).
Maniatis, T. and Reed, R. (2002) ‘An extensive network of coupling among gene expression machines’, *Nature*, 416(6880), pp. 499–506. doi: 10.1038/416499A.
Markmiller, S. *et al.* (2018) ‘Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules’, *Cell*. Cell Press, 172(3), pp. 590-604.e13. doi: 10.1016/J.CELL.2017.12.032.
Mateju, D. *et al.* (2017) ‘An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function’, *The EMBO Journal*, 36(12), pp. 1669–1687. doi: 10.15252/embj.201695957.
Matheny, T. *et al.* (2021) ‘RNA partitioning into stress granules is based on the summation of multiple interactions.’, *RNA (New York, N.Y.)*. Cold Spring Harbor Laboratory Press, 27(2), pp. 174–189. doi: 10.1261/rna.078204.120.
McCann, C. *et al.* (2011) ‘The Ataxin-2 protein is required for microRNA function and synapse-specific long-term olfactory habituation.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 108(36), pp. E655-62. doi: 10.1073/pnas.1107198108.
McMahon, A. C. *et al.* (2016) ‘TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins’, *Cell*. Cell Press, 165(3), pp. 742–753. Available at: http://www.cell.com/article/S009286741630263X/fulltext (Accessed: 29 September 2022).
Moon, S. L. *et al.* (2019) ‘Multicolour single-molecule tracking of mRNA interactions with RNP granules’, *Nature Cell Biology*, 21(2), pp. 162–168. doi: 10.1038/s41556-018-0263-4.
Nakayama, K. *et al.* (2017) ‘RNG105/caprin1, an RNA granule protein for dendritic mRNA localization, is essential for long-term memory formation’, *eLife*, 6(5030732). doi: 10.7554/eLife.29677.
Neuwald, A. F. and Koonin, E. V (1998) ‘Ataxin-2, global regulators of bacterial gene expression, and spliceosomal snRNP proteins share a conserved domain.’, *Journal of molecular medicine (Berlin, Germany)*, 76(1), pp. 3–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9462862 (Accessed: 13 May 2019).
Ng, H., Pulst, S.-M. and Huynh, D. P. (2007) ‘Ataxin-2 mediated cell death is dependent on domains downstream of the polyQ repeat’, *Experimental Neurology*,

208(2), pp. 207–215. doi: 10.1016/j.expneurol.2007.07.023.
Nihei, Y., Ito, D. and Suzuki, N. (2012) 'Roles of ataxin-2 in pathological cascades mediated by TAR DNA-binding protein 43 (TDP-43) and fused in sarcoma (FUS)', *Journal of Biological Chemistry*, 287(49), pp. 41310–41323. doi: 10.1074/JBC.M112.398099.
Nonhoff, U. *et al.* (2007) 'Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules.', *Molecular biology of the cell*. Edited by J. Weissman, 18(4), pp. 1385–96. doi: 10.1091/mbc.e06-12-1120.
Nonis, D. *et al.* (2008) 'Ataxin-2 associates with the endocytosis complex and affects EGF receptor trafficking', *Cellular Signalling*, 20(10), pp. 1725–1739. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18602463 (Accessed: 13 May 2019).
Patel, A. *et al.* (2015) 'A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation', *Cell*. Elsevier Inc., 162(5), pp. 1066–1077. doi: 10.1016/j.cell.2015.07.047.
Paul, S. *et al.* (2018) 'Staufen1 links RNA stress granules and autophagy in a model of neurodegeneration', *Nature Communications*. Nature Publishing Group, 9(1), p. 3648. Available at: http://www.nature.com/articles/s41467-018-06041-3 (Accessed: 3 July 2019).
Petrauskas, A. *et al.* (2022) 'Structured and disordered regions of Ataxin-2 contribute differently to the specificity and efficiency of mRNP granule formation', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2022.02.15.480566. doi: 10.1101/2022.02.15.480566.
Protter, D. S. W. *et al.* (2018) 'Intrinsically Disordered Regions Can Contribute Promiscuous Interactions to RNP Granule Assembly.', *Cell reports*, 22(6), pp. 1401–1412. doi: 10.1016/j.celrep.2018.01.036.
Protter, D. S. W. and Parker, R. (2016) 'Principles and Properties of Stress Granules', *Trends in Cell Biology*, 26(9), pp. 668–679. doi: 10.1016/J.TCB.2016.05.004.
Pulst, S.-M. M. *et al.* (1996) 'Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2.', *Nature genetics*. Presse Dienstleistungsgesellschaft mbH und Co. KG, 14(3), pp. 269–76. doi: 10.1038/ng1196-269.
Ralser, M. *et al.* (2005) 'Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways', *Human Molecular Genetics*, 14(19), pp. 2893–2909. doi: 10.1093/HMG/DDI321.
Ramanan, V. K. and Saykin, A. J. (2013) 'Pathways to neurodegeneration : mechanistic insights from GWAS in Alzheimer ' s disease , Parkinson ' s disease , and related disorders', 2(3), pp. 145–175.
Ramaswami, M., Taylor, J. P. and Parker, R. (2013) 'Altered ribostasis: RNA-protein

granules in degenerative disorders.', *Cell*. Elsevier, 154(4), pp. 727–36. doi: 10.1016/j.cell.2013.07.038.
Rowe, R. G. and Daley, G. Q. (2019) 'Induced pluripotent stem cells in disease modelling and drug discovery', *Nature Reviews Genetics*. Springer US. doi: 10.1038/s41576-019-0100-z.
Sanders, D. W. *et al.* (2020) 'Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization', *Cell*. Cell Press, 181(2), pp. 306-324.e28. doi: 10.1016/J.CELL.2020.03.050.
Sanpei, K. *et al.* (1996) 'Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT', *Nature Genetics*. Presse Dienstleistungsgesellschaft mbH und Co. KG, 14(3), pp. 277–284. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8896556 (Accessed: 12 May 2019).
Satterfield, T. F., Jackson, S. M. and Pallanck, L. J. (2002) 'A Drosophila Homolog of the Polyglutamine Disease Gene SCA2 Is a Dosage-Sensitive Regulator of Actin Filament Formation', *Genetics*. Oxford Academic, 162(4), pp. 1687–1702. doi: 10.1093/GENETICS/162.4.1687.
Satterfield, T. F. and Pallanck, L. J. (2006) 'Ataxin-2 and its Drosophila homolog, ATX2, physically assemble with polyribosomes', *Human Molecular Genetics*. Narnia, 15(16), pp. 2523–2532. doi: 10.1093/hmg/ddl173.
Scoles, D. R. *et al.* (2017) 'Antisense oligonucleotide therapy for spinocerebellar ataxia type 2', *Nature*. Nature Publishing Group, 544(7650), pp. 362–366. doi: 10.1038/NATURE22044.
Seidel, K. *et al.* (2017) 'On the distribution of intranuclear and cytoplasmic aggregates in the brainstem of patients with spinocerebellar ataxia type 2 and 3', *Brain pathology (Zurich, Switzerland)*. Brain Pathol, 27(3), pp. 345–355. doi: 10.1111/BPA.12412.
Sen, N. E. *et al.* (2016) 'Search for SCA2 blood RNA biomarkers highlights Ataxin-2 as strong modifier of the mitochondrial factor PINK1 levels', *Neurobiology of disease*. Neurobiol Dis, 96, pp. 115–126. doi: 10.1016/J.NBD.2016.09.002.
Shibata, H., Huynh, D. P. and Pulst, S. M. (2000) 'A novel protein with RNA-binding motifs interacts with ataxin-2.', *Human molecular genetics*, 9(9), pp. 1303–13. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10814712 (Accessed: 13 May 2019).
Singh, A. *et al.* (2021) 'Antagonistic roles for ataxin-2 structured and disordered domains in rnp condensation', *eLife*. eLife Sciences Publications Ltd, 10. doi: 10.7554/ELIFE.60326.
Strassnig, M. and Ganguli, M. (2005) 'About a Peculiar Disease of the Cerebral Cortex: Alzheimer's Original Case Revisited', *Psychiatry (Edgmont)*. Matrix Medical Communications, 2(9), p. 30. Available at: /pmc/articles/PMC2993534/ (Accessed: 31

August 2022).
Stubenvoll, M. D. *et al.* (2016) ‘ATX-2, the C. elegans Ortholog of Human Ataxin-2, Regulates Centrosome Size and Microtubule Dynamics’, *PLoS Genetics*. Public Library of Science, 12(9). doi: 10.1371/JOURNAL.PGEN.1006370.
Sudhakaran, I. P. *et al.* (2014) ‘FMRP and Ataxin-2 function together in long-term olfactory habituation and neuronal translational control.’, *Proceedings of the National Academy of Sciences of the United States of America*, 111(1), pp. E99–E108. doi: 10.1073/pnas.1309543111.
Swisher, K. D. and Parker, R. (2010) ‘Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1 on stress granules in Saccharomyces cerevisiae’, *PloS one*. PLoS One, 5(4). doi: 10.1371/JOURNAL.PONE.0010006.
Tartakoff, A. *et al.* (2022) ‘The dual nature of the nucleolus’, *Genes & Development*. Cold Spring Harbor Laboratory Press, 36(13–14), pp. 765–769. doi: 10.1101/GAD.349748.122.
Tillotson, R. *et al.* (2017) ‘Radically truncated MeCP2 rescues Rett syndrome-like neurological defects.’, *Nature*. Nature Publishing Group, 550(7676), pp. 398–401. doi: 10.1038/nature24058.
Van Treeck, B. *et al.* (2018) ‘RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 115(11), pp. 2734–2739.
Venken, K. J. T. *et al.* (2011) ‘MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes’, *Nature Methods*, 8(9), pp. 737–743. doi: 10.1038/nmeth.1662.
Verdile, V., De Paola, E. and Paronetto, M. P. (2019) ‘Aberrant Phase Transitions: Side Effects and Novel Therapeutic Strategies in Human Disease’, *Frontiers in genetics*. Front Genet, 10. doi: 10.3389/FGENE.2019.00173.
Vianna, M. C. B. *et al.* (2016) ‘Drosophila ataxin-2 gene encodes two differentially expressed isoforms and its function in larval fat body is crucial for development of peripheral tissues’, *FEBS Open Bio*, 6(11), pp. 1040–1053. doi: 10.1002/2211-5463.12124.
Vogler, T. *et al.* (no date) ‘TDP-43 and RNA form amyloid-like myo-granules in regenerating muscle’, *nature.com*. Available at: https://idp.nature.com/authorize/casa?redirect_uri=https://www.nature.com/articles/s41586-018-0665-2&casa_token=PHf_EPuyXvEAAAAA:XkJVrj_q17vns3pzJFv3-gwEKRqnD6XfJFI4TygiL7cSAcJkaaqlhMJyw8FeR0fK52-T2TMN4k47tdTYWw (Accessed: 30 September 2022).

Wegmann, S. *et al.* (2018) ‘Tau protein liquid-liquid phase separation can initiate tau aggregation.’, *The EMBO journal*, 37(7), p. e98049. doi: 10.15252/embj.201798049.
Wolozin, B. and Ivanov, P. (2019) ‘Stress granules and neurodegeneration’, *Nature Reviews Neuroscience*. Nature Publishing Group, 20(11), pp. 649–666. doi: 10.1038/S41583-019-0222-5.
Xie, J., Kozlov, G. and Gehring, K. (2014) ‘The “tale” of poly(A) binding protein: the MLLE domain and PAM2-containing proteins.’, *Biochimica et biophysica acta*, 1839(11), pp. 1062–8. doi: 10.1016/j.bbagrm.2014.08.001.
Yi, H. *et al.* (2018) ‘PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay’, *Molecular Cell*. Cell Press, 70(6), pp. 1081-1088.e5. doi: 10.1016/J.MOLCEL.2018.05.009.
Yokoshi, M. *et al.* (2014) ‘Direct binding of Ataxin-2 to distinct elements in 3’ UTRs promotes mRNA stability and protein expression.’, *Molecular cell*. Cell Press, 55(2), pp. 186–98. doi: 10.1016/j.molcel.2014.05.022.
Yoshida, M. *et al.* (2006) ‘Poly(A) binding protein (PABP) homeostasis is mediated by the stability of its inhibitor, Paip2’, *EMBO Journal*, 25(9), pp. 1934–1944. doi: 10.1038/SJ.EMBOJ.7601079.
Zhang, Y. *et al.* (2013) ‘A role for Drosophila ATX2 in activation of PER translation and circadian behavior’, *Science*. American Association for the Advancement of Science, 340(6134), pp. 879–882. doi: 10.1126/SCIENCE.1234746.

Appendix

Abstract

Ataxin-2 (ATXN2) is a gene implicated in spinocerebellar ataxia type II (SCA2), amyotrophic lateral sclerosis (ALS) and Parkinsonism. The encoded protein is a therapeutic target for ALS and related conditions. ATXN2 (or Atx2 in insects) can function in translational activation, translational repression, mRNA stability and in the assembly of mRNP-granules, a process mediated by intrinsically disordered regions (IDRs). Our work has shown that the C-terminal IDR of Atx2 is essential for both long-term habituation and degeneration model progression in flies. On the other hand, we subsequently show the structured LSm domain, which can help stimulate mRNA translation, antagonizes mRNP-granule assembly. We built on previous research through a series of experiments on *Drosophila* and human Ataxin-2 proteins to highlight a poly-A tail – polyA-binding protein (PABP) – Atx2 interaction driven localisation mechanism for the protein. Results of Targets of RNA-Binding Proteins Identified by Editing (TRIBE), co-localization and immunoprecipitation experiments indicate that the PABP interacting, PAM2 motif of Ataxin-2 may be a major determinant of the mRNA and protein content of Ataxin-2 mRNP granules. Transgenic experiments in Drosophila indicate that while the Atx2-LSm domain may protect against neurodegeneration, structured PAM2- and unstructured IDR- interactions both, through distinct modes of action, support Atx2-induced cytotoxicity. Examining radically truncated Atx2 constructs in cells and transgenic flies allowed to separate and independently assay the effects of structured and unstructured domains and their respective interactions. Taken together, the data lead to a proposal for how Ataxin-2 interactions are remodelled during translational control and how structured and non-structured interactions contribute differently to the specificity and efficiency of RNP granule condensation as well as to neurodegeneration, with powerful implications for future therapeutic approaches